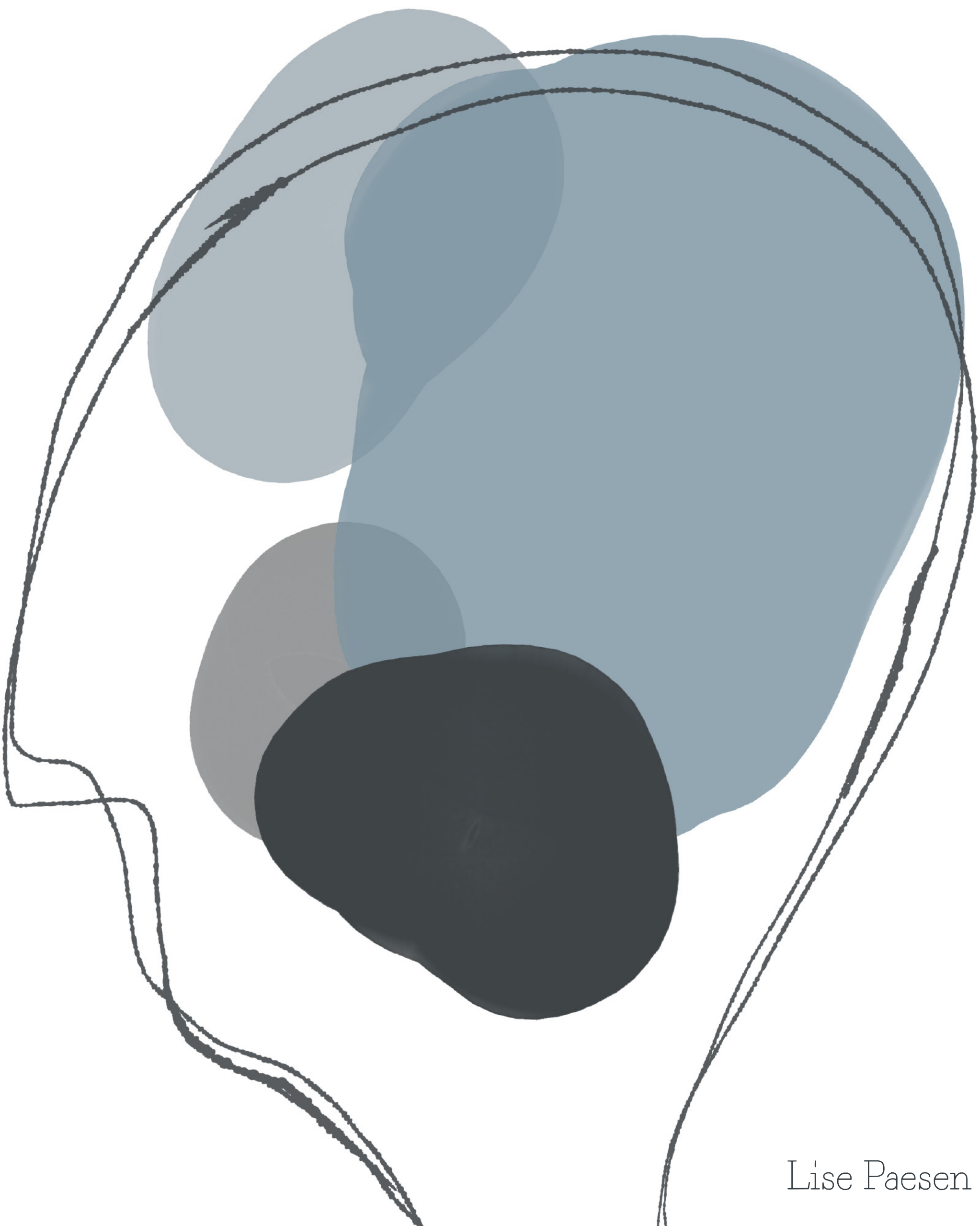
# THE STORY OF A LIFETIME

A linguistic approach to studying changes in spontaneous written language due to ageing



Lise Paesen

# THE STORY OF A LIFETIME

A linguistic approach to studying changes in spontaneous written language due to ageing

Lise Paesen

Universiteit Antwerpen

# THE STORY OF A LIFETIME

A linguistic approach to studying changes in spontaneous written language due to ageing

Proefschrift voorgelegd tot het behalen van de graad van doctor in de taalkunde aan de Universiteit Antwerpen, te verdedigen door

Lise Paesen

Faculteit Letteren en Wijsbegeerte

Departement Taalkunde

Promotoren:
Prof. dr. Mariëlle Leijten
Prof. dr. Miet De Letter

Antwerpen, 2022

The story of a lifetime: A linguistic approach to studying changes in spontaneous written language due to ageing

Nederlandse titel: Het levensverhaal: Een taalkundige benadering van het bestuderen van veranderingen in spontaan geschreven taal ten gevolge van veroudering

The research in this dissertation was supported by a grant from BOF-DOCPRO 2015, the ‘Bijzonder Onderzoeksfonds’, grant number: 31661

ISBN: 9789057287329
Depotnummer: D/2022/12.293/04

Cover design and layout: © Lise Paesen

Printed by: Drukkerij Room | www.drukkerijroom.be

| | |
|---|---|
| Supervisors | prof. dr. Mariëlle Leijten (University of Antwerp) |
| | prof. dr. Miet De Letter (Ghent University) |
| | prof. dr. Peter Mariën (Vrije Universiteit Brussel) (†) |
| Doctoral Committee | prof. dr. Dirk Van Hulle (University of Antwerp) |
| | prof. dr. Sarah Bernolet (University of Antwerp) |
| Doctoral Jury | prof. dr. Peter Petré (University of Antwerp) |
| | prof. dr. Iris Schrijver (University of Antwerp) |
| | prof. dr. Victoria Johansson (Lund University) |
| | prof dr. Arnaud Szmalec (Université Catholique de Louvain & Ghent University) |

# Table of Contents

# Chapter one

## General introduction

# 1. ORIGIN OF THIS DISSERTATION

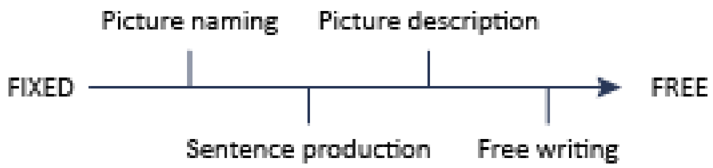
Language philosopher Ludwig Wittgenstein once stated '[t]he limits of my language mean the limits of my world' (1922). Even though Linguistic Determinism is not what we want to preach within this dissertation, we can state that when language is suddenly limited or impaired, the world of a patient with a language disorder becomes compromised. Research shows that patients who suffer from a acquired neurogenic language disorder will often suffer from an impacted quality of life (Ashaie et al., 2019; Hakim, 2011; Salary & Moghadam, 2013). Speeding the diagnostic process could benefit the general wellbeing of those patients, increase their life quality and even slow down the deterioration or aid the rehabilitation process. Given that age is a predominant risk factor for acquired neurogenic disorders, and provided that language changes already occur at the onset of several progressive disorders (e.g., Alzheimer's disease), research into a disruption of linguistic skills could provide additional insight into the onset of acquired neurogenic language disorders. In that respect, speech in the early stages of Alzheimer's disease has often been studied, in order to detect initial language changes that could aid the diagnostic process (Kavé & Dassa, 2018; Pekkala et al., 2013; Tsantali et al., 2013).

Within the context of pathological ageing, this dissertation started out with a focus on linguistic changes found in adults with Mild Cognitive Impairment (MCI) and Alzheimer's disease (AD); we sought to develop a non-intrusive writing test to describe the onset of cognitive decline in Alzheimer's disease. We strove to characterise the complete writing process focussing on the hand motor, cognitive and linguistic capacities of the patients and healthy controls. However, even in healthy controls, language changes can already be found due to neuroplasticity of the brain upon ageing (Burianová et al., 2013; Park & Reuter-Lorenz, 2009). More specifically, research points towards the vulnerability of language production which, in contrast to language comprehension, shows signs of age-related language decline (Burke & Mackay, 1997; Shafto et al., 2007). Therefore, the focus of this dissertation shifted. The decision was made to focus on linguistic changes in spontaneous language upon healthy ageing, in order to provide insight in the onset of various (progressive) disorders, thereby expanding our focus and contributing to various fields at once.

Mapping linguistic changes in healthy ageing and ageing with a cognitive impairment can be achieved by studying language on either a word, sentence or text level, or combination of those. Various experimental screening tasks have already been successfully developed to monitor some of these changes; depending on the target 'level', they range from the very controlled, fixed picture naming tasks that allow research insight on word level, to free writing tasks in which spontaneous speech without boundaries is studied (cf. Figure 1).

**Figure 1**

*Overview Of Different Experimental Screening Tasks, Depicted In Terms Of Task Flexibility*




*Note:* Examples: PNT (Chapter two, Cuetos et al., 2005; Goodglass et al. 1983); Picture Description Task (Forbes-McKay et al., 2014; Leijten et al., 2015); Sentence production (Catherine Meulemans et al., n.d.), Free writing (Garrard et al., 2005; Le et al., 2011).

One frequent method to study language changes on a **word level** is with the use of picture naming tasks (PNT). These tasks require the speeded naming (either spoken or written) of given prompts and allow researchers to study the effects of certain word characteristics of these prompts on naming speed and/or naming accuracy. Ample studies have proven that lexical knowledge and size continue to increase throughout life (Hardy et al., 2020; Kreuz & Roberts, 2019; Verhaeghen, 2003) and only decline in late adulthood (Kemper, Greiner, et al., 2001). In a study by Alantie and colleagues (2021), semantic fluency was only found to decrease from the age of 85 onwards. Given that the semantic system remains unaffected (Burke & Mackay, 1997; Clark-Cotton et al., 2007; Shafto et al., 2007), it is noteworthy that one of the most commonly noticed age-related linguistic changes is caused by a deficit in the phonological neural system: the 'tip-of-the-tongue' issue. This phenomenon leads to the temporary inability to remember a certain word, even though that word can be recalled later on without any issue (Burke & Mackay, 1997; Clark-Cotton et al., 2007; Shafto et al., 2007; Shafto & Tyler, 2014; Silagi et al., 2015). As a result, picture naming data has revealed that elderly become slower and more error-prone, especially with low-frequency words (Hardy et al., 2020).

In order to study language on a **sentence level**, lengthier output needs to be triggered. This can be done through studying larger language samples, leaving ample room for interpersonal differences, or for instance with the use of sentence production tasks and tests. The latter type



offers researchers the most controlled environment, as sentences are elicited by a range of prompted constructions, lexical items or images. The output can be interpreted from various perspectives, ranging from sentences produced (sentence construction, type of clauses) to studying the production process (focussing for instance on the fluency with which various types of clauses are produced). Note that these tasks can even be used to focus on specific target words elicited by the prompts. As a result, these tasks are used in various fields such as aphasia research (for instance Wilshire et al., 2014) and research into language changes upon ageing with a cognitive impairment and healthy ageing (such as Meulemans et al., n.d.). With regard to language in healthy ageing, research has found that syntactic skills are generally well-preserved, with dominant and subordinate clauses being processed equally effectively (Shafto & Tyler, 2014) and understanding complex sentences only becoming an issue very late in life (Clark-Cotton et al., 2007). Nevertheless, in terms of production, elderly tend to rely on simpler structures (Clark-Cotton et al., 2007), with fewer embedded and subordinate clauses (Croisile et al., 1996; Hardy et al., 2020; Williams et al., 2003). Also an increase in syntactic errors (for instance tense use) can be found (Hardy et al., 2020; Kemper, Greiner, et al., 2001). Relatively stable are the switching skills between syntactic structures (Altmann & Kemper, 2006; Hardy et al., 2020).

Studying language changes on a **text or discourse level** can be done by relying on different types of tasks and language samples. When opting for a more controlled setting, researchers have used image stimuli such as the Frog Story (for instance in Holmqvist & Johansson, 2005) or picture description tasks to elicit longer discourse about a predefined topic. Picture description tasks (PDTs) elicit language by presenting a depiction of for instance a domestic scene (e.g., the Cookie theft picture (Goodglass & Kaplan, 1983)) and requiring this image to be described by the participant. Other studies have opted to select language samples of certain authors (e.g., Iris Murdoch (Le et al., 2011) or ageing adults (e.g., diaries kept by nuns - The Nun Study (Kemper, Greiner, et al., 2001), as these samples mimic spontaneous language and thus free writing/speech most. As a result, language can be studied from both a product perspective, with a focus on for instance output quantity, number of filler words and pauses, word type use and the syntactic structure of the text, or a process perspective, focussing on pause durations within different parts of the spontaneous discourse. Note, however, how some studies have also focussed on the sentence and even word level when using these types of tests.

Previous studies within this broad field have pointed towards an impacted discourse, both in speech and writing. Given that word finding issues might arise and elderly are less able to filter distractions, more filler words will be used (Kreuz & Roberts, 2019), speech can become more disfluent with increased word repetition and with prolonged, frequent and empty pauses (Burke & Mackay, 1997; Shafto & Tyler, 2014). Word use becomes more vague (Shafto & Tyler, 2014), with unspecified referential pronouns or more errors when making references (Clark-Cotton et al., 2007; Juncos-Rabadán et al., 2005).

Additionally, some studies found that elderly are more verbose when describing personal situations as they have the tendency to include more irrelevant content. Nevertheless, elderly do use the same amount of semantic content and they tend to stay on topic when describing pictures (Juncos-Rabadán et al., 2005). Drawing up a story could pose issues later on in life, with difficulty organising the storylines into a cohesive structure, decreased referential ties and essential elements to the story that are being omitted (Clark-Cotton et al., 2007; Juncos-Rabadán et al., 2005; Sherratt & Bryan, 2019). Research on the quantity of speech is more ambiguous, with studies on ageing finding either the quantity of speech to increase in narratives, no effect, increased loquacity, or even reduced speech (Juncos-Rabadán et al., 2005).

## 2. METHODOLOGY

Despite the fact that ample studies focus on spoken language changes in ageing (Antonsson et al., 2021; Farias et al., 2012; Gosztolya et al., 2019; Hardy et al., 2020; Orimaye et al., 2017) and that tools to map these changes are available, the current test battery sets have three major shortcomings for mapping changes to spontaneous written language generation. In what follows, (1) these shortcomings and how they resulted in the creation of narrative writing tasks are discussed; (2) the reasoning behind the selected language analysis tools is explained.

### 2.1 Narrative writing tasks

The first shortcoming for mapping changes to spontaneous written language generation with the available test battery sets is the duality of **task control** and studying language changes. On the one hand, researchers want to create a writing task in which every produced word is anticipated for. This allows for key variables to be determined beforehand to analyse and differentiate linguistic processes of cognitively impaired patients and healthy controls. On the other hand, letting a participant write a text on a topic of their choosing - allowing spontaneous language generation - triggers other important variables to distinguish healthy cognitive changes from

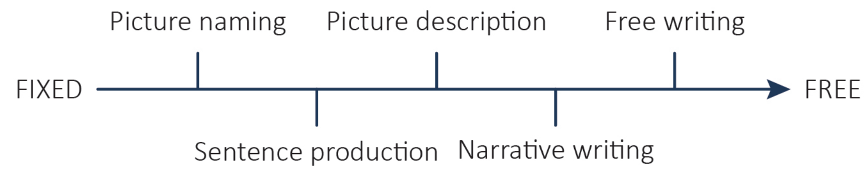




unhealthy cognitive decline. Concepts such as idea generation can be measured that way. Results from the pilot by Leijten and colleagues (2015), Paesen (2015) & Meulemans and colleagues (n.d.) indicate that a combination of the existing picture description tasks did not manage to balance this duality: they did not have the inherent possibility to fully control the output of certain word types, and the produced texts were too constraint in topic. Within this dissertation, writing tasks were developed that control for specific target words and at the same time allow for a diverse narrative output. Figure 2 exemplifies how we position ourselves within the framework of free vs restricted language generation.

**Figure 2**

*Overview Of How We Position Ourselves Between The Different Experimental Screening Tasks, Depicted In Terms Of Task Flexibility*




Secondly, profound linguistic analysis relies on sufficient **text length**. Written output obtained by means of a picture description task is too short for automated language analysis. Research by Faroqi-Shah et al. (2020) has indicated that the Cookie Theft Picture elicited short and simple stories, that only contain a limited number of ideas. As a result, the lexical diversity and tense use were limited. They state the need for tests that elicit longer samples, so as to potentially make a more sensitive differential diagnosis. Similarly, a pilot project also indicated that the current picture description tasks trigger texts of insufficient and incomparable text length (Paesen, 2015). Bearing the need for iterative research in mind, this dissertation, therefore, strives to obtain texts of at least 100 words, in order to conduct relevant automated analyses.

Lastly, we want to stress the need for **congeneric tasks**. Since we are interested in the linguistic changes upon ageing, longitudinal data collection with congeneric materials is needed so as to make the tasks interchangeable on certain key variables (Forbes-McKay et al., 2014). As proven by our pilot study (Paesen, 2015), the current set of picture description tasks is too diverse in nature to be used interchangeably in longitudinal research.

Given these shortcomings and the need for a non-intrusive, low-cost and time effective tool, we decided to create narrative writing tasks that elicit spontaneous written language whilst still controlling for certain key variables. Inspiration was found in combining the strength of the picture description tasks / story retelling tasks and the picture naming tasks. The former allows for a storyline to be created, the latter for specific target nouns to be triggered. We opted for a selection of multiple images to be presented at once, provided that all images depict a single living or non-living thing. In consequence, a narrative would need to be written without constraints about the topic, as long as the depicted images were used in the story. In other words, a spontaneous narrative is elicited that still controls for specific target words, in the form of the nouns, triggered by the individual single-subject images.

### *2.1.1 Step 1: creation of new images*

Provided that images would form the basis of our narrative writing tasks, the selection of images needed to be a thoroughly controlled process. One of the strengths of the current study is that we created a new image database, which combines the findings from prominent studies in the field and overcomes some of the issues the existing image databases pose with regards to usage within narrative writing tasks. The main issues we tried to counter were that the existing databases contained images:

- without colour, with only black-and-white line drawing;
- that depict dated objects (e.g., a rotary phone);
- that depict multiple subjects (e.g., a pirate with a sword, or a dolphin and water);
- that depict the subject doing an action (e.g., a dolphin jumping out of water);
- that are not consistent in their depiction (e.g., a dolphin looking left, a dog looking right and a cat looking straight forward)

Bearing these issues in mind, we believed a database that uses images that follow strict depiction rules would result in a better control of the target noun that is triggered by the image. Therefore, based on previous studies, existing single-subject images were selected and used as an inspiration. They were selected based on linguistic variables, such as age-of-acquisition, frequency, and name agreement in other studies. Previous reading studies in Dutch (Severens et al., 2005) served as a basis in order to predict the outcome for picture naming in this Dutch writing experiment. Additionally, we also checked for the occurrence of the selected images in other studies in Dutch, such as the CAT-NL, as it could be possible for a patient to receive both





that test and ours and learning effects ought to be avoided. The words should also be easily typed, avoiding words that contain for instance an umlaut (such as ‘pinguïn’ – penguin in Dutch).

Additionally, the target nouns were differentiated based on two categories, as suggested by Forbes-McKay and colleagues (2014). They advised the use of two task types, a simpler task and a more complex one, in order to avoid both floor and ceiling effects. These tasks should differentiate in the amount of information and number of pictorial themes to be triggered. Given that this dissertation focusses on creating the basis of a screening tool for language changes in healthy ageing, it could serve as a diagnostic aid for differentiating between healthy ageing and ageing with a cognitive impairment. The latter group might find the simple task easy and will be challenged by the complex task. Patients with a more severe impairment may already experience the simple task as quite complex. In order to differentiate the tasks between simple and complex (respectively Narrative Writing Task 1 or NWT1 and Narrative Writing Task 2 or NWT2), we imposed additional strict criteria on the target nouns to distinguish between the two types (NWT1 and 2). See Table 1 and Chapter two for a full review.

The final selection consisted of 50 target nouns that needed to be transformed into images; 20 of those were selected to create NWT1, the easier narrative writing task, and 30 were selected to create NWT2, the more complex narrative writing task. Based on this selection, we redrew the existing images from previous studies and imposed various criteria on the visual requirements. These visual requirements served to overcome the issues we found with previous image databases, specifically with regard to usage within narrative writing tasks. We believed the following criteria would result in the highest image recognition and task control within narrative writing tasks:

- the images need to be coloured;
- the images require a simplistic drawing style, without too many details;
- the images can only depict a single object;
- the images cannot depict any movement; and
- per category, the images all need to point in the same direction.

**Table 1**

*Overview Of Picture Selection Criteria Per NWT In Dutch*

| | Images for NWT1 | Images for NWT2 |
|---|---|---|
| Name Agreement (%) | > 50% | > 50 % |
| Age-of-acquisition (years) | Young (-6) | Older (+6) |
| Frequency (per million words) | >30 | <10 |
| Reaction Times (in ms) | 631.5 < X < 981,34 | 652 < X < 1191.48 |
| Occurrence in other studies | Ignored | Accounted for |
| Spelling | Normal | Normal |

A pre-test was set up to ascertain if the name agreement of the newly developed images was sufficiently high. A cohort of 200 healthy participants, comprised of both students and adults, received a set of 50 images that had to be named. Their set consisted of a random combination of 25 original black-and-white images and 25 of our new images. The results led to the deletion of two images, due to poor name agreement (below 50%). The resulting image set held a total of 48 images, 19 of the NWT1 images and 29 NWT2 images. The full list of images can be found in **Appendix A**.

### 2.1.2 *Step 2: creation of narrative writing tasks*

A narrative writing task is a task that requires the participant to write a story based on prompted images. They are a new type of tasks, designed to generate spontaneous written language while still controlling for key variables; those key variables centre around target nouns triggered by the images. The controlled design of the images and dataset enhanced task control and additionally the congeneric character of the task by allowing us to randomise the images for every task and every participant. The design of the narrative writing tasks contributed to the controlled character of the test, while also allowing for free narration and interchangeability in longitudinal settings. By distinguishing between NWT1 and NWT2 in our task design and instruction, we enabled greater text length and overcame both floor and ceiling effects. In what follows, a full description of the logic behind the narrative writing tasks is given.

The strength of the narrative writing tasks springs from the fact that the distinction between the two tasks was not only made through the use of images; it was also enforced by our task design. A distinction was made in number and type of images and task instructions. The first or simple

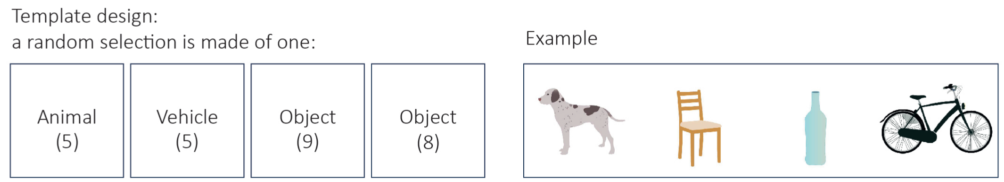




narrative writing task (NWT1) is inspired by activities from daily life. Its images have been selected to contain everyday objects/subjects; the narrative that is generated is one that is close to the daily lives of the participants. Given that patients with a cognitive impairment tend to tackle narrative difficulties by narrating from an autobiographical point of view (Hydén & Örulv, 2009), the use of the first person personae could enable most patients with a mild language impairment to participate in this test without suffering from a ceiling effect. For each NWT1, a combination of four images of the simple subset was made. This combination was semi-random, always picking one of the five animals, one out of the five vehicles and two of the ten objects for each template (Figure 3).

**Figure 3**

*Template Design And An Example Of The Selection Of Pictures In A NWT1*

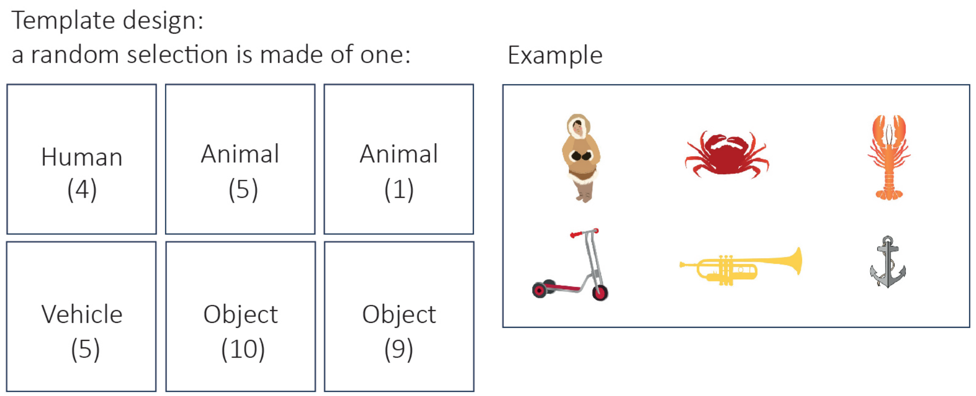



The second narrative writing task (NWT2) was designed to be more challenging. The difficulty was added through the images, tense use and required personae use. Within this task, six images needed to be incorporated that did not necessary fit together: one human, two animals, a vehicle and two objects (an example of the template and a random selection of images can be found in Figure 4). The two animals were predetermined pairs as they needed to be visually and/or semantically related (e.g., a dolphin and a shark) and shared a categorical relationship, increasing task difficulty. Previous studies found longer reaction times for semantically / visually related pairs in healthy adults compared to unrelated pairs (Campos-Magdaleno et al., 2020; Kumar Sanju, 2017; Laisney et al., 2011), and slower reaction times for words that share a categorical relationship (Cocquyt et al., 2021). Additionally, since the six concepts depicted by images were most of the time unrelated, participants needed to rely on their imagination and step away from the familiar, possibly triggering difficulties for patients with a cognitive impairment (Hydén & Örulv, 2009). This effect was further enhanced by the required use of the past tense and the use of the third person personae.

The versatile yet congeneric character of the narrative writing task contributes to its strength. Because the task is made up of a randomised set of images, the experiment can be done multiple times by the same patient, without receiving the same selection of pictures to elicit a narrative. More specifically, 1.800 and 9.000 variations to respectively the NWT1 and NWT2 are possible. As a result, in theory, a person could redo this test every year throughout their lives, to accurately assess their potential language changes without suffering from a learning effect, while still being evaluated based on the same variables. During our experiments, each participant in the longitudinal study received a selection of pictures that was sampled without replacement, i.e.: a participant never received the same images in their NWT twice. However, throughout the various studies within this dissertation, the order of NWT1 – NWT2 never changed.

**Figure 4**

*Template Design And An Example Of The Selection Of Pictures In A NWT2*




*Note*: The number in brackets below each image type indicates the number of variants there are for that image type within the template.





## 2.2 Selection of language analysis tools

Gaining insight into spontaneous written language changes on a text level can be achieved by studying language samples; previous studies have therefore relied on for instance bibliographies (e.g., the nun study (Kemper et al., 2001) or novels produced by a certain author at a certain time (e.g., research into Iris Murdoch's changing cognition due to Alzheimer's disease (Le et al., 2011)). However, the processes behind text construction provide insight in language ability as well. For instance, in writing processes as well as in oral speech production, pauses in the production flow are considered signs of cognitive complexity (Olive, 2012; Van Waes & Leijten, 2015). With this dissertation, we want to adopt a combined approach, in which we include both product and process-oriented variables. We strive to characterise the writing process and product by focussing on the hand motor and linguistic capacities of our healthy target group and possibly several cases with a possible pathology.

### *2.2.1 Writing processes - keystroke logging tools*

One possible method to provide an adequate description of these linguistic capacities and the resulting writing process is keystroke logging. Keystroke logging tools are some form of computer software that – when activated - record every keystroke activity, mouse movement and could even record speech samples. As these actions/movements are time stamped, the text production process can be accurately reconstructed. These tools are made to be unobtrusive, as they are programmed to run in the background and, therefore, other methods of observation can often be run in parallel, such as thinking aloud protocols. As a result, keystroke logging tools can be used in various fields of study. Three main research categories can be found: theoretical, methodological and applied research. Within the (1) theoretical category, both writing development is studied (e.g., research by Baaijen and colleagues (2014) and Lindgren and colleagues (2011)), and theoretical models of writing are tested (e.g., research by Thierry Olive (2014) or Van Waes & Leijten (2015)). The studies focussing on (2) methodology aim to make suggestions as to the way in which the data should be addressed in various fields of study (e.g., studies by Galbraith and colleagues (2012) and Van Waes & Leijten (2015)), and typing measures are related to the cognitive effort behind the writing (see studies by (Chukharev-Hudilainen (2014), or Van Waes and colleagues (2017)). Lastly, the (3) applied studies aim to contribute to ongoing studies into various subprocesses within translation (e.g., Schrijver and colleagues (2016)), L1-L2 research (e.g., Van Waes & Leijten, 2015) or even writing with learning difficulties or a cognitive impairment (see for instance Galbraith and colleagues (2012) or Van Waes and colleagues (2017)). This latter aspect is of crucial importance to the current dissertation. Exploratory studies conducted by Leijten and colleagues suggest that process data can be used in aiding the diagnosis of dementia (Leijten et al., submitted; Leijten et al., 2015; Van Waes et al., 2017). In consequence, the decision to use computer literacy as a participant inclusion criterion proved to be feasible and adequate (Van Waes & Leijten, 2015). Due to the increased use of computers in daily life, even among the elderly, the participants had little difficulties performing the writing tasks on the computer. Since the group of computer literate elderly will continue to grow, this technique will be even more broadly applicable in the future.

For this dissertation, the decision was made to use two complementary keystroke logging tools simultaneously: Inputlog (Leijten & Van Waes, 2013) and ScriptLog (Frid et al., 2014). Although both tools offer both logging and analysis opportunities, we believe both programs had their merits with regard to the different studies presented in this dissertation. Inputlog is one of the most prominent keystroke logging tools, due to its versatility and analysing strength. To this day, sixteen different types of analyses are available, providing various perspectives on the writing process. It allows researchers to study both more fine-grained variables, such as within word pauses as more robust variables such as total time in the word document. Additionally, the tool offers a typing test which we used to gain insight into the typing skills and computer literacy of the participants.

ScriptLog offered a custom-made module to log the entire writing process of both the narrative writing tasks and picture naming task. The advantages of this custom-made module were twofold. Firstly, it enabled the researcher to receive automatically made and randomised templates (instead of by hand), catering for longitudinal research more effectively. Secondly, the module interface displayed the text box and image (e.g., of the narrative writing task or picture naming tasks) in a fixed frame, ensuring sufficient visibility and the elimination of possible interferences in the logging / writing process.



### *2.2.2 Writing product - T-Scan*



Another potential method to describe the linguistic capacities of a participant is studying their written output in terms of its coherence and / or cohesiveness. In the early days of text analysis, researchers tried to automate the analysis of text difficulty by studying features such as word length & frequency (e.g., Dale & Chall (1948)). Later on, the focus shifted and a greater emphasis was placed on text coherence, which eventually lead to the creation of Coh-Metrix, an English 'computational linguistic tool that measures text cohesion and text difficulty on a range of word, sentence, paragraph, and discourse dimensions' (McNamara et al., 2010, p.1). Similarly, T-Scan was designed for the analysis of Dutch texts, and offers over 300 variables that provide insight into lexical / sentence complexity, referential cohesion and lexical diversity, relational coherence, concreteness, personal style, verbs & time, parts-of-speech, and probability features (Pander Maat et al., 2014). The wide range of variables allowed us to fine-grain our analysis and target specific cohesion measures.

## 3. MAIN AIMS

Throughout this dissertation, we have endeavoured to find out how spontaneous written language changes upon ageing, by adopting a linguistic approach. To reach this goal, we defined the aim of the current research project to be fourfold:

1. to establish how age, image characteristics and repeated testing affect picture naming;
2. to find out how ageing and sex affect the cohesion of written narratives;
3. to examine the test-retest reliability of spontaneous written language in healthy ageing;
4. to explore how spontaneous written language can be used to make a differential diagnosis.

In order to answer the main aims of this dissertation, we decided to construct a new clinical toolkit for the evaluation of spontaneous written language. This toolkit needed to be non-intrusive, low-cost and time effective. Three additional, technical, aims were constructed for this purpose. We strove:

5. to create a toolkit that bridges the gap within differential diagnostics with regard to spontaneous written language generation;
6. to create tasks within that toolkit that are congeneric in nature and that are interchangeable on certain key characteristics;
7. to create and standardise the toolkit within a healthy ageing population and pilot its validity within a clinical setting.

The entire process, different studies and chapters strive to formulate an answer to our main research question and also contribute to the construction of our toolkit, as visualised by Figure 4. In order to create narrative writing tasks, new images needed to be created that could elicit a narrative without predefining the content of that narrative apart from certain target words. In our image validity study (chapter two), new images were created, pre-tested and standardised in a cross-sectional study. As a result, the effect of ageing, image characteristics and repeated testing on naming could be tested as well. These results gave rise to the narrative writing tasks used in the continuation of this dissertation (chapters 3 – 5). More specifically, with the use of these images, narrative writing tasks were constructed, pre-tested and standardised in both a cross-sectional (chapter 3) and longitudinal (chapter 4) study. These chapters provide insight into the effect of ageing, sex and test-retesting on narrative writing tasks. Additionally, results from the cross-sectional study were used to construct a normative set, which gave input for all the following studies within this dissertation (chapters 4 – 5). To establish insight into potential

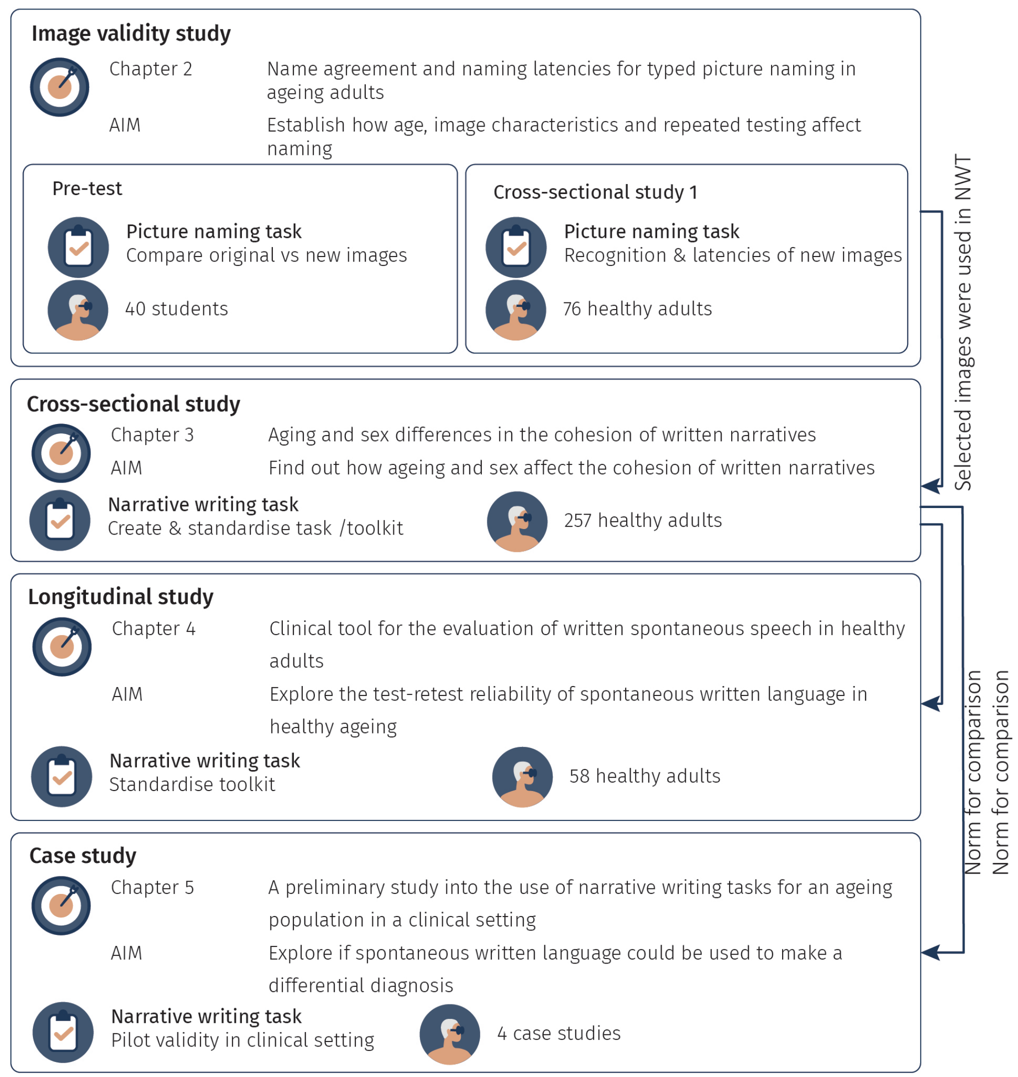




language biomarkers, we built on insights from the cross-sectional study (chapter 3) and specific examples from the case-studies (chapter 5).

**Figure 4**

*The Entire Process, Different Studies And Chapters That Contribute To Our Understanding Of Changes To Spontaneous Written Language Upon Ageing And The Construction Of A Toolkit*



## 4. OVERVIEW OF THIS DISSERTATION

This dissertation comprises of a total of six chapters: this general introduction, four studies that each address part of the main research question: 'How does spontaneous written language change upon ageing?', and a general discussion. Given that each chapter is constructed as a separate journal article, each chapter can be read on its own. Nevertheless, these different chapters all contribute to this research question, thereby partly sharing the same theoretical framework and method sections. In what follows, the aforementioned research question will be addressed with the use of our four different chapters.

Chapter one (the current chapter) is the general introduction. It provides insight into the structure of this dissertation, how the different chapters follow-up and link to each other, and hopefully guides the reader into a clearer understanding of the topic. Furthermore, it serves to familiarise the reader with some key elements that are crucial for understanding research into spontaneous written language studies.

**Chapter two** presents the study titled '*Name agreement and naming latencies for typed picture naming in ageing adults*', and aims to find an answer to the research question:

*How do age, image characteristics and repeated testing affect the naming accuracy and latencies within a picture naming task?*

This research project started off with the creation of a new, coherent set of images that could be used to elicit a narrative without predefining the storyline and that could trigger certain target words. Chapter two addresses the creation of this image database, and aims to establish how age, image characteristics and repeated testing affect naming. Therefore, the image validity study set out to develop a coherent set of images that could be used to distinguish healthy language changes from pathological language decline upon ageing. The focus was not specifically on the use within narrative writing tasks; the image database was designed for broader use, wherefore the images can also be used by themselves, for instance in a picture naming task. Differentiation from existing databases was needed in order to generate a greater control over the written output and combine the findings from previous studies and create more cohesion. This image validity study and the thereby created images served as a baseline for the continuation of this dissertation.





After the careful selection of relevant images, dividing them into two categories (highly vs less relatable, respectively HR and LR) based on certain image characteristics (e.g., word frequency, age-of-acquisition, colour), a typed picture naming task was given to a representative sample of 60 healthy ageing adults aged 50 and over. We hypothesized that:

- In terms of the naming product (naming accuracy)
  - name agreement & object recognition would be equal for the HR and LR categories;
  - name agreement & object recognition would not differ between the two age groups (AG 50 – 64 and AG 65+).
- In terms of the naming process (latencies)
  - naming & interkey latencies would be longer for LR than HR nouns;
  - naming & interkey latencies would be longer for the older age group (AG 65+);
  - naming & interkey latencies would remain the same with a three-month interval.
- In terms of the connection between product and process
  - name agreement could be used to predict naming & interkey latencies due to its high correlation with the latencies.

Given these hypotheses, the results were studied both in light of the naming product and naming process, respectively referred to as naming accuracy (name agreement & object recognition) and latencies (naming latencies and interkey latencies). The naming product was manually filtered and reduced in order to categorise spelling mistakes, synonyms, incorrect naming and blanks. The naming process was logged with the use of the keystroke logging tool ScriptLog (Frid et al., 2014) and analysed with Inputlog (Leijten & Van Waes, 2013).

In **Chapter three**, '*Ageing and sex differences in the cohesion of written narratives*', we wanted to (1) explore potential effects ageing and/or sex has on the cohesion of written narratives and (2) create congeneric narrative writing tasks. We set out to find an answer to the following research question,

*'Do ageing & sex affect the cohesion of written narratives?'.*

Based on previous studies, we hypothesized that (1) grammatical and lexical complexity would decline with age and that (2) the grammatical and lexical output would significantly differ between men and women. Based on the images created in chapter two, narrative writing tasks were constructed. Those tasks combine the strength of spontaneous language generation and the controlled environment of picture naming tasks, in which certain target words are elicited with the use of images. More specifically, in such a task, a story must be written based on the prompted coloured images. By doing so, our study adds to the existing test batteries by creating a controlled narrative writing task for spontaneous language generation that can be used in longitudinal research and even possibly serve as a language screening tool for differentiating healthy language changes and language changes due to a pathology.

Given that this study will form the baseline for future studies into narrative writing tasks, we were interested in the written product. We focussed on grammatical and lexical complexity and explored how age and sex affected the written narrative. Additionally, we also wanted to create a normative dataset that could be used as reference in future studies.

In order to answer the posited research questions, we selected a sample of 257 self-reported cognitively healthy participants (students, N = 39; adults aged 50+, N = 218) who each were given two narrative writing tasks (respectively NWT1 and NWT2). Based on the findings from the previous study, the healthy ageing adults were not treated as a homogeneous group in the analysis, but rather as a set of individuals through a multi-level analysis in RStudio. For the presentation of the normative dataset, the findings were divided per decade. Cohesion was measured with the use of lexical and grammatical variables, extracted from the narratives with T-Scan (Pander Maat et al., 2014). Multi-level analysis provided insight into the language changes upon ageing and between sexes. The findings provided both a normative database and valuable insights that were used in our follow-up studies, respectively chapters four and five.

**Chapter four** presents the study titled: '*Clinical tool for the evaluation of written spontaneous speech in healthy adults*'. 58 healthy ageing adults participated thrice in the study. The third study in this dissertation focusses on determining the test-retest reliability of spontaneous written speech in healthy adults and establishing a new clinical tool based on these findings. We centre our findings around the following research question:

*To what extent does spontaneous written language affect the test-retest reliability of healthy ageing?*

We posited the hypothesis that more pauses would be needed by the older participants due to word finding issues, even though the writing product will not be affected (e.g., in terms of total word count). Additionally, we hypothesized that no change in both product and process data would be visible over the course of the three trials.





In other words, we defined the aim as being twofold. Firstly, we wanted to standardise the narrative writing tasks in a longitudinal setting, explore the effect of repeated testing on spontaneous narrative writing. Secondly, we aimed at exploring the writing processes triggered by the narrative writing tasks. The answer was sought through a study with a longitudinal design applying the same design as chapter three, using the narrative writing tasks. Results were studied with a focus on both the written product and the writing process. A subsample of 58 participants from the healthy cohort described in chapter three partook in this longitudinal study. In total, the experiment was administered three times with an interval of three months in between the test moments.

**Chapter five**, '*A preliminary study into the use of narrative writing tasks for an aging population in a clinical setting*', provides an answer to the research question:

*In terms of spontaneous language generation, which parameters are needed for a differential diagnosis between healthy and pathological?*

Given that our participants were self-reported healthy individuals, additional screening assured their cognitive or psychological health. The cases presented in the current study were eliminated from the longitudinal dataset due to disturbances in either their cognitive or psychologic health; however, their data did provide additional insights into possible pathologies and their writing processes with regards to narrative writing tasks. In order to find an answer to the main research question, we hypothesised that:

- the picture naming task would not shed light on different pathologies;
- both lexical and grammatical measures would be needed in order to make a differential diagnosis;
- grammatical markers, such as the density of personal pronouns, could be used to make a differential diagnosis; and
- aphasic language will be differentiated based on lexical markers.

Even though the participants in these cases originally enrolled as healthy participants and therefore followed the same procedures as in chapter three, issues were found in their cognitive and/or psychological health with the use of the MoCA, GDS or even their medical history reported in their questionnaire. A clinical language pathologist studied the narrative writing tasks in order to confirm potential pathologies – of course without clinical testing. These tasks were also compared to the normative dataset created in chapter three. Additionally, their picture naming tasks were studied, in line with the study presented in chapter two. This study allowed us to contribute to the two aims of this dissertation: explore if spontaneous written language could be used to make a differential diagnosis and by piloting the tasks within a clinical setting

The final chapter, **chapter six**, consists of the general discussion. The main findings of our conducted studies will be explained and their implications for future studies will be addressed. Both the limitations and strengths of the developed toolkit will be pinpointed, and the applicability of a narrative writing task that elicits spontaneous language will be discussed for both a healthy ageing adults and for adults ageing with a cognitive impairment. The final part of the discussion is the conclusion.



# 5. BIBLIOGRAPHY



Alantie, S., Tyrkkö, J., Makkonen, T., & Renvall, K. (2021). Is Old Age Just a Number in Language Skills? Language Performance and Its Relation to Age, Education, Gender, Cognitive Screening, and Dentition in Very Old Finnish Speakers. *Journal of Speech, Language, and Hearing Research*, 1–18. https://doi.org/10.1044/2021_JSLHR-21-00178

Altmann, L. J. P., & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes*, 21(1–3), 322–354. https://doi.org/10.1080/01690960544000006

Antonsson, M., Lundholm Fors, K., Eckerström, M., & Kokkinakis, D. (2021). Using a Discourse Task to Explore Semantic Ability in Persons With Cognitive Impairment. *Frontiers in Aging Neuroscience*, 12(January), 1–11. https://doi.org/10.3389/fnagi.2020.607449

Ashaie, S. A., Hurwitz, R., & Cherney, L. R. (2019). Depression and Subthreshold Depression in Stroke-Related Aphasia. *Archives of Physical Medicine and Rehabilitation*, 100(7), 1294–1299. https://doi.org/10.1016/j.apmr.2019.01.024

Baaijen, V. M., Galbraith, D., & de Glopper, K. (2014). Effects of writing beliefs and planning on writing performance. *Learning and Instruction*, 33, 81–91. https://doi.org/10.1016/j.learninstruc.2014.04.001

Burianová, H., Lee, Y., Grady, C. L., & Moscovitch, M. (2013). Age-related dedifferentiation and compensatory changes in the functional network underlying face processing. *Neurobiology of Aging*, 34(12), 2759–2767. https://doi.org/10.1016/j.neurobiolaging.2013.06.016

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philos Trans R Soc Lond B Biol Sci*, 352(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

Campos-Magdaleno, M., Leiva, D., Pereiro, A. X., Lojo-Seoane, C., Mallo, S. C., Nieto-Vieites, A., Juncos-Rabadán, O., & Facal, D. (2020). Longitudinal Patterns of the Tip-of-the-Tongue Phenomenon in People With Subjective Cognitive Complaints and Mild Cognitive Impairment. *Frontiers in Psychology*, 11(March), 1–14. https://doi.org/10.3389/fpsyg.2020.00425

Chukharev-Hudilainen, E. (2014). Pauses in spontaneous written communication: A keystroke logging study. *Journal of Writing Research*, 6(1), 61–84. https://doi.org/10.17239/jowr-2014.06.01.3

Clark-Cotton, M. R., Williams, R. K., Goral, M., & Obler, Loraine, K. (2007). *Language and Communication in Aging* (pp. 1–8). https://doi.org/10.1016/B0-12-370870-2/00103-7

Cocquyt, E. M., Santens, P., van Mierlo, P., Duyck, W., Szmalec, A., & De Letter, M. (2021). Age- and gender-related differences in verbal semantic processing: the development of normative electrophysiological data in the Flemish population. *Language, Cognition and Neuroscience*, 0(0), 1–27. https://doi.org/10.1080/23273798.2021.1957137

Croisile, B., Ska, B., Brabant, M.-J., Duchene, A., Lepage, Y., Aimard, G., & Trillet, M. (1996). Comparative Study of Oral and Written Picture Description in Patients with Alzheimer's Disease. *Brain and Language*, 53(53), 1–19. https://doi.org/10.1006/brln.1996.0033

Dale, E. & J.S. Chall (1948). A formula for predicting readability. *Educational Research Bulletin* 27, 37-54

Farias, S. T., Chand, V., Bonnici, L., Baynes, K., Harvey, D., Mungas, D., Simon, C., & Reed, B. (2012). Idea density measured in late life predicts subsequent cognitive trajectories: Implications for the measurement of cognitive reserve. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 67 B(6), 677–686. https://doi.org/10.1093/geronb/gbr162

Faroqi-Shah, Y., Treanor, A., Ratner, N. B., Ficek, B., Webster, K., & Tsapkini, K. (2020). Using narratives in differential diagnosis of neurodegenerative syndromes. *Journal of Communication Disorders*, 85(January 2019), 105994. https://doi.org/10.1016/j.jcomdis.2020.105994

Forbes-McKay, K., Shanks, M., & Venneri, A. (2014). Charting the decline in spontaneous writing in Alzheimer's disease: a longitudinal study. Acta *Neuropsychiatrica*, 26(04), 246–252. https://doi.org/10.1017/neu.2014.2

Frid, J., Johansson, V., Johansson, R., Wengelin, Å., & Johansson, M. (2014). Developing a keystroke logging program into a writing experiment environment. Presentation at Writing Across Borders, WRAB, Paris February 2014.

Galbraith, D., Baaijen, V. M., Smith-Spark, J., & Torrance, M. (2012). The Effects of Dyslexia on the Writing Processes of Students in Higher Education. *In Learning to Write Effectively: Current Trends in European Research* (pp. 195–198). BRILL. https://doi.org/10.1163/9781780529295_044

Goodglass, H., & Kaplan, E. (1983). The Boston Diagnostic Aphasia Examination. Lea and Febiger.

Gosztolya, G., Vincze, V., Tóth, L., Pákáski, M., Kálmán, J., & Hoffmann, I. (2019). Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech and Language*, 53, 181–197. https://doi.org/10.1016/j.csl.2018.07.007

Hakim, A. M. (2011). Depression, strokes and dementia: New biological insights into an unfortunate pathway. *Cardiovascular Psychiatry and Neurology*, 2011(Mci). https://doi.org/10.1155/2011/649629

Hardy, S. M., Segaert, K., & Wheeldon, L. (2020). Healthy Aging and Sentence Production: Disrupted Lexical Access in the Context of Intact Syntactic Planning. *Frontiers in Psychology*, 11(February). https://doi.org/10.3389/fpsyg.2020.00257

Holmqvist, K., & Johansson, V. (2005). Perceiving and Producing the Frog Story. *Perspectives on Language and Language Development,* January. https://doi.org/10.1007/1-4020-7911-7

Hydén, L. C., & Örulv, L. (2009). Narrative and identity in Alzheimer's disease: A case study. *Journal of Aging Studies*, 23(4), 205–214. https://doi.org/10.1016/j.jaging.2008.01.001





Juncos-Rabadán, O., Pereiro, A. X., & Rodríguez, M. S. (2005). Narrative speech in aging: Quantity, information content, and cohesion. *Brain and Language*, 95(3), 423–434. https://doi.org/10.1016/j.bandl.2005.04.001

Kavé, G., & Dassa, A. (2018). Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology*, 32(1), 27–40. https://doi.org/10.1080/02687038.2017.1303441

Kemper, S., Greiner, L., Marquis, J., Prenovost, K., & Mitzner, T. (2001). Language decline across the life span: Findings from the nun study. *Psychology and Aging*, 16(2), 227–239. https://doi.org/10.1037/0882-7974.16.2.227

Kreuz, R., & Roberts, R. (2019). Changing minds. How Aging Affects Language and How Language affects aging. MIT Press

Kumar Sanju, H. (2017). Semantic Priming and Reaction Time Measures in 5-6 Year Old Normal Subjects. *Open Access Journal of Neurology & Neurosurgery*, 6(2). https://doi.org/10.19080/OAJNN.2017.06.555683

Laisney, M., Giffard, B., Belliard, S., de la Sayette, V., Desgranges, B., & Eustache, F. (2011). When the zebra loses its stripes: Semantic priming in early Alzheimer's disease and semantic dementia. *Cortex*, 47(1), 35–46. https://doi.org/10.1016/j.cortex.2009.11.001

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing*, 26(4), 435–461. https://doi.org/10.1093/llc/fqr013

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication*, 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Lindgren, E., Leijten, M., & Van Waes, L. (2011). Adapting to the reader during writing. W*ritten Language & Literacy,* 14(2), 188–223. https://doi.org/10.1075/wll.14.2.02lin

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes*, 47(4), 292–330. https://doi.org/10.1080/01638530902959943

Meulemans, C., Leijten, M., Van Waes, L., Engelborghs, S., & De Maeyer, S. (n.d.). Cognitive writing process characteristics in Alzheimer's disease. *Submitted.*

Meulemans, Catherine, Leijten, M., & De Maeyer, S. (n.d.). The influence of ageing and verb transitivitiy on written sentence production. *Submitted.*

Olive, T. (2012). Working Memory in Writing. In V. Beringer (Ed.), Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology (pp. 485–506.). *Psychology Press*. https://doi.org/10.1007/978-94-010-0804-4_6

Olive, Thierry. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, 6(2), 173–194. https://doi.org/10.17239/jowr-2014.06.02.4

Rimaye, S. O., Wong, J. S. M., Golden, K. J., Wong, C. P., & Soyiri, I. N. (2017). Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. *BMC Bioinformatics*, 18(1), 1–13. https://doi.org/10.1186/s12859-016-1456-0

Paesen, L. (2015). De verbetering van schrijfprocesanalyse ter diagnostisering van alzheimer : vooronderzoek naar cognitieve beperking met behulp van vernieuwde descriptieve taken en de invloed die typevaardigheid erop heeft [*Master's Thesis*, University of Antwerp].

Pander Maat, H., Kraf, R., Bosch, A. van den, Dekker, N., Gompel, M. van, Kleijn, S., Sanders, T. J. M., & Sloot, K. van der. (2014). T-Scan : a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain : Aging and Neurocognitive Scaffolding. *Annual Review of Psychology*, 60, 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656. The

Pekkala, S., Wiener, D., Himali, J. J., Alexa, S., Obler, L. K., Liu, Y., & Kee, A. N. N. M. C. (2013). Lexical retrieval in discourse: An early indicator of Alzheimer's dementia. *Clinical Linguistics & Phonetics,* 27(December), 905–921. https://doi.org/https://doi.org/10.3109/02699206.2013.815278

Salary, S., & Moghadam, M. A. A. (2013). Relationship Between Depression and Cognitive Disorders in Women Affected with Dementia Disorder. *Procedia - Social and Behavioral Sciences*, 84, 1763–1769. https://doi.org/10.1016/j.sbspro.2013.07.028

Schrijver, I., Van Vaerenbergh, L., Leijten, M., & Van Waes, L. (2016). The impact of writing training on transediting in translation, analyzed from a product and process perspective. *Perspectives*, 24(2), 218–234. https://doi.org/10.1080/0907676X.2015.1040034

Severens, E., Lommel, S. Van, Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, 119(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the Tip-of-the-Tongue: Neural Correlates of Increased Word-finding Failures in Normal Aging. *Journal of Cognitive Neuroscience*, 19(12), 2060–2070. https://doi.org/10.1162/jocn.2007.19.12.2060

Shafto, M. A., & Tyler, L. K. (2014). Decline and Preservation. *Science*, 346(6209), 583–588. https://doi.org/10.1126/science.1254404

Sherratt, S., & Bryan, K. (2019). Textual cohesion in oral narrative and procedural discourse: the effects of ageing and cognitive skills. *International Journal of Language & Communication Disorders*, 54(1), 95–109. https://doi.org/10.1111/1460-6984.12434





Silagi, M., Bertolucci, P., & Ortiz, K. (2015). Naming ability in patients with mild to moderate Alzheimer's disease: what changes occur with the evolution of the disease? *Clinics*, 70(6), 423–428. https://doi.org/10.6061/clinics/2015(06)07

Tsantali, E., Economidis, D., & Tsolaki, M. (2013). Could language deficits really differentiate Mild Cognitive Impairment (MCI) from mild Alzheimer's disease? *Archives of Gerontology and Geriatrics,* 57(3), 263–270. https://doi.org/10.1016/j.archger.2013.03.011

Van Waes, L., & Leijten, M. (2015). Fluency in Writing: A Multidimensional Perspective on Writing Fluency Applied to L1 and L2. *Computers & Composition*, 29.

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior*, 73, 311–319. https://doi.org/10.1016/j.chb.2017.03.050

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, 18(2), 332–339. https://doi.org/10.1037/0882-7974.18.2.332

Williams, K., Holmes, F., Kemper, S., & Marquis, J. (2003). Written language clues to cognitive changes of aging: An analysis of the letters of King James VI/I. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 58(1), 58–61. https://doi.org/10.1093/geronb/58.1.P42

Wilshire, C. E., Lukkien, C. C., & Burmester, B. R. (2014). The sentence production test for aphasia. *Aphasiology*, 28(6), 658–691. https://doi.org/10.1080/02687038.2014.893555

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. Reprinted, New York: Harcourt, Brace, 1933

Based on: Paesen, L. & Leijten, M. (2019) *Name agreement and naming latencies for typed picture naming in ageing adults*, Clinical Linguistics & Phonetics, 33:10-11, 930-948, DOI: 10.1080/02699206.2019.1590734

# Chapter two

## Name agreement and naming latencies for typed picture naming in ageing adults

## ABSTRACT

This baseline study aimed to create a coherent set of images that can be used to describe language decline found in healthy elderly and to compare this to the language change found in the early stages of Alzheimer's disease. To this extend, a typed picture naming task was created, in which visual complexity, age-of-acquisition, frequency and name agreement were controlled for. 76 healthy elderly participated in the test; their data will be used in follow-up studies to compare with cognitively impaired patients. The entire typing process was logged with keystroke logging tools Inputlog and Scriptlog; the obtained results were analysed in light of the typing product (name agreement and object recognition) and the writing process (naming latencies and interkey latencies). Results showed that the latencies increased with age and that the older participants had longer latencies for images with a lower frequency and higher age-of-acquisition. Hence, our results indicate the need to take both the latencies and the typing product into consideration.



Keywords: language decline; keystroke logging; picture naming




## 1. INTRODUCTION

Due to the gradual language impairment in the early stages of Alzheimer disease, research into the nature of those language changes and the difference with normal language decline could provide valuable insight into the onset of the disease and into normal cognitive health of the ageing elderly. To that effect, researchers have often focussed on language production which, in contrast to language comprehension, shows signs of specific age-related impairments (Burke & Mackay, 1997; Shafto, Burke, Stamatakis, Tam, & Tyler, 2007). Older people often complain about their inability to come up with a specific word, even though they are certain they know that word (Burke & Mackay, 1997; Clark-Cotton, Williams, Goral, & Obler, Loraine, 2007; Shafto et al., 2007; Shafto & Tyler, 2014; Silagi, Bertolucci, & Ortiz, 2015). These so-called tip-of-the-tongue errors are not the result of a decline in semantic skills since the word is almost always retrieved somewhat later; they are due to an impaired access to phonological and/or orthographical information (Burke & Mackay, 1997; Clark-Cotton et al., 2007; Shafto et al., 2007). To measure the extent of the changes in word retrieval, researchers have often relied upon picture naming tasks and/or studied the effects of certain word/image characteristics on naming speed. These tasks provide valuable insight in the language production processes that are required to name the prompted image as fast as possible. By manipulating the given prompts, researchers hope to influence naming latencies and naming agreement. In this study we aimed to create a coherent set of images that can be used in follow-up studies to compare the language decline found in healthy ageing controls and the clinical language decline found in the early stages of AD. Therefore, the main aim of this study was to determine how the naming accuracy and latencies are affected by age and variable characteristics so as to create a baseline for further research.

It is important to note that the studies on which we will base ourselves have made use of spoken naming, mostly disregarding the written language production. Nevertheless, studying written language production is crucial to fully grasp both the naming latencies and the cognitive processes underlying the picture naming. Therefore, it is necessary to test the foregoing theories on the written product, keeping in mind that the theories developed for spoken picture naming should not uncritically be applied to written picture naming for reasons of spelling and the change of communicative medium (Torrance et al., 2017). Since the onset of AD is not characterised by motor abnormalities (Cummings & Benson Boston, 1983), follow-up studies will be able to correctly compare the data of AD patients with healthy age-matched controls. In what follows we will translate the existing theories on oral picture naming to an alternative that makes use of written – more specifically typed – picture naming. We will do so by using keystroke logging tools; computer programs that "log and time stamp keystroke activity to reconstruct and describe text production processes" (Leijten & Van Waes, 2013, p359). Our baseline study will focus on the difference in image characteristics, on the language changes that might occur upon ageing, and on the coherence of the results in a longitudinal study. Two participant groups, an age group (AG) with younger adults (AG 50 - 64) and a group with older adults (AG 65+), were tested two times with a time interval of three months (moment1 and moment 2) using a within and between subject design.

### 1.3 Overview of word characteristics

Previous studies have pinpointed four eminent word characteristics that influence naming latencies: age of acquisition, frequency, familiarity and name agreement. A first word characteristic which influences naming accuracy and latencies is the age-of-acquisition (AoA) of words. Early-acquired words can be triggered more rapidly and thoroughly than later-acquired words (Brysbaert, Stevens, et al., 2014; Rodríguez-Ferreiro et al., 2009). Even in healthy age-matched controls early acquired words are named with a higher speed and more accuracy than those words that are acquired later on in life (Brysbaert, Warriner, et al., 2014; Gilhooly & Logie, 1980; Kremin et al., 2001). This effect can be contributed to the organisation of the mental lexicon, in which earlier-acquired words tend to have more connections than later-acquired words. Nevertheless, that is only in the condition that those early-acquired words are also frequently used throughout someone's life and are not only used early on in life (Brysbaert, Stevens, et al., 2014; Lété & Bonin, 2013). Therefore, age-of-acquisition is often studied in interaction with frequency, the second determining word characteristic (Brysbaert, Stevens, et al., 2014). Although some studies did not find any frequency effect on the reaction times (Bonin et al., 2001), more recent studies argue that frequency effects can indeed be found, if the AoA is lower (Barry et al., 1997; Severens et al., 2005). More specifically, reaction times have been found to improve with words that have an early AoA and a high frequency (Scaltritti et al., 2016). Studies that focus on AD showed similar results: AoA had an effect on the naming accuracy, resulting in faster naming latencies when presented with an early-acquired word (Garrard et al., 2005; Rodríguez-Ferreiro et al., 2009). Frequency did predict picture naming in AD. Naming accuracy improve when words occur more in the language (Garrard et al., 2005; Kremin et al., 2001), and words with a low frequency tend to lead to more tip-of-the-tongue errors for patients (Astell & Harley, 1996; Garrard et al., 2005). These characteristics are closely related to familiarity. Research in both healthy and cognitively impaired patients revealed





that familiarity to a concept helps to speed naming (Brysbaert, Stevens, et al., 2014; Garrard et al., 2005; Kremin et al., 2001; Rodríguez-Ferreiro et al., 2009; Scaltritti et al., 2016; Zannino et al., 2010). However, this effect can be minimalised due to the concept's high relatedness to AoA and frequency (Brysbaert, Stevens, et al., 2014; Brysbaert & Cortese, 2011). A very familiar concept is often a concept learned at a very young age and that is encountered frequently in daily life (Brysbaert, Stevens, et al., 2014). A fourth influencing characteristic is name agreement, which refers to the extent to which different people agree on the name of a certain thing. In studies that focus on either healthy or cognitively impaired participants, higher naming accuracy and faster latencies were found for words with a higher naming agreement (Garrard et al., 2005; Kremin et al., 2001; Rodríguez-Ferreiro et al., 2009; Scaltritti et al., 2016; Severens et al., 2005).

### 1.4 Overview of image characteristics

Naming latencies in picture naming tasks are not only influenced by the characteristics of the targeted word. They are also influenced by the image characteristics, such as visual complexity, colour and what is depicted. In general, researchers argue that both healthy and cognitively impaired patients have greater difficulty naming living things than man-made objects (Adlington et al., 2009; Duarte & Robert, 2014), even though AD patients are thought to have a deficit in both categories (Adlington et al., 2009). Some studies attribute the advantage for non-living things to the lower number of features that are needed to distinguish between the different man-made things (Duarte & Robert, 2014). The naming of living things would benefit from more features and surface detail (Adlington et al., 2009). In that respect colour showed to have a positive effect on naming latencies, especially if the colour was the diagnostic characteristic of the object (e.g., an orange orange or a yellow banana). Therefore, researchers argue that the recognition of living things benefits from the addition of colour; non-living things have a lower colour diagnosticity and do not need the addition of colour to be easily recognised (Adlington et al., 2009; Moreno-Martinez & Rodriguez-Rojo, 2015; Rodríguez-Ferreiro et al., 2009; Zannino et al., 2010). Since cognitively impaired patients often suffer from a visual impairment, an increase in colour contrasts even aids AD patients leading to comparable results with healthy controls (Adlington et al., 2009). Nevertheless, this effect cannot be replicated when adding photographic detail (Zannino et al., 2010). For the remainder of this article we will discuss how we interpreted the existing theories for the purpose of our study, how we collected the data and what the language product and process data looked like.

## 1.5 Selection of images for the picture naming task

For the selection of the images, we opted for images that meet strict criteria, based on existing theories described in the previous section. Those images are divided into two categories: the first category, 'highly relatable images' (HR), will contain 'easier' images; the second, 'Less relatable images' (LR), will contain more complex images in order to account for both floor and ceiling effects in follow-up studies with AD patients (Forbes-McKay et al., 2014). Even though the names HR and LR do not do justice to the depth and extend of those categories, those terms will be used in the remainder of this article for convenience's sake.

**Table 1**

*Overview Of Picture Selection Criteria 1 To 5*

| | | Highly Relatable | Less relatable |
|---|---|---|---|
| 1 | Name agreement (%) | > 50% | > 50% |
| 2 | AoA (years) | < 6 (young) | > 6 (older) |
| 3 | Frequency (% per million words) | > 30 | < 10 |
| 4 | Reaction Times (ms) | 631,5 < X < 981,34 | 652 < X < 1191,48 |
| 5 | Occurence in other studies | Ignored | Accounted for |

For the image selection, we used the picture naming norms described by Severens and colleagues (2005) for 590 pictures in Dutch. We subjected the images of their study to several elimination rounds, always distinguishing between the HR and LR categories. We first selected only those images that are easily nameable (1) and imageable in order to allow for an increased naming speed and accuracy. Then we selected images on AoA (2) and the frequency (3) with which they appear in Dutch. For the remainder of the images we evaluated the mean spoken reaction times (4). We selected those images that could be found within two standard deviations of the mean latencies per category to create a coherent set of images and to discard outliers. Keeping in mind that some AD patients in our follow-up study could have been subjected to other picture naming studies (e.g., Boston naming test), we decided to eliminate images that appear in other tests (5) for the LR category and keep the images of the HR category, since we presume that participants are more accustomed to the latter. An overview of the five selection criteria is provided in Table 1.





After this first selection of images, we decided to eliminate words that are difficult to type (6) such as words containing a diaeresis (e.g., 'pinguïn' - penguin in Dutch). The remaining images were divided into several sub-categories, distinguishing between living things and man-made objects (7).

**Table 2**

*An Overview Of The Selection Criteria Based On Variable 7.*

| | Highly Relatable | Less relatable |
|---|---|---|
| Living | Humans – 0 | Humans – 5 (*4) |
| | Animals – 5 | Animals – 10 |
| Non-living | Vehicles – 5 | Vehicles – 5 |
| | Objects – 10 (*9) | Objects – 10 |

*Note*: Due to pre-testing results we deleted two images; the affected categories and their new image count are marked with a '*')

The resulting selection consisted of 50 images; these images were redrawn for reasons of visual complexity (8) and coherence in a longitudinal design. We deleted actions from the images, deleted 'decorative' objects, made images of the same sub-category 'look' in the same direction and added colour (9). A few examples of the new images are provided below (cf. figures 1 to 3).

**Figure 1**

*The Dolphin (Dolfijn) ©*

**Figure 2**

*The Shark (Haai) ©*

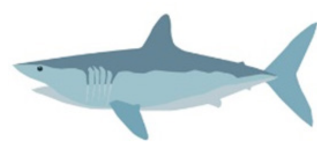


**Figure 3**

*The Pirate (Piraat) ©*

After carefully pre-testing those images in a typed picture naming experiment on 200 students and healthy elderly, we deleted two images due to insufficiently high name agreement (<50%). All images that were created for the purpose of this study will be placed in the 'Open Linguistic Picture Database' or OLPD, and can be found in **Appendix A**.

## 2. AIM AND HYPOTHESES

This study aimed to create a coherent image set that can be used to distinguish healthy language decline from pathological language decline upon ageing. These images were studied in light of the naming accuracy (name agreement & object recognition) and latencies (naming latencies & interkey latencies).

*Naming accuracy*

- We hypothesize that name agreement & object recognition will be equal for the HR and LR categories
- We hypothesize that name agreement & object recognition will not differ between the two age groups (AG 50 – 64 and AG 65+)

*Latencies*

- We hypothesize that naming & interkey latencies will be longer for LR than HR nouns
- We hypothesize that naming & interkey latencies will be longer for the older age group (AG 65+)
- We hypothesize that naming & interkey latencies will remain the same with a three-month interval

*Naming accuracy x latencies*

We believe that name agreement can be used to predict naming & interkey latencies due to its high correlation with the latencies.



## 3. METHOD

### 3.1 MATERIALS AND PROCEDURE

Each participant followed the same procedural steps during both test moments. Firstly, they had to complete the (1) typing test (Van Waes et al., 2017), followed by the (2) picture naming test, a (3) questionnaire, the (4) Montreal Cognitive Assessment (MoCA), and the (5) Geriatric Depression Scale (GDS). The typing test and picture naming test were administered on the researcher's computer on which two keystroke logging tools were installed: tools are programs that log and time stamp every keyboard activity. The tools were respectively Inputlog 7.05 (Leijten & Van Waes, 2013) and Scriptlog (Frid et al., 2014).

#### *3.1.1 Typing test*

In order to account for interpersonal differences in typing speed, we decided to subject all of our participants to a typing test. One specific module of Inputlog 7.0 is the 'copy task', which consists of different typing tests in which specific words/letters/sentences – made up of specific bigram combinations - need to be copied. The assignment is presented left on the screen; hence the letters/words/sentences can be copied and do not need to be remembered, keeping the cognitive load to a minimum (Leijten & Van Waes, 2013; Van Waes et al., 2017). Seven consecutive assignments need to be fulfilled: (1) Repetition of two letters for a time span of 15 seconds, (2) a sentence repetition task for a time span of 30 seconds, (3 to 6) copying a combination of two/three words seven times, (7) copying four blocks of six consonants. A final questionnaire enquired after the handedness of the participant.

#### *3.1.2 Picture Naming Test*

A picture naming test was created, using 48 coloured images as described in the introduction. The keystroke logging tool Scriptlog (Frid et al., 2014) was used to log and time stamp the entire writing process. Before commencing the picture naming task, the participants were instructed to name the prompted picture by typing. They were told it was crucial that they used only one word to name the image (no adjectives, no articles) and that they had to name the image as fast as they could. Every participant was given three images as a practise session to make sure they understood the instruction and afterwards a random combination of the 48 images (not including those three practise images) followed.

### 3.1.3 Questionnaire

With the use of a questionnaire we enquired after the background of the participants. Participants had to answer questions on their studies, job, possible visual impairments, language impairments, bilingualism and possible disorders.

### 3.1.4 MoCa

The Montreal Cognitive Assessment (Nasreddine et al., 2005) is a cognitive screening tool developed to differentiate healthy cognitive ageing from Mild Cognitive Impairment (MCI) - the clinical stage before dementia. We decided to opt for this test rather than its famous counterpart the MMSE (Folstein et al., 1975) for three main reasons: the shorter administration time, its higher sensitivity to the earliest stages of AD and its high reliability in re-testing within a period of three months (Nasreddine et al., 2005). During the test, tasks are given on short-term memory recall, visuospatial abilities, executive functions, attention, concentration, working memory, language and orientation to time and place (Nasreddine et al., 2005). We administered the Dutch version 7.1 on the first encounter with our participants and version 7.2 on the second. In order to be considered cognitively healthy, a participant needs to achieve a score of 26 out of 30 or higher. A score between 18 and 25 out of 30 indicates that a participant might suffer from mild cognitive impairment.

### 3.1.5 GDS

The Geriatric Depression Scale is a questionnaire developed for screening especially elderly on possible signs of a depression (Sheikh & Yesavage, 1986). In order to keep the total test time for our study to a minimum, we decided to select the version containing 15 questions. A score of 6 out of 15 or more indicates that the participant shows signs of a depression.



## 3.2 PARTICIPANTS

To form the base-line of this study, we recruited 76 healthy participants; 68 of them completed the two contact moments (moment 1 and moment 2 which was approximately 3 months later). They were contacted through several elderly organisations, the network of other participants and the researcher's own network. Participants were native Dutch speakers, aged 50 or older on the moment of the first test and had to be sufficiently computer literate. Participants were excluded if they had a history of neurological/psychological illnesses and they were not allowed to have a writing disability. Participants were tested individually at a location of their choice – mostly their home. After careful selection of the participants, we used the data of 62 participants. Note that eight of those 62 participants had a lower MoCA score on one of the two test moments, leading to usable data of 58 participants for moment 1 and 60 for moment 2 (see table 3). Participants with both a positive and a negative MoCA score on the two test moments will be excluded for some analyses (cf. statistical analysis). We divided the participants into two age groups based on the retirement age in Belgium, which is 65: a younger age group (AG 50 - 64) with people aged between 50 and 64 and an older age group (AG 65+) with people aged 65+.

**Table 3**

*Description Of The Demographic Data Of The Participants In Mean (Sd) On The First Contact Moment And Three Months Late*

| | Moment 1 (n = 58) | | Moment 2 (n=60) | |
|---|---|---|---|---|
| | AG 50 - 64 (n=34) | AG 65+ (n=24) | AG 50 - 64 (n=30) | AG 65+ (n=30) |
| Age (years) | 56.88 (3.22) | 70.63 (5.36) | 56.80 (3.20) | 69.77 (4.96) |
| Years of education | 14.94 (2.86) | 13.54 (2.75) | 15.33 (2.58) | 13.17 (2.60) |

*Note*: AG= Age Group

## 3.3 STATISTICAL ANALYSIS

The data that was logged with Scriptlog (Frid et al., 2014) was transformed into a file that can be analysed with the keystroke logging tool Inputlog 7.05 (Leijten & Van Waes, 2013; free download for research purposes on www.inputlog.net). The latter tool allows for more detailed analysis of the naming latencies via the so-called Word Pause analysis, which enabled us to define the words that were written, the naming latencies (pauses before the nouns) and the interkey latencies (pauses between letters). We selected the relevant variables (cf. Table 4) and used SPSS for further analyses.

**Table 4**

*The Selected Naming Accuracy And Latency Variables For The Analysis Of Written Picture Naming.*

| Variables | Description |
|---|---|
| Name agreement | The percentage of words that are named correctly in a strict sense, not including variants or synonyms. |
| Object recognition | The percentage of words that are named correctly in a wider sense: including synonyms, morphological variants (e.g., dolphin – dolphins), more specific names (e.g., Dalmatian instead of dog) and more general names (e.g., boat instead sailboat) |
| Naming latency nouns | The onset reaction time needed by the participant before typing the target noun. This time (in ms) is the 'x' in the examples below |
| Interkey latency | The latencies between the letters of the target noun, divided by the number of letters to account for differences in word length. This can be exemplified by the latencies between the letters for the noun 'ball' (= y1 + y2 + y3) versus 'wheelbarrow' (= y1 + ... + y10); those results cannot be compared due to the number of letters that differ. In order to make those results comparable, the interkey latencies must be calculated divided by the number of letters, which leads to an interkey latency for "ball" of ((y1 + y2 + y3)/3) ms and for "wheelbarrow" of ((y1 + ... + y 10)/10) ms. |
| Examples | _b_a_l_l (x y1 y2 y3) _w_h_e_e_l_b_a_r_r_o_w (x y1 y2 y3 y4 y5 y6 y7 y8 y9 y10) |

*Note*: in Inputlog the 'Naming latency' is referred to as a 'Within word pause' and the 'Interkey latency' is known as 'Between word pause'





### *3.3.6 Data reduction*

In total, we expected 6912 nouns to be typed over the course of the two picture naming tasks; 48 images were named by 76 participants during the first moment and by 68 participants the second time. However, due to technical errors, only 6854 nouns were typed. These data were also produced by participants whose MoCA or GDS scores violated the aforementioned requirements on one or both test moments. We deleted their data and used the resulting data of 58 healthy elderly on the first moment and 60 on the second moment. Furthermore, not all of these nouns were target nouns; some participants added adjectives (e.g., black cat) and even entire descriptions (e.g., pirate with wooden leg and eye patch), thus increasing the word count. Therefore, nouns were excluded if they were proceeded by an adjective or description. Nouns that were followed by a description were kept for some analyses since there were no proceeding words to influence the naming latency of the target noun. The remaining data consisted of 5469 nouns.

These nouns were divided into six categories (cf. Table 5):

1. Named correctly: images that were named correctly without typing/spelling mistakes.
2. Named correctly – with typo correction: images that were named correctly; a typing/spelling mistake was made during the writing process and this error was corrected
3. Named correctly – containing typo: images that were named correctly; a typing/spelling mistake can be found in the final word product (e.g., 'sharrk' instead of 'shark')
4. Synonyms: images that were recognised correctly but given an alternative name or morphological variant (e.g., cat = 'kat' or 'poes' or 'katje')
5. Empty answers: images that were not or barely named due to technical errors, by pressing the 'enter' button too quickly by accident or by pressing the 'enter' button while correcting a mistake in the middle of the word
6. Incorrect names: images that were given the wrong name (e.g., typing 'fish' instead of 'shark') or adjectives

**Table 5**

*The Percentage Of Words On The Total Of 5469 Words - Per Score, Test Moment And Age Group.*

| | | AG 50 - 64 | | AG 65++ | | |
|---|---|---|---|---|---|---|
| | | M1 | M2 | M1 | M2 | Total |
| 1 | Named correctly | 23.42 | 19.91 | 16.15 | 16.95 | 76.43 |
| 2 | Named correctly - with typo correction | 2.10 | 1.98 | 1.24 | 1.35 | 6.67 |
| 3 | Named correctly – containing typo | 0.84 | 0.71 | 0.75 | 0.49 | 2.80 |
| 4 | Synonyms | 2.58 | 2.08 | 2.21 | 2.03 | 8.91 |
| 5 | Empty answers | 0.15 | 0.46 | 0.33 | 0.35 | 1.28 |
| 6 | Incorrect names | 1.02 | 0.38 | 1.63 | 0.88 | 3.91 |

*Note*: AG= Age Group, M1 = Moment 1; M2 = Moment 2

Depending on the analyses, we in- or excluded some of these scores:

- When looking at the product variable 'Name Agreement', we only took words into consideration that were named correctly, independent of typing/spelling errors (85.9% of words included);
- The product variable 'Object recognition' was be calculated using the data from the correctly named words (independent of errors) and synonyms (94.81% of words included);
- The naming latencies for the different words (before word pause), indicating how long it took participants to start typing the word, were only measured for words that were named correctly, independent of typing/spelling errors (85.9% of words included);
- The time it took participants to write a certain word (interkey latencies) – taking into account their personal typing speed and the number of letters per word – could only be calculated for words that were named and typed completely correct. Words that contain(ed) typing/spelling errors were discarded (76.43% of words included);
- The difference in naming latencies before words between the two tests moments was only be calculated for participants that enrolled in both test moments and had a sufficient MoCA score both moments. Words that were named correctly, independent of spelling, were included (85.9% of words included);
- The difference in naming latencies within words per letter between the two different tests moments was only be calculated for participants that enrolled in both moments and had a sufficient MoCA score both moments. Words that were named correctly – without typing/spelling mistakes in the process and/or product – will be included (76.43% of words included).



# 4. RESULTS

## 4.1 Naming accuracy



In order to measure the naming accuracy for a certain object/animal/human/vehicle we made a distinction between two variables. On the one hand, there is name agreement where we only allowed those objects that were named completely correctly, independent of typing/spelling errors. On the other hand, there is object recognition, a variable in which synonyms were also taken into consideration.

**Table 6**

*Overview Of The Naming Accuracy Results*

| | Highly Relatable | Less Relatable | *p* |
|---|---|---|---|
| Name Agreement (%) (N = 4759) | 82.47 (15.18) | 84.54 (13.78) | 0.627 |
| Object recognition (%) (N = 5246) | 95.09 (1.28) | 91.73 (4.93) | 0.006* |

'*Name agreement will be equal for the two categories HR and LR*'. We presumed that name agreement would be equal for both categories since it was one of the selection criteria. Results affirm this hypothesis, $p = 0.627$ (cf. table 6). More specifically, a Mann-Whitney test showed no differences within the age groups for the HR and LR words; AG 50 - 64: $U = 188.50$, $p = 0.065$, $r = 0.267$; AG 65+: $U = 243.00$, $p = 0.491$, $r = 0.099$, (cf.: table 7).

'*Name agreement will not differ between the two age groups, for neither HR and LR*'. A paired-samples t-test indicated that there was no difference in name agreement between the younger age group and the older age group, $t(18) = 1.63$, $p = 0.120$, $d = 0.37$ for HR images. However, a significant difference for LR images was found between the first age group and the older age group, $t(28) = 2.367$, $p = 0.025$, $d = 0.44$. This result indicates that the 65+ year olds did have more difficulty in correctly naming the LR images compared to their younger counterparts.

**Table 7**

*Name Agreement Scores Both Within And Between The Two Age Groups In %*

| | AG 50 - 64 | AG 65+ | *p* |
|---|---|---|---|
| Highly Relatable | 81.33 (14.33) | 83.82 (16.87) | 0.120 |
| Less Relatable | 86.32 (12.55) | 82.44 (16.44) | 0.025 |
| *p* | 0.065 | 0.099 | |

*Note*: AG= Age Group

'*Object recognition will be equal for the two categories HR and LR*'. The general object recognition scores as shown in table 6 indicate that there a significant difference in object recognition between HR images and LR images, $p = 0.006$. In addition to those results, nearly but no significant differences were found in within each age group: AG 50 - 64 ($U = 187.500$, $p = 0.053$, $r = 0.280$) and a significant difference between high and low frequency words was found for AG 65+ $U = 131.000$, $p = 0.002$, $r = 0.448$ (cf.: table 8).

'*Object recognition will not differ between the two age groups, for neither HR and LR*'. With the paired-samples t-test, no significant differences were found for LR words $t(28) = 1.046$, $p = 0.304$, $d = 0.194$. However, contrary to the results of the name agreement, there was a significant difference in naming accuracy for HR images $t(18) = -3.327$, $p = 0.004$, $d = 0.764$.

**Table 8**

*Object Recognition Between The Two Age Groups In %*

| | AG 50 - 64 | AG 65+ | *p* |
|---|---|---|---|
| Highly Relatable | 94.16 (1.15) | 96.20 (2.43) | 0.004* |
| Less Relatable | 92.30 (3.39) | 91.06 (7.84) | 0.304 |
| *p* | 0.053 | 0.002* | |

*Note*: AG= Age Group



## 4.2 LATENCIES

When comparing the naming and interkey latencies in a between-subject design, it is important to take the computer literacy and hence the personal typing speed of all participants into consideration. The analysis of the typing task, particularly the analysis of bigram intervals as provided by Inputlog 7.05 (Leijten & van Waes, 2013), provided a measure that exemplifies the typing literacy of the participants (cf. table 9).



A Mann-Whitney test revealed that there is no significant difference in typing speed between the two groups and the two test moments; moment 1: $U = 421.5$, $p = 0.173$, $r = 0.169$; moment 2: $U = 436.0$, $p = 0.242$, $r = 0.145$. Since the typing speed of the two groups is not significantly different, we decided not to take the typing speed into consideration in further analyses of the latencies.

**Table 9**

*A Measure Of Typing Speed In Ms*

| | AG 50 - 64 | AG 65+ | *p* |
|---|---|---|---|
| Moment 1 | 207 (53) | 240 (84) | 0.242 |
| Moment 2 | 207 (52) | 240 (85) | 0.173 |

*Note*: AG= Age Group

'*Naming & interkey latencies will be longer for LR than HR nouns*'. Significant differences in naming latencies were observed within both AG 50 - 64 $U = 660650.000$, $p < 0.001$, $r = 0.122$ and AG 65+ $U = 445890.500$, $p < 0.001$, $r = 0.163$. Furthermore, a significant difference in interkey latencies was observed within AG 50 - 64 $U = 438273.000$, $p < 0.001$, $r = 0.258$ and AG 65+ $U = 360877.000$, $p < 0.001$, $r = 0.175$ (cf. table 10).

**Table 10**

*The Naming And Interkey Latencies In Ms Contrasting The Two Age Groups And Image Categories.*

| | | AG 50 - 64 | AG 65+ | *p* |
|---|---|---|---|---|
| Naming latencies | Highly Relatable | 1611 (833) | 1901 (1506) | 0.000* |
| | Less Relatable | 1845 (1322) | 2261 (1747) | 0.000* |
| Interkey latencies | Highly Relatable | 212 (140) | 241 (137) | 0.000* |
| | Less Relatable | 254 (107) | 276 (141) | 0.000* |

Note: AG= Age Group

‘*Naming & interkey latencies will be longer for the older age group*’. When looking at the naming latencies, results of the Mann-Whitney test show that there is a significant difference for the LR images *U* = 740842.000, *p* <0.001, *r* = 0.227 and HR images *U* = 340865.500, *p* <0.001, *r* = 0.169 (cf. table 10). Looking at the data of interkey latencies, we see the following results with a Mann-Whitney test: in the naming of LR images, the interkey latencies differed significantly *U* = 714278.500, *p* <0.001, *r* = 0.072; the same goes for HR images *U* = 314575.00, *p* <0.001, *r* = 0.124 (cf. table 10).

‘*Naming & interkey latencies will remain the same with a three-month interval*’. Since the previously described results might be induced by our repeated testing with an interpose of three months, the data of the two test moments will also be compared with a Mann-Whitney test (cf. table 11). For AG 50 - 64, there was no significant difference in naming latencies neither for LR nouns *U* = 289713.000, *p* = 0.264, *r* = 0.028 nor for HR nouns *U* = 120557.500, *p* = 0.390, *r* = 0.027. Similar results can be found for AG 65+, with the difference that for both LR images *U* = 198359.500, *p* = 0.057, *r* = 0.053 and HR images *U* = 83612.000, *p* = 0.071, *r* = 0.062 the results are nearly significant.

Mann-Whitney test revealed that there is no significant difference in interkey latencies in AG 50 - 64 for both LR images *U* = 224496.00, *p* = 0.135, *r* = 0.040 and HR images *U* = 103869.500, *p* = 0.951, *r* = 0.002. The opposite is true for AG 65+, for both LR images *U* = 143902.500, *p* = 0.013, *r* = 0.074 and HR images *U* = 71981.500, *p* = 0.026, *r* = 0.078 (cf. table 11).

**Table 11**

*Statistics Per Age Group, Per Moment, Per Frequency Type In Ms.*

| | | | Moment 1 | Moment 2 | *p* |
|---|---|---|---|---|---|
| Naming latencies | AG 50 - 64 | Highly Relatable | 1728 (891) | 1769 (2181) | 0.390 |
| | | Less Relatable | 1961 (1422) | 1844 (1003) | 0.264 |
| | AG 65+ | Highly Relatable | 2166 (1751) | 2004 (1564) | 0.071 |
| | | Less Relatable | 2635 (2784) | 2542 (4158) | 0.057 |
| Interkey latencies | AG 50 - 64 | Highly Relatable | 207 (90) | 217 (180) | 0.135 |
| | | Less Relatable | 260 (115) | 247 (96) | 0.951 |
| | AG 65+ | Highly Relatable | 244 (115) | 239 (151) | 0.026 |
| | | Less Relatable | 283 (139) | 271 (142) | 0.013 |

*Note*: AG= Age Group



## 4.3 NAMING ACCURACY X LATENCIES

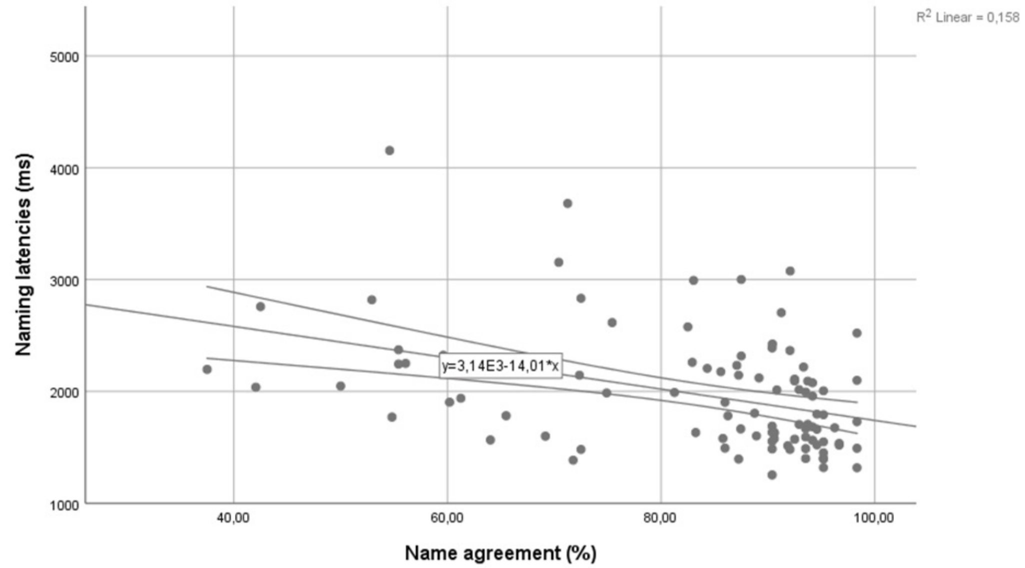
In order to measure if latencies can be predicted with the use of the name agreement, a Pearson correlation coefficient was computed.

‘Naming & interkey latencies can be predicted using name agreement’. There was a negative correlation between the two variables for naming latencies, *r* = -0.397, n = 96, *p* < 0.001. This result is summarized with a scatterplot (cf. figure 4). No significant correlation was found for interkey latencies: *r* = -0.139, n = 96, *p* = 0.177.



**Figure 4**

*The Correlation Between Name Agreement And Naming Latencies*



# 5. DISCUSSION

With this study we aimed to create a coherent set of images that can be used to distinguish healthy cognitive ageing from pathological ageing. Therefore, the influence of age and image characteristics on naming accuracy and naming latencies was studied as a baseline for further research.

## 5.1 NAMING ACCURACY

We studied naming accuracy by distinguishing between name agreement and object recognition; two variables that account for the number of images that are named correctly in either a strict sense (name agreement) or a wider sense (object recognition) in which synonyms and morphological variants were allowed. Since we controlled for name agreement upon selecting the images from previous studies, we presumed that neither image characteristics nor increasing age would trigger differences in naming accuracy. Image characteristics did not influence the name agreement nor object recognition scores within the younger age group. Accordingly, they did not influence name agreement in the older age group. However, this older age group showed lower object recognition scores for LR images compared to HR images. We attribute this effect to the higher number of synonyms that were found for the HR images. With respect to the influence of age, the results are more ambiguous to interpret. Whereas the name agreement scores did not differ for HR images, the object recognition scores indicated that the younger age group did have more difficulty with those images than the older age group. Surprisingly, the name agreement scores for LR images showed greater difficulty for the older age group compared to the younger, a result that was evened out when looking at object recognition. When disregarding the results of the object recognition scores, our results are in line with previous studies where healthy ageing elderly tended to have more tip-of-the-tongue errors in low frequent words than in high frequent words (Clark-Cotton et al., 2007). Moreover, our study adds that it is only with increasing age that word retrieval issues for Less relatable (and thus lower frequent) words arise. Since AD patients in the early stages of the disease show an increased impairment in word retrieval, we believe that the discrepancy in naming accuracy between the HR and LR images will only grow. The issues encountered in object recognition scores can be contributed to the differences in synonyms and morphological variants used. Therefore, we would suggest that name agreement is a more stable determiner of differences in picture naming than object recognition.



## 5.2 LATENCIES

We presumed that both image characteristics and age would influence the latencies (naming and interkey), with a stable result between different test moments. As expected, image characteristics were correctly predicted to influence latencies, with faster reaction times for the easier images (HR vs LR). Furthermore, latencies were also correctly hypothesized to increase with age, even though the typing speed of the two groups was comparable. This finding is consistent with previous findings on spoken picture naming and is further proof of the increasing difficulty healthy ageing elderly have with lexical retrieval (Kemper & Sumner, 2001; Shafto & Tyler, 2014). Furthermore, our results are also in line with studies that have found an effect of AoA and frequency on the naming latencies in picture naming tasks (Barry et al., 1997; Boukadi et al., 2016; Levelt et al., 1999; Scaltritti et al., 2016). With respect to the change in latencies over time, our results are less straight forward. The younger age group showed no difference in naming and interkey latencies between the two moments. Accordingly, the older age group showed similar naming latencies, indicating that naming latencies are a stable measure to use in longitudinal research. The interkey latencies, however, triggered faster results in the older age group for the second test moment. Possibly, we triggered a learning effect; once the required word form is retrieved, the processing within the word takes less time because the word has been used before in the previous test session.

With respect to future research, we tested the correlation between latencies and name agreement in order to find out if name agreement can be used to predict naming latencies. Our results indicate that name agreement can be used as a predictor for naming latencies, with faster reaction times to words with a higher name agreement. Naming latencies are also found to be influenced by name agreement in other studies; with faster reaction times for words that were both HR and had a low AoA, which we manipulated in our image selection (Scaltritti et al., 2016). This finding could not be replicated for interkey latencies, possibly triggered by the ongoing search for the correct orthographical form for the retrieved word.

## 5.3 IMPLICATIONS

This baseline study set out to create a coherent set of images for a typed picture naming task that can be used in future research to differentiate healthy cognitive ageing from pathological ageing. To that extend, our results imply that even in healthy ageing, differences in typed picture naming can be found. Therefore, we find it necessary never to treat the healthy controls group as a whole, but rather as a set of individuals with different ages. By logging the entire writing process we were able to provide additional and more in-depth details to the cognitive processes that are required to name the prompted images. These latencies are needed in addition to the naming accuracy data. We believe that for studies with a longitudinal design, naming latencies provide a stable measure in differentiating between AD patients and healthy controls. However, the change in interkey latencies must be studied more thoroughly in future studies, since it provides more details on the ongoing writing process and the difficulty encountered during this process. It is important to stress the exploratory nature of this study. We found differences in latencies both between the different image categories and age groups. This implies that there is a change in language capacities in healthy ageing elderly; further studies with cognitively impaired patients are needed to define more clearly which language change is still to be considered healthy and which is pathological.

## 6. ACKNOWLEDGEMENTS

This research was supported by the research fund BOF-DOCPRO 2015, PS ID: 31661 of the University of Antwerp. We would also like to thank Eric Van Horenbeeck, the technical coordinator of Inputlog, and Luuk Van Waes, co-founder of Inputlog and the typing task. We thank Johan Frid and Victoria Johansson, respectively the technical coordinator and co-founder of Scriptlog. Further thanks go to Peter Mariën (†) and Dorien Vandenborre for their input during the design of the study. Last but not least we would like to thank the elderly who enthusiastically participated in this study.

## 7. DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.



## 8. BIBLIOGRAPHY

Adlington, R. L., Laws, K. R., & Gale, T. M. (2009). Visual processing in Alzheimer's disease: Surface detail and colour fail to aid object identification. *Neuropsychologia*, 47(12), 2574–2583. https://doi.org/10.1016/j.neuropsychologia.2009.05.004

Astell, A. J., & Harley, T. A. (1996). Tip-of-the-tongue states and lexical access in dementia. *Brain and Language*, 54(2), 196–215. https://doi.org/10.1006/brln.1996.0071

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart Pictures: Effects of Age of Acquisition, Frequency, and Name Agreement. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 50(3), 560–585. https://doi.org/10.1080/783663595

Bonin, P., Fayol, M., & Chalard, M. (2001). Age of Acquisition and Word Frequency in Written Picture Naming. *The Quarterly Journal of Experimental Psychology Section A*, 54(2), 469–489. https://doi.org/10.1080/713755968

Boukadi, M., Zouaidi, C., & Wilson, M. A. (2016). Norms for name agreement, familiarity, subjective frequency, and imageability for 348 object names in Tunisian Arabic. *Behavior Research Methods,* 48(2), 585–599. https://doi.org/10.3758/s13428-015-0602-3

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology,* 64(3), 545–549. https://doi.org/10.1080/17470218.2010.503374

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, 150, 80–84. https://doi.org/10.1016/j.actpsy.2014.04.010

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Clark-Cotton, M. R., Williams, R. K., Goral, M., & Obler, Loraine, K. (2007). *Language and Communication in Aging* (pp. 1–8). https://doi.org/10.1016/B0-12-370870-2/00103-7

Duarte, L. R., & Robert, C. (2014). Semantic richness influences naming pictures in old age and in Alzheimer's disease. *Journal of Neurolinguistics,* 31, 55–68. https://doi.org/10.1016/j.jneuroling.2014.06.004

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research,* 12(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Forbes-McKay, K., Shanks, M., & Venneri, A. (2014). Charting the decline in spontaneous writing in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatrica*, 26(04), 246–252. https://doi.

org/10.1017/neu.2014.2

Frid, J., Johansson, V., Johansson, R., Wengelin, Å., & Johansson, M. (2014). Developing a keystroke logging program into a writing experiment environment. *Presentation at Writing Across Borders,* WRAB, Paris February 2014.

Garrard, P., Maloney, L. M., Hodges, J. R., & Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain*, 128(2), 250–260. https://doi.org/10.1093/brain/awh341

Gilhooly, K. J., & Logie, R. H. (1980). METHODS &DESIGNS Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, 12(4), 395–427. https://doi.org/10.3758/BF03201693

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging,* 16(2), 312–322. https://doi.org/10.1037//0882-7974.16.2.312

Kremin, H., Perrier, D., Wilde, M. De, Dordain, M., Bayon, A. Le, Gatignol, P., Rabine, C., Corbineau, M., Lehoux, E., & Arabia, C. (2001). Factors predicting success in picture naming in Alzheimer's disease and primary progressive aphasia. *Brain and Cognition*, 46(1–2), 180–183. https://doi.org/10.1006/brcg.2000.1270

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication*, 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Lété, B., & Bonin, P. (2013). Does frequency trajectory influence word identification? A cross-task comparison. *Quarterly Journal of Experimental Psychology*, 66(5), 973–1000. https://doi.org/10.1080/17470218.2012.723725

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(01), 1–75. https://doi.org/10.1017/S0140525X99001776

Moreno-Martinez, F. J., & Rodriguez-Rojo, I. C. (2015). On colour, category effects, and Alzheimer's disease: A critical review of studies and further longitudinal evidence. *Behavioural Neurology,* 2015, 1–12. https://doi.org/10.1155/2015/960725

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment , MoCA : A Brief Screening. *Journal of the American Geriatrics Society,* 53(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Rodríguez-Ferreiro, J., Davies, R., González-Nosti, M., Barbón, A., & Cuetos, F. (2009). Name agreement, frequency and age of acquisition, but not grammatical class, affect object and action naming in Spanish speaking participants with Alzheimer's disease. *Journal of Neurolinguistics,* 22(1), 37–54. https://doi.org/10.1016/j.jneuroling.2008.05.003

Scaltritti, M., Arfé, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition*, 156, 16–29. https://doi.org/10.1016/j.cognition.2016.07.006



Severens, E., Lommel, S. Van, Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, 119(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shafto, M. A., & Tyler, L. K. (2014). Decline and Preservation. *Science*, 346(6209), 583–588. https://doi.org/10.1126/science.1254404

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 5(1–2), 165–173. https://doi.org/10.1300/J018v05n01_09

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior,* 73, 311–319. https://doi.org/10.1016/j.chb.2017.03.050

Zannino, G. D., Perri, R., Salamone, G., Di Lorenzo, C., Caltagirone, C., & Carlesimo, G. A. (2010). Manipulating color and other visual information influences picture naming at different levels of processing: Evidence from Alzheimer subjects and normal controls. *Neuropsychologia*, 48(9), 2571–2578. https://doi.org/10.1016/j.neuropsychologia.2010.05.003

Will be submitted to: *International Journal of Speech Language Pathology*

# Chapter three

## Ageing and sex differences in the cohesion of written narratives

## ABSTRACT

Purpose – There is strong evidence that cohesion in narratives changes when people get older. We set out to establish a normative dataset based on a cross-sectional study and to develop normative data for future research that focus on Alzheimer's disease and possibly also other clinical studies with degenerative diseases.

Method – Cognitively healthy volunteers (N = 257) completed two narrative writing tasks, additional cognitive screening tests, working memory tests and a questionnaire to inquire after their socio-economic background. The resulting narratives were analysed with T-Scan in light of how cohesive they were. We hypothesized that age would negatively affect grammatical complexity and lexicon and that sex would lead to significant differences in cohesion.

Results – Results indicate that age did not have a negative effect on sentence complexity, with increased distance between subject and verb upon ageing, and an increase in word complexity. A decrease in text and word length was noticeable in our older age groups. In terms of sex differences, women use a higher density of conjunctions, and the density of nouns decreases whereas the density of (abstract) verbs increases.

Conclusion – Cohesion is not negatively affected by ageing; more research into writing processes should provide more insight into potential struggles during the writing process. Sex differences in cohesion are apparent; we suggest to always differ between the sexes upon analysis. The additional normative table created in this study will allow future research, both in Alzheimer's disease and other neurodegenerative diseases, to develop a diagnostic aid.




## 1. INTRODUCTION

An estimate of 50 million people are suffering from some form of dementia worldwide, with Alzheimer's disease occurring as its most prevailing form. Since the risk of suffering from the disease increases with age, the number of patients will only increase as baby boomers are reaching the crucial age of 65. Even though it is often thought to be part of normal ageing, Alzheimer's is a neurodegenerative disease that can start with small changes in memory loss and language issues and will gradually affect the patient's emotions, behaviour, and other cognitive functions. Moreover, women are found to suffer from a greater decline compared to men (Laws et al., 2018). In a study published by the Alzheimer association, 82% of the respondents indicated that they would take a test in order to learn about their potential risk (Alzheimer's Association, 2019; Alzheimer's Disease International, 2019). However, the current diagnostic procedure is rather complex; it is conducted by a team of specialists and includes a variety of cognitive medical tests. This negatively impacts the willingness of interested elderly to go and have themselves checked.

Research into linguistic changes in Alzheimer's disease could aid the diagnostic process since a gradual language impairment can be noticed in the early stages of the disease. By comparing the language changes in AD to the changes that occur upon healthy ageing, valuable insight can be gained into the onset of the disease. Previous studies into healthy ageing have refrained from using language comprehension which is unaffected by age; rather, language production is studied since small age-related impairment can be noticed (Burke & Mackay, 1997; Shafto et al., 2007). One of the most common and noticeable impairments are issues with word recall, the so-called tip-of-the-tongue mistakes which leads to a state in which the person feels the word is 'on the tip of their tongue' (Burke & Mackay, 1997; Clark-Cotton et al., 2007; Shafto et al., 2007; Silagi et al., 2015). The word will be retrieved later on, since it is not the semantic knowledge about the word that has been lost; rather, the delay in retrieval reflects a deficit in the phonological neural system (Burke & Mackay, 1997; Clark-Cotton et al., 2007; Shafto et al., 2007).

### 1.1 Neural changes in healthy ageing

Entrenchment is a central concept to describe the extent to which linguistic networks have been developed in the brain. It is generally described as 'the process of strengthening a speaker's representation of linguistic knowledge through repeated processing, in comprehension as well as production' (Steinkrauss & Schmid, 2016, p. 368). More concretely, this entails that every encounter with a linguistic structure, either through comprehension or production, contributes to existing knowledge of specific linguistic events and, therefore, leads to a deeper entrenchment of that structure. Needless to say, the exposure to and use of the language is crucial for this lifelong cognitive reorganisation. Entrenchment is a usage-based, bottom-up process: linguistic structures and similarities between them contribute to a deeper entrenchment of certain constructions (H. J. Schmid, 2007, 2017; Steinkrauss & Schmid, 2016). The ease with which cognitive units are activated, therefore, depends on the extent to which they were previously encountered and thus how familiar they are. For example: upon the encounter with 'a cat' and 'a cockatoo', recognition times will be longer for the latter animal. Processes that link the target conceptualisation to its phonological unit are different for both animals, predictably with a better entrenchment for the conceptual unit 'cat' compared to 'cockatoo' (H. J. Schmid, 2007). Apart from entrenchment, rapid transfer of information is needed to accommodate normal motor, sensory and cognitive functioning.

Information transfer is facilitated in the white matter by the insulation around axons, myelin, which allows for impulses to be transferred faster, compared to their unmyelinated counterparts. From birth onwards, the number of oligodendrocytes continues to grow, allowing the myelinisation process to commence in a diffused pattern. Earlier myelination can be found in the motoric and sensory systems, and the maturational effect results in later myelination of the temporal and frontal lobes (Mu et al., 2017; Nasrabady et al., 2018). More specifically, up until the age of 26, no effects of grey matter changes can be found in the temporal lobe. After that age, the number of myelinated axons or white matter will increase, leading to a decrease in grey matter in the temporal lobe compared to the total mass (Mu et al., 2017). From the age of 50 onwards, this process reverses, with an increase in the ratio grey on white matter, which implies that there is a loss in myelinated axons. The areas that were myelinated later on in life tend to be more vulnerable to demyelination; this process known as 'neuropathologic retrogenesis' (Nasrabady et al., 2018; Williamson & Lyons, 2018).

Due to demyelination, together with a general white matter thinning and volumetric shrinkage (e.g., in the prefrontal cortex), the affected structural integrity leads to a transmission deficit upon ageing (Park & Reuter-Lorenz, 2009). These structural changes in the brain go hand in hand with changes in cognitive functioning. For instance, a change in working memory capacity can be noticed; with slower reaction times, less accuracy and an overall decline in ability for the elderly adult (Mogle & Sliwinski, 2013). However, some studies have found that the only



effect of an increased working memory load on syntactic processing resulted in longer reaction times; accuracy remained the same (Alatorre-Cruz et al., 2018). A possible explanation for these increased reaction times is that due to a gradual inhibitory dysfunction, irrelevant information will also be selected and the content stored in the working memory will be inefficiently removed (Park & Reuter-Lorenz, 2009).

Furthermore, dedifferentiation is hypothesized to take place, a process in which dedicated neural circuitry becomes less distinctive in both domain-specific and domain-general cognitive areas. A deficit in the neural resource allocation leads to less precise and trustworthy neural representations (Burianová et al., 2013; Carp et al., 2011; Koen & Rugg, 2019; Park & Bischof, 2013; Park & Reuter-Lorenz, 2009). In order to compensate for the declining circuitry and continue facilitating the transfer of information, spontaneous neuroplasticity will take place. Function non-specific brain areas will be activated in order to aid in the network connectivity (Meunier et al., 2014). Several studies concluded that with age, compensatory bilateral activation (in the prefrontal and other brain areas) will aid the declining task-specific neural circuitry and maintain performance in older adults (Burianová et al., 2013; Park & Reuter-Lorenz, 2009).



Note that the compensatory neural activation serves as an aid to a failing dedicated circuitry. Even though performance will always be negatively affected, the performance would be worse without the compensatory system. Furthermore, the need for compensatory systems is not due to ageing per se. Ageing affects the structural and functional characteristics of the brain which leads to the need for supporting neural activations (Meunier et al., 2014). Nevertheless, due to the notion of cognitive reserve, these ageing effects, and even effects of a neurodegenerative disease do not show up equally among peers. Cognitive reserve refers to the 'protective mechanism that supports sustained cognitive function following damage to the physical brain associated with age, injury, or disease' (Fleck et al., 2017, p.1). In other words, the greater the size of the reserve, the more it allows somebody to cope with the damage from, for instance, a neurodegenerative disease before the symptoms of that disease will show and clinical diagnosis will follow. This cognitive reserve is affected by various socio-economic factors such as education and occupation; however, even sex has been found to play a significant role (Fleck et al., 2017; Laws et al., 2018). Women have been found to have an advantage in verbal memory tasks, which according to Laws and colleagues (2018) could indicate that women benefit from a cognitive reserve that delays verbal memory decline in Alzheimer's disease (AD).

## 1.2 Language changes in healthy ageing

Along the lines of the posited theories described above, structural, and functional specificity will decline with age. The structural change is further accommodated by the growth in corpus callosum (Rosselli et al., 2014); its white matter facilitates rapid conduction of electrical impulses between the two hemispheres. As a result, a change in hemispheric activation for language comprehension and production will occur with ageing. The specific left hemispheric lateralisation in younger adults will gradually become bilateral, resulting in, for instance, activation of the right hemisphere during language production and comprehension tasks. Cognitive tasks have been found to activate the prefrontal lobes of both hemispheres in elderly, whereas there is a more specific lateralisation of one hemisphere in younger adults (Rosselli et al., 2014).

In terms of grammar and ageing, this greater bilateral hemispherical involvement can be noticed. The Nun Study was one of the first longitudinal studies that focused on how language changes, in order to establish possible risk factors for Alzheimer's disease (Boyer, 1996; Kemper et al., 2001). Language samples from 180 nuns were collected; those samples were gathered from biographies written by the nuns during their time in the convent. A nun typically wrote her first biography when she took her vows, between 1931 and 1943, being aged between 18 and 32 years old. The researchers were provided with two samples from 90 nuns, 36 nuns delivered three samples and 20 of them had four. Furthermore, after the year 1990 the nuns were subjected to annual tests; this provided the researchers with information about the cognitive health of the participants. In analysing the language samples, a distinction was made between grammatical complexity and idea density (Kemper et al., 2001). The former represents measurements of sentences that consist of one simple clause and sentences that comprise various subclauses. The latter refers to the number of ideas that can be found within the text sample (Sirts et al., 2017). Together with the bilateral differentiation that occurs upon ageing, a gradual decline in grammatical complexity and idea density was found. Less complex sentences were found in samples that were written later on in life and the texts of the nuns who developed Alzheimer's disease were described as repetitive and vague. Moreover, Kemper and colleagues stated that those nuns already scored lower on idea density and grammatical complexity at a young age (Kemper et al., 2001).

In a study by Le, Lancashire, Hirst and Jokel (2011) the novels of three British authors were studied in search for changes in cohesion by looking at lexical and syntactical characteristics. The first novelist was Iris Murdoch, who showed signs of cognitive decline later on in her life and was



post-mortem found to have had Alzheimer's disease. In a previous study by Garrard, Maloney, Hodges, & Patterson (2005), three of her novels had already been scanned for signs of language deterioration, but their work was criticised later on for their lack of detail and their focus on only parts of the texts. Therefore, Le et al. (2011) decided to focus on 20 of her novels which she wrote between the age of 35 and 76 years old. The two other novelists were Agatha Christie and Phillis Dorothy James (who published under P.D. James) of whom respectively 16 novels (written between the ages 28 and 82) and 15 novels (written between 42 and 82 years old) were analysed. Even though it was never established, Agatha Christie was suspected of having suffered from Alzheimer's disease; Phillis Dorothy James lived a healthy life and did not suffer from any cognitive illnesses (Le et al., 2011).



In terms of syntax, the grammatical complexity was measured with the use of the variable 'D-Level' which indicates how complex a sentence is. Murdoch showed a decrease in grammatical complexity from the age of 50 onwards, whereas with Agatha Christie and James only a small and non-significant decrease was noticeable. Furthermore, the use of the passive voice was measured. Bates et al. (1995) found that Alzheimer patients and the older healthy controls used less passives compared to their younger and healthy controls. This result was also found for Murdoch, whose use of the passive voice decreased significantly with age. Le et al. (2011) concluded that Christie showed the initial signs of Alzheimer's disease in her last novels and that Murdoch already showed a decrease in vocabulary and syntax from the age of 40 onwards and reaching its peak at the age of 60.

Lexically, the vocabulary of the authors was studied with the use of the type/token ratio (TTR) and the word-type introduction rate (WTIR). These respectively represent: 'the number of unique lemmatized word-types divided by the total number of word-tokens' and 'the cumulative number of unique lemmatized types computed at every 10,000-token interval' (Le et al., 2011, p. 440). Phillis Dorothy James showed no differences in vocabulary over the course of time; Agatha Christie showed a gradual decline in her final two books and Murdoch showed an abrupt decrease in vocabulary knowledge in her last novel 'Jackson's Dilemma' (Le et al., 2011). The lexical repetitions were also studied since they might indicate possible word recall issues. Again, no changes were found in the novels of Phillis Dorothy James compared to the statistically significant effect in repetitions in the final novels of Murdoch and Christie. The fact that Phillis Dorothy James showed no changes in vocabulary does not imply that she did not suffer from tip-of-the-tongue errors. However, these studies focused merely on the writing product and not the writing process. Further insight into the writing and revision process that the books go through, could reveal traces of word search. After all, the meaning of the words is retained and the temporary inaccessibility of the phonological form of that word is resolved after a while (Burke & Mackay, 1997; Clark-Cotton et al., 2007; Shafto et al., 2007). A study of word types (nouns, pronouns, content verbs, adjectives) indicated that there was no significant difference to be found in the novels of Phillis Dorothy James; Christie and Murdoch showed a gradual decrease in nouns and increase of content verbs (Le et al., 2011). This result stands in contrast to the study by Garrard et al. (2005), who found no difference in word types between the three novels of Murdoch in their study.

### 1.3 Research questions

In our study we observed participants who conducted two narrative writing tasks. Based on the changes that happen in the neural circuitry and the functional changes that go along with it, we know that there is a shift in cohesive abilities of elderly. Therefore, we posit the following research question: How do ageing and sex affect the cohesion of written narratives? Note that the term 'cohesion' is closely related to and often intermixed with the term 'coherence' between disciplines. The cohesion of a story refers to the grammatical aspects of the written narratives and therefore reflects the relationship between the different sentence elements; the term coherence refers to the more subjective argument development within the text (Min, 2015). For the purpose of this study, we will focus on the grammatical structure or cohesion of texts.

In order to answer our main research question, the following sub questions were composed: '*To what extent does ageing have an effect on the grammatical complexity of a typed narrative?*' and '*To what extent does ageing have an effect on the lexical complexity of a typed narrative?*'. Based on the knowledge that sex, next to education, leads to a difference in cognitive reserve, we also pose the research questions: '*To what extent does sex have an effect on the grammatical complexity of a typed narrative?*', and '*To what extent does sex have an effect on the lexical complexity of a typed narrative?*'.

In line with findings from the nun study, wherein the ageing nuns show a decline in grammatical complexity (Kemper et al., 2001), we formulate the hypothesis (1) that grammatical complexity will decline with age. Based on the tip-of-the-tongue issues that are found due to compensational strategies of neural processes, we propose the hypothesis (2) that lexical complexity will change with age, with a preference for easier structures and formulations. Considering the differences



between male and female participants and their cognitive reserve, we suggest that (3) the grammatical and lexical output will differ significantly between men and women.

In order to answer these research questions, we designed a cross-sectional study, in which participants aged 50 and older were asked to write two narratives on a computer. We decided on the age of 50 as a cut-off point due to the neural changes that occur between the ages of 40 and 50 (Nasrabady et al., 2018; Williamson & Lyons, 2018). An additional baseline set containing narratives from students was gathered. Given that brain development is not completed until the age of 25 (Casey et al., 2008; Sharma et al., 2013), and young onset dementia may already occur at 30 years old (Rossor et al., 2010; Sampson et al., 2004), we decided this participant group should also be included in our study. In analysing those narratives, we will focus on the changes that occur in terms of cohesion (through studying grammar and lexicon) and compare language changes in healthy ageing, and between sexes.



## 2. METHOD

### 2.1 PARTICIPANTS

Trying to depict how language usage changes over the course of a lifetime, we created an experiment in which a total of 257 healthy volunteers participated (cf. Table 1). The majority of this cohort – 218 healthy ageing adults - belong to the target group and are between 50 and 100 years old; they were divided into different decades for this study. Additionally, we created a baseline group consisting of 39 students. The healthy elderly in the target group were recruited through the researcher's own network, as well as with the help of the Free University of Brussels (VUB) master's course Neuro- and psycholinguistics, and through senior organisations. The baseline group were all masters in Multilingual Professional Communication from the University of Antwerp. All participants needed to be native Dutch speakers without any neurological or psychological issues. We excluded participants with dyslexia and insufficient computer literacy. The older participants were tested individually at a location of their choice – mostly at their home; students were tested in group at university. At the start of the experiment, every participant signed a consent form, approved by the ethics committee, and additional information about the study was provided. The study was performed in accordance with the Declaration of Helsinki. Note that within this study, the participants were recruited and divided in two subsets. The first subset only partook in the cross-sectional study; the second subset will partake in an additional future longitudinal study. In what follows, these two subsets will respectively be referred to as the cross-sectional subset and the longitudinal subset.

**Table 1**

*Participants' Demographics Based On Age, Sex And Education*

| | 50 – 59 | | 60 -69 | | 70 - 79 | |
|---|---|---|---|---|---|---|
| | F | M | F | M | F | M |
| N | 16 | 7 | 12 | 11 | 7 | 5 |
| Years of education | | | | | | |
| Mean | 14.50 | 16.00 | 13.92 | 15.54 | 12.43 | 14.00 |
| SD | 2.63 | 3.27 | 1.73 | 2.66 | 3.69 | 3.16 |
| Range | 8-19 | 12-22 | 11-16 | 10-21 | 8-19 | 10-18 |
| MoCA score M1 (/30) | | | | | | |
| Mean | 28.63 | 28.43 | 26.83 | 28.00 | 26.71 | 27.20 |
| SD | 1.31 | 1.40 | 2.12 | 2.00 | 1.38 | 1.30 |
| Range | 26-30 | 26-30 | 24-30 | 24-30 | 25-29 | 26-29 |
| MoCA score M2 (/30) | | | | | | |
| Mean | 28.69 | 28.71 | 28.25 | 28.27 | 28.57 | 27.60 |
| SD | 1.49 | 1.60 | 0.97 | 1.79 | 1.51 | 1.82 |
| Range | 25-30 | 26-30 | 27-30 | 25-30 | 27-30 | 25-30 |
| MoCA score M3 (/30) | | | | | | |
| Mean | 28.50 | 28.43 | 28.00 | 28.00 | 27.86 | 27.20 |
| SD | 1.51 | 1.13 | 1.41 | 1.34 | 1.68 | 1.30 |
| Range | 26-30 | 27-30 | 25-30 | 26-30 | 25-30 | 26-29 |

*Note*: M1= moment 1, M2 = moment 2, M3 = moment 3, M = male, F = female

## 2.2 MATERIALS

### *2.2.1* Narrative writing tasks

In order to elicit coherent stories, we developed Narrative Writing Tasks (NWTs). Those tasks required the participant to write a story based on prompted images. The images all depict a single coloured object/being and were selected from the Open Linguistic Picture Database (OLPD, www.



olpd.eu). The OLPD contains both 'easier' images (e.g., 'dog' and 'car') and more 'complex' images (e.g., 'kangaroo' and 'wheelbarrow'); those categories were made using the word characteristics in Dutch behind the images such as the frequency with which they occur in the language and their age-of-acquisition while both having a high name agreement (for more information see Paesen & Leijten, 2019). For each NWT, a random combination of those images was made for every participant. As a result of this discrepancy between easier and more complex images, two types of NWTs were created. The simple NWT (NWT1) in which a combination of four 'easier' images is made (see Figure 1) and a complex NWT (NWT2) that combines six 'more complex' images (see Figure 2). The discrepancy between the easier and more complex NWT was further accentuated by the perspective of the NWT. The easier NWT had to be written in the first person singular; the complex NWT had to be written in the third person singular and past tense



**Figure 1**

*An Example Of A Possible NWT1, Consisting Of 4 Images (One Animal, One Vehicle, Two Objects)*

### *2.2.2* *Questionnaire*

A questionnaire was constructed, in order to gather information on the participants' background. We enquired after the participants' education, job (if applicable), language and computer skills, possible neuro- or psychological issues and possible visual impairments.

### *2.2.3* *Cognitive screening test*

Two different types of cognitive screening tests were used in order to assure the cognitive health of the participants: the Mini-Mental State Examination (Folstein et al., 1975) for the cross-sectional subset and the Montreal Cognitive Assessment (Nasreddine et al., 2005) for the longitudinal subset. The Mini-Mental State Examination (MMSE) comprises of tests of orientation, registration, recall, calculation and attention, naming, repetition, comprehension, reading, writing and drawing (Cockrell & Folstein, 2002). Participants with a score of 24 out of 30 or higher were considered to be cognitively healthy and were therefore included in our study. The main reasons for switching from the MMSE to the Montreal Cognitive Assessment (MoCA) in the longitudinal subset is the shorter admission time and the higher reliability for re-testing within a period of three months (Nasreddine et al., 2005). The items of the MoCA include tests of short-term memory recall, visuospatial abilities, executive functions, attention, concentration, working memory, language, and orientation. Participants need to score 26 out of 30 or higher to be considered cognitively healthy. A score between 18 and 25 out of 30 indicates that the participant might suffer from mild cognitive impairment. We administered the Dutch version 7.1.

**Figure 2**

*An Example of A Possible NWT2, Consisting of 6 images (One Person, Two Similar Animals, One Vehicle, Two Objects)*

### *2.2.4* Working memory tests

In order to assess the working memory capacities of each participant, working memory tests were used: the forward digit span and the backward digit span (Wechsler, 2008). The tests comprise of an increasing set of numbers that have to be repeated, with two sets of numbers for each 'level'. If a participant is wrong on both sets of the same level, the test is stopped and the next type of digit span task is commenced.

### *2.2.5* Geriatric Depression Scale

The last task was the Geriatric Depression Scale (GDS; Sheikh & Yesavage, 1986). This questionnaire, containing 15 questions, can be used as a screening tool for possible signs of depression. Participants needed a score of 5 out of 15 or lower to be considered psychologically healthy.





## 2.3 DESIGN AND PROCEDURE

The basis for both subsets was the same: two computerised tasks – (1) two narrative writing tests – and (2) a questionnaire on paper. The elderly in the first subset were provided with (3) an additional MMSE, (4) two digit span tasks and (5) the GDS. The elderly in the second subset were provided with (3) the MoCA instead of the MMSE, (4) three digit span tasks and (5) the GDS. Keystroke logging tools Inputlog 7.0.0.0 (Leijten & Van Waes, 2013) and ScriptLog (Frid et al., 2014) logged and time-stamped every keyboard activity and mouse movement, providing researchers insight into the writing and pausing behaviour of the participants.

### *2.3.1* Narrative writing tasks

Participants wrote two narrative writing tasks; they all received their easier NWT1 first, followed by the complex NWT2. They were instructed to write (type) a coherent story, using the images at the top of their screen. We requested the story to be written in the first person singular and showed an exemplifying text to indicate how we expected the images to be incorporated and to indicate the desired text length. However, an exact word count or time constraint was not given. The participants were also told that the images needed to be named explicitly in the story, either in singular or plural form, to make sure they did not just describe the images. If the participant had no further questions, the instructor started the NWT. When finished, the participant received the instructions for the second, complex NWT. This time the story needed to be written in the third person singular and past tense. Even though the instructions concerning text length and naming of the images remained unchanged, they were repeated before the start of the second NWT.

### *2.3.2* Questionnaire, cognitive screening test, working memory tests and GDS

In the following phase of the experiment, we conducted the questionnaire on paper. Participants were instructed to fill it in, with the researcher helping out when needed. The questionnaire was the last task for the students in the baseline set; the healthy elderly participants received additional testing. The cognitive screening tests (the MoCA and the MMSE), the working memory tests (the digit span tasks), and the GDS were conducted in accordance with the official guidelines.

**Table 2**

*Selection Of Product Variables That Give Insight Into The Cohesion Within The Narratives*

| Variable | Explanation |
|---|---|
| **Lexicon** | |
| Cohesion | A score that indicates how cohesive a text is perceived to be. |
| Total words | The total number of words per text, including corrections if needed. |
| Word length | The average number of letters per word |
| Word frequency of nouns | The log frequency of the nouns in the text |
| Type token ratio of words | The number of different words (types) divided by the total number of words (tokens) |
| Density of adjectives | Proportion of adjectives on the total number of words (i.o.l.r.) |
| Density of nouns | Proportion of nouns on the total number of words (i.o.l.r.) |
| Density of verbs | Proportion of verbs on the total number of words (i.o.l.r.) |
| Density of concrete verbs | Proportion of concrete verbs on the total verbs (i.o.l.r.) (example verbs: to smell, to freeze) |
| Density of abstract verbs | Proportion of abstract verbs on the total verbs (i.o.l.r.) (example verbs: to contribute, to hope) |
| **Grammar** | |
| D-Level | The 'Development Level' or a measure of syntactical complexity. In a text, every sentence receives a complexity score, which results in a text average. |
| Distance subject-verb | The distance between two elements of the sentence that belong together (subject-verb). The further those two are apart, the more difficult a sentence is to interpret. The score in T-Scan reflects the average number of words between the subject and verb in the text. |
| Distance determiner-noun | The distance between two elements of the sentence that belong together (determiner-noun). The further those two are apart, the more difficult a sentence is to interpret. The score in T-Scan reflects the average number of words between the determiner and noun in the text. |
| Density of personal references | Density of references to people (personal and possessive pronouns, nouns referring to a person and names) |
| Density of conjunctions | Proportion of conjunctions on the total number of words (i.o.l.r.) |

*Note*: i.o.l.r. = indicator of lexical richness

## 2.4 DATA PREPARATION

The obtained texts from the narrative writing tasks were proofread by the first author and spelling or typing errors in the text were corrected. This was necessary so as disambiguate the given words and texts for the automated processing with the computer program T-Scan: a tool for analysing



Dutch texts (Pander Maat et al., 2014). T-Scan provides information on word, sentence, and text level. For the purpose of this study, we decided not to focus on sentence level given that some participants did not use punctuation, which reduced the entire text to one long sentence. We did, however, focus on cohesion variables such as the frequency of word occurrence, the ratio of word type occurrence, the Type Token Ratio, etc. We divided the variables into two categories: lexicon and grammar, cf. Table 2.



## 2.5 ANALYSES

Due to the complex structure of the data, we relied on multilevel modelling in RStudio to analyse the variables. Our data is hierarchically structured, with texts nested in participants. Hence, the influence of sex, task and/or age could be incorporated when studying the effects of cohesion changes upon ageing, providing a richer basis for interpretation and a smaller opportunity of Type-I errors to occur (Hox, 2002; Quené & van den Bergh, 2008). We developed six different models (see Table 3) to measure the (interaction) effect of age, sex, and task on the cohesion characteristics of the text. As can be seen in Table 3, Model 0 does not contain any explanatory variables, only random variance. Models 1, 2, and 3 take into account the main effects. In models 4 and 5, interaction effects are added to the model with main effects. Note that the healthy ageing adults were not treated as a homogeneous group in the analysis, but rather as a set of individuals through a multi-level analysis in RStudio. For the presentation of the normative dataset, however, the findings were divided per decade.

**Table 3**

*Overview of the multilevel models used in the statistical analyses*

| Model | General overview of the data |
|---|---|
| 0 | Random variance |
| 1 | Effect of age |
| 2 | Effect of age and sex |
| 3 | Effect of age, sex and task |
| 4 | Effect of task and interaction effect (age x sex) |
| 5 | Interaction effect (age x sex x task) |

# 3. RESULTS

Results were analysed in light of the changes in narrative cohesion that occur upon ageing using lexical and grammatical variables and in order to create a normalised table that can be used in clinical settings and future studies. Table 4 and 5 provide an overview of respectively the lexical and grammatical results, showing the best fitting model for each variable under study. In addition, parameter estimates for the best-fitting model were determined in R. An overview of the significant effects and theire estimates of these best-fitting models can also be found in Table 4 and 5. A full overview of the results with model fits and comparisons of the models can be found in **Appendix B**, along with estimates of fixed effects of the best fitting model. The models were compared based on the change in log-likelihood ratio. To determine the model with the best fit, Chi-square test statistic was used. The normative datatables can be found in **Appendix C**.

**Table 4**

*An Overview Of The Models That Described Most Lexical Variance In The Data, A Summary Of The Significant Results And Their Estimates.*

| | M | Significant findings | Est. |
|---|---|---|---|
| Total words | 5 | Age: decrease of total words with age | -1.639 |
| Word length | 3 | Sex: shorter words for women | -.011 |
| | | Task: longer words for NWT2 | .000 |
| Word frequency of nouns | 3 | Age: decrease in frequency with age | -.003 |
| | | Task: decrease in NWT2 | -.265 |
| Type token ratio of words | 1 | Age: increase in TTR | .002 |
| Density of adjectives | 0 | / | |
| Density of nouns | 3 | Sex: decrease for women | -11.107 |
| | | Task: increase in NWT2 | 8.147 |
| Density of verbs | 2 | Sex: higher density for women | 1.265 |
| Density of concrete verbs | 0 | / | |
| Density of abstract verbs | 2 | Sex: higher density for women | 6.632 |

*Note*: M = model



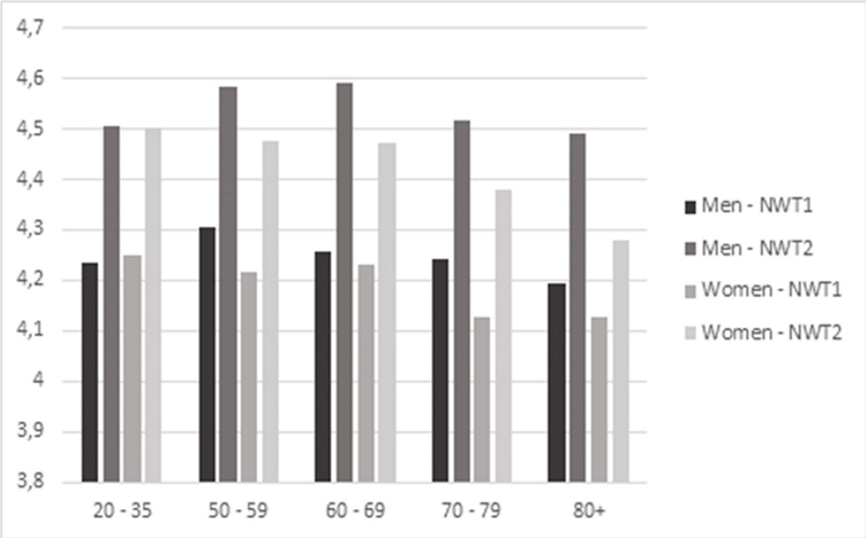
## 3.1 LEXICON

*Total number of words:* the fifth model explained most of the variance in the data: $\chi2(10) = 8.509$, $p = .037$, implicating that there is an interaction effect between age, task and sex. The effect of age is significant, with a decrease of the total words used when the participant was older.

*Word length*: The first model was not significant; models two and three were. Hence there is no main effect of age; there is a main effect of sex and task. The best general model is the third model: $\chi2(6) = 134.902$, $p < .001$. There was a significant increase in letters in the second narrative writing task (see Figure 2) and women tend to write shorter words. There was a non-significant increase with age.

**Figure 2**

*The Average Word length, Shown Per Sex, Task And Age Group*




*Word frequency of nouns*: The first, second and third model were signicant; there was a main effect of age, sex and task. The best model was the third model (effect of age, sex and task): $\chi2(6) = 131.275$, $p < 0.001$. There is a significant decrease in frequency with age and in the NWT2; women have a higher log frequency of nouns compared to men (Est. .053), but the difference is non-significant, even though it was significant in models 1 and 2.

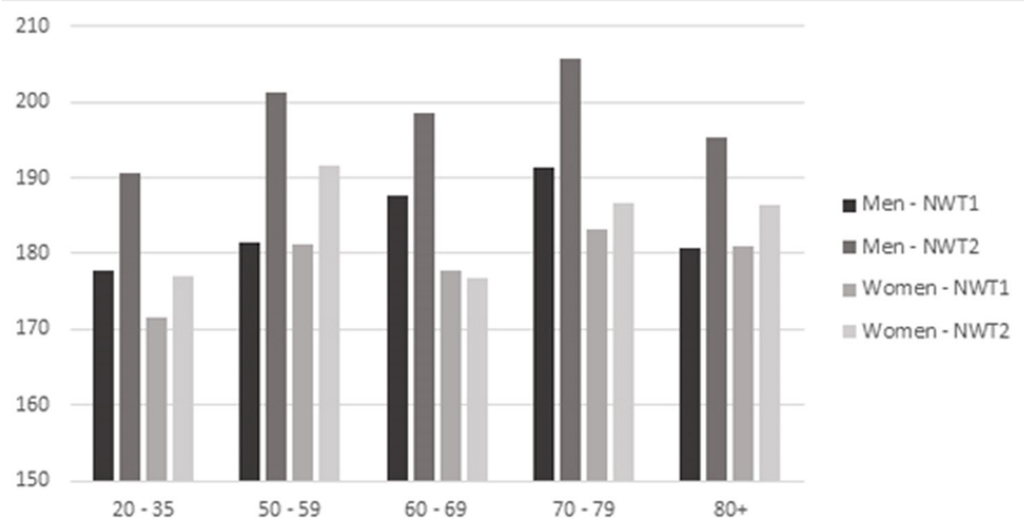
The *Type token ratio of words* can best be explained by the first model χ2(4) = 53.497, $p < 0.001$, which implies there is a main effect of age. There is an increase of TTR with age.

For some variables, the null model turned out to be the best model, leading to the conclusion that neither age, task nor sex can explain the variance in the data. This is the case for: *Density of adjectives* & *Density of concrete verbs*.

*Density of nouns*: The first model was not significant, the second and third word. The best model to describe the variance within the density of nouns was the third model: χ2(6) = 9.644, $p = .002$, implying a main effect of sex and task but not age (see Figure 3). There was a significant increase in the density of nouns in the second narrative writing task and a significant decrease for women compared to men. Ageing only led to a small, non-significant increase.

**Figure 3**

*The Average Density Of Nouns, Per Sex, Task And Age Group*




*Density of verbs* & *Density abstract verbs*: For both variables only the second model (Effect of age and sex) was significant, which signifies a main effect of sex but not age; respectively χ2(6) = 110.742, $p = 0.001$ and χ2(6) = 4.388, $p = 0.036$. For both models, results indicate that there is an increase in density for women compared to men.



## 3.2 GRAMMATICAL

*D-Level:* the first model (effect of age) explains most variance in the data, χ2(4) = 5.219, $p = .022$. A main effect of age can be found, namely, there is a significant increase in D-Level when people age.

**Table 5**

*An Overview Of The Models That Described Most Grammatical Variance In The Data, A Summary Of The Significant Results And Their Estimates.*

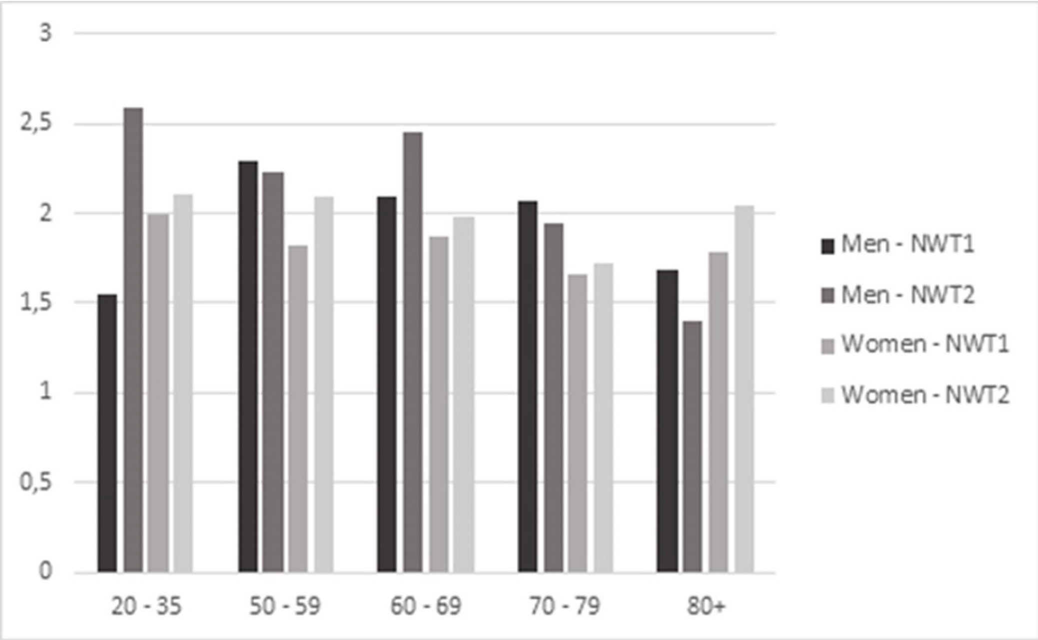
| | M | Significant findings | Est. |
|---|---|---|---|
| D-Level | 1 | Age: higher D-Level with age | .009 |
| Distance subject-verb | 5 | Task: Greater distance in NWT2 | 1.395 |
| | | Age*Task: decrease for elderly in NWT2 | -.020 |
| | | Sex*task: decrease for women in NWT2 | -1.235 |
| | | Age*Sex*Task: increase for older women, NWT2 | .020 |
| Distance determiner-noun | 0 | / | |
| Density of personal references | 3 | Age: decrease in density with age | -.270 |
| | | Task: less density in NWT2 | -6.726 |
| Density of conjunctions | 2 | Sex: The density is higher for women | 7.488 |

*Note*: M = model

*Distance subject-verbs*: The second, third and fifth models significant, whereas the first and third are not. This implies there is a main effect of sex, effect of task and an interaction effection between age, sex and task. The best model to explain most of the variance in the data is the fifth model: χ2(10) = 8.212, $p = .042$ (see Figure 4). In the second narrative writing task, there was a greater distance between subject and verb. However, when looking at the interaction between ageing and the second narrative task, a decrease in distance can be noticed. The same goes for the interaction between women and the second narrative writing task. In contrast: the interaction between age, sex and task reveals that ageing women in NWT2 write sentences with longer distances between subject and verb.

**Figure 4**

*The Average Distance Between The Subject And Verb, Per Sex, Task And Age Group*




*Distance determiner-noun*: The null model was the only significant model, revealing that neither age, task nor sex could contribute to explaining the variance in the data.

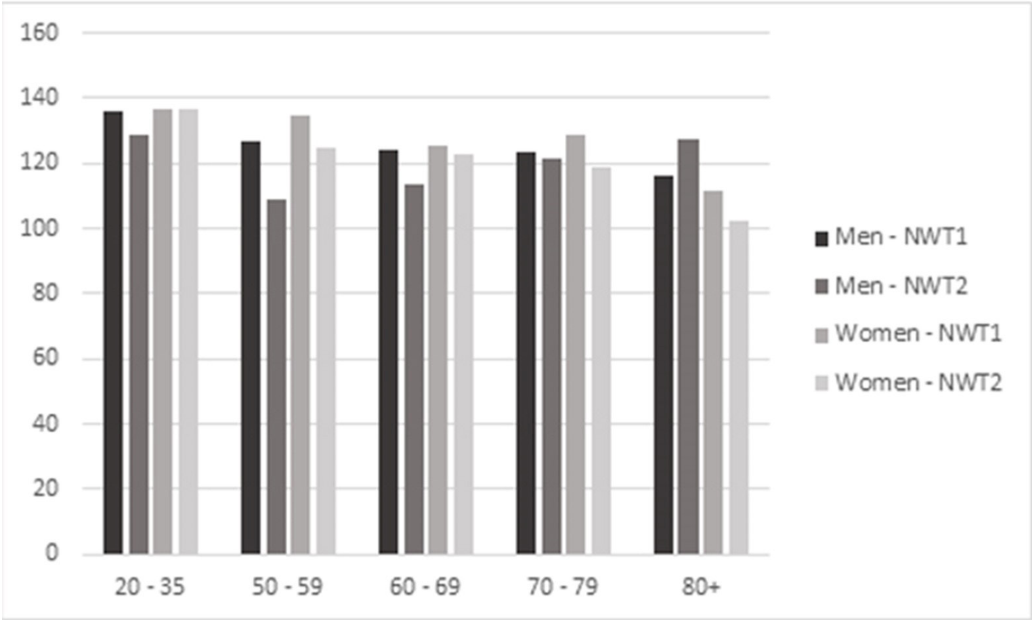
*Density of personal references*: Both the first and third model were significant, the second was not. This implies there is a main effect of age and task, but not sex. The third general model (Effect of age, sex and task) explains most of the variance in the data: $\chi2(6) = 5.704$, $p = .017$ (see Figure 5). The results indicate that with age and in the NWT2, there is a significant decrease in use of personal references.

*Density conjunctions*: Only the second model was significant, which implies that sex but not age has a main effect: $\chi2(5) = 12.231$, $p < .001$. There was a significant increase in the use of conjunctions for women compared to men.



**Figure 5**

*The Average Density Of Personal References, Per Sex, Task And Age Group*




In sum, we hypothesized that (1) grammatical complexity would decline with age and that (2) lexical complexity would change, with a preference for easier structures and formulations. With regards to the grammatical complexity, results indicate that with age, the distance between the subject and verb increases, personal references become less dense, and women have a higher density of conjunctions in their texts than men. In answer to our hypothesis that lexical complexity would change in favour of easier structures, results indicate that on the one hand age has a negative effect on the number and length of words that are used, and on the other hand age has a positive effect on complexity with older people using more difficult words. The NWT2 was more complex in nature and results indicate that this task did indeed elicit more complex words. The density of certain word types was affected by age and sex, with an increase in density of nouns in NWT2. Women - compared to men - show a decrease in noun density, and an increase in both verb and abstract verb density. Furthermore, based on the lexical and grammatical variables and with the use of our data, we were able to create a normative table (cf. **Appendix C**). A distinction has been made between sex, task, and age. The student group was added as a baseline.

## 4. DISCUSSION

With this study we aimed to describe if and how healthy ageing affects the cohesion in stories. To that purpose, we created two narrative writing tasks that also bear in mind the task characteristics needed for future research into the comparison of healthy ageing and ageing with a neurodegenerative disease. In the two narrative writing tasks, stories are elicited by prompted coloured images. The first and less complex NWT (NWT1) triggers words that are learned early on in life, that occur frequently in the language and that require participants to narrate the story from their own perspective. The more complex NWT2 uses images that trigger words learned somewhat later in life, that occur less frequent in the language and that require the story to be narrated from a third person singular perspective. In order to answer our research question, the following research questions were composed: '*Does ageing have an effect on the grammatical complexity of a typed narrative?*', '*Does ageing have an effect on the lexical complexity of a typed narrative?*', '*Does sex have an effect on the grammatical complexity of a typed narrative?*', and '*Does sex have an effect on the lexical complexity of a typed narrative?*'.

Based on our literature review, we hypothesized that ageing would negatively affect cohesion, as can be seen by (1) a diminishing grammatical complexity and (2) a preference for easier lexical structures. With the use of T-Scan, we selected both grammatical and lexical variables that contribute to the understanding of text cohesion. Results indicate that there is an increase in grammatical complexity with age, with little to no sex differences. Lexically, the most prominent phenomenon is the decrease in word count with age and increase in word difficulty. When looking at sex specific effects, women tend to use more personal references in their texts compared to men.

More specifically, with regards to the **grammatical** aspect, the results are homogeneous. *D-Level* results indicate that age positively affects sentence complexity; the *Distance between subject – verb* confirms this finding, with even an increase in distance with age and in the NWT2. The *Distance determiner - noun* shows no effect of age, task nor sex. These results corroborate findings of previous studies; Phillis Dorothy James and even Agatha Christie showed little to no decrease in their scores for D-Level during their lifetime (Le et al., 2011), and only very complex sentences or counterintuitive sentences are found to pose a problem to healthy ageing elderly (Clark-Cotton et al., 2007). In other words, these variables indicate that no negative change in sentence complexity is to be expected upon ageing.



In terms of the *Density of conjunctions*, our results indicate that neither age nor task have an effect; women use more conjunctions compared to men. Not all studies agree with the fact that the density of conjunctions remain unaffected; some do state that elderly have problems using cohesive conjunctions, such as Juncos-Rabadán, Pereiro, & Rodríguez (2005). Our finding that the *Density of personal references* is affected by the type of task might be explained by the nature of the task. The difference in instruction and images might trigger a different style of narrating, with a bigger focus on creating a coherent story in the second NWT, compared to the first NWT in which participants were instructed to write the story using the first-person perspective. Moreover, it is also affected by sex, with a higher general density of personal references for women compared to men.



**Lexically**, we expected to find a preference for easier sentence and word choices with age. Our results differ from that hypothesis and even effects of task design were found on the data. The *Total number of words* and the *Word length* indicate that, indeed, words become shorter and less numerous in the later stages of life (from age 70 onwards), even though all participants did tend to write more and longer words in the second narrative writing task. From the age of 80 onwards, the discrepancy between NWT1&2 in terms of word count increases; a possible explanation is that these participants have a greater need for semantic descriptions. The word cannot be recalled temporarily and participants will therefore rely on descriptions rather than the intended word, leading to more words and sentences in the complex task (Kavé & Goral, 2018). Future studies into the writing process behaviour could corroborate these findings with longer pauses before certain descriptions or by further evaluating the semantics of the samples.

We found that with age, people tend to use more difficult words (as seen in the decrease in *Word frequency of nouns*) and use a larger variety of word types (greater *Type token ratio of words*). This result stands in contrast to the findings of Le et al. (2011), where the novels of Phillis Dorothy James showed no change in type token ratio. Nevertheless, this comparison might not necessarily be appropriate. This result might be due to her experience as a writer, whereby her type token ratio remained on the same high level. Again, we find an effect of task, with less frequent words in NWT2 compared to NWT1; a result that might be elicited by the task design. With regards to certain word types, we found that neither age, task nor sex influenced *Denisty of adjectives* and *Density of concrete verbs*. Results indicate that the texts from women contained a lower *Density of nouns* and an increased *Density of verbs* and *Density of abstract verbs.* These findings mirror those found by (Le et al., 2011), whereby an analyses of the novels of writers Christie and Murdock showed a decrease in nouns and increase in verbs.

When we set out with this study, we expected to find that age had a negative impact on text cohesion. Our results indicate the opposite, with an increase in sentence and word complexity as people age. The only variables that were negatively affected by age concerned word length and number of words. One of the possible explanations for these findings is that our study was cross-sectional and not longitudinal in nature, unlike the nun-study. On the one hand, we had a student population with a uniform background (same age and educational level); on the other hand, the older participant group in the study contained people from all walks of life, with different educational backgrounds and experiences. Nevertheless, we believe that this diversity in the latter group is a strength for the normative data representation. The student population provide insight into the brain in an almost fully developed state (Casey et al., 2008; Sharma et al., 2013), and the older population starts at age 50 – the moment cognitive changes occur and a change in linguistic becomes noticeable (Alenius et al., 2019).

The nun-study revealed that the educational background of the nuns affected their grammar even in the early stages of their lives. Recent studies confirm that education aids in preventing/coping with the onset stages of the disease, with a higher risk for people with lower education (Roe et al., 2007; Wada et al., 2018). Therefore, it is crucial that further studies with a larger participant set should take these differences in schooling into consideration upon creating a normative data table. Additionally, given that young onset dementia may already occur at 30 years old (Rossor et al., 2010; Sampson et al., 2004), we do believe that a future study including 30 to 49 year olds would shed even more light on the linguistic changes.

Another explanation for these counterintuitive ageing effects can be found in the compensatory mechanisms in the brain that aid the more wide-spread neural circuitry. The change in structural integrity is accompanied by a change in cognitive functioning; researchers found an increase in working memory load, resulting in slower reaction times. Nevertheless, other studies have reported that the accuracy remains unaffected (Alatorre-Cruz et al., 2018) and that for instance semantic fluency is only affected from age 85 onwards (Alenius et al., 2019). It could be that, performance in our study appears to be maintained, even though the time leading to these responses might have been affected. We suggest that future studies into writing process research might shed more light on the effect ageing has on the cognitive processes during the writing instead of solely on the writing product. Furthermore, we believe that the participants with



well-developed cognitive processes are triggered to use words that – even though they are less frequent – are still well entrenched; no jargon nor unfamiliar terms were requested. Therefore, text cohesion in this field might not have been impacted.

With regards to sex, we expected differences to be found between men and women. Results indicate that indeed, both lexically and grammatically the two groups display different behaviour. For women we noticed a decrease in *Word length* and *Density of nouns* compared to men, and an increase in *Density of verbs* and *Density of conjunctions*; results also pointed towards an increase in *Distance subject-verb* as women get older. Other studies, too, have found considerable difference in the processing and production of language between men and women. In a study by Aerts and colleagues (2015) on phoneme discrimination and word recognition, clear sex differences were found, with a greater sensitivity to phonemic contrasts in women and a greater advantage for women in word-pseudoword processing. Others reported an advantage for women in verbal memory tasks, possibly attributable to the fact that they have a greater cognitive reserve compared to men which delays verbal memory decline in AD (Laws et al., 2018). Considering the neuroanatomical differences between men and women, and the non-simultaneous degeneration of certain brain structures, these functional discrepancies are to be expected. Women have been found to have greater cortical thickness, more specifically in frontal and parietal regions of the brain, and the hippocampus has been reported to be larger in women, which gives them an advantage in memory. Furthermore, regions connected to language are reported to be larger in women compared to men, whereas men generally have a greater hypothalamus, thalamus and amygdala (Aerts et al., 2015; Beltz et al., 2020). Note, however, that the spontaneous neuroplasticity might compensate for a part of these findings. Understanding the neuroanatomical sex differences and their effect on language might also provide insight into the understanding of neurological diseases and their progression. Therefore, in accordance with previous studies (Aerts et al., 2015), we recommend future normative studies to distinguish between the two sexes, as sex is clearly an influencing factor.

The normative data table (cf. **Appendix C**) has a broad range of applications within the clinical setting. We designed the current study with the linguistic discrepancy between healthy ageing and ageing due to AD as our focal point. Further research should shed more light on which cohesion variables will turn out to be indicative of a language pathology. However, we do want to stress that by no means this task replaces the current tools for early detection of cognitive decline or dementia. Rather, we believe that it can be an aid in diagnosing language issues as a primary symptom of cognitive deterioration due to its accessibility and low cost. We also believe that our test can contribute to other fields. A longitudinal retesting of our design with people who suffer from primary progressive aphasia (PPA) might provide diagnostic insight in the differential diagnosis of PPA. Nevertheless, we realise that the test is not accessible for all degenerative disease; for some patients the accompanying motoric issues will prohibit them from participating in the test. Furthermore, we also recommend future studies that consider using the narrative writing tasks to look into possible task reduction. The current experiment in its entirety required full concentration of the participants for almost 90 to 120 minutes. Of course, this time also incorporated the additional screening and memory testing and questionnaire; the narrative writing tasks alone often lasted up to 40 minutes. Reducing the task size by opting for either the NWT1 or NWT2 instead of both, will aid the participants focus and determination to finish the experiment in its entirety. Finally, we want to note that we do realise that the representativeness of our participants might be biased due to the exclusion of people with poor computer literacy.



ACKNOWLEDGEMENT

We would like to thank Eric Van Horenbeeck and Luuk Van Waes for their technical assistance with Inputlog and the typing task. We thank Johan Frid and Victoria Johansson, respectively the technical coordinator and co-founder of ScriptLog, for creating a ScriptLog module especially for our research purposes. Further thanks go to Sven De Maeyer for his input during the analysis of the data. Additionally, Peter Mariën (†) and Dorien Vandenborre are thanked for their input during the design of the study. Last but not least we would like to thank the elderly who enthusiastically participated in this study.



DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

FUNDING STATEMENT

This research was supported by the research fund BOF-DOCPRO 2015, PS ID: 31661 of the University of Antwerp. Recipient of the funding is Prof. Dr. M. Leijten.

## 5. BIBLIOGRAPHY


Aerts, A., van Mierlo, P., Hartsuiker, R. J., Santens, P., & De Letter, M. (2015). Sex Differences in Neurophysiological Activation Patterns During Phonological Input Processing: An Influencing Factor for Normative Data. *Archives of Sexual Behavior,* 44(8), 2207–2218. https://doi.org/10.1007/s10508-015-0560-y

Alatorre-Cruz, G. C., Silva-Pereyra, J., Fernández, T., Rodríguez-Camacho, M. A., Castro-Chavira, S. A., & Sanchez-Lopez, J. (2018). Effects of age and working memory load on syntactic processing: An event-related potential study. *Frontiers in Human Neuroscience*, 12(May), 1–11. https://doi.org/10.3389/fnhum.2018.00185

Alenius, M., Koskinen, S., Hallikainen, I., Ngandu, T., Lipsanen, J., Sainio, P., Tuulio-Henriksson, A., & Hänninen, T. (2019). Cognitive Performance among Cognitively Healthy Adults Aged 30-100 Years. *Dementia and Geriatric Cognitive Disorders Extra*, 9(1), 11–23. https://doi.org/10.1159/000495657

Alzheimer's, A. (2019). 2019 Alzheimer's Disease Facts and Figures. *Alzheimer's Dement*, 15(3), 321–387.

Alzheimer's Disease International, X. (2019). Attitudes to dementia. In World Alzheimer Report 2019: *Attitudes to dementia*. Alzheimer's Disease International.

Bates, E., Harris, C., Marchman, V., Wulfeck, B., Bates, E., Harris, C., Marchman, V., & Wulfeck, B. (1995). Production of Complex Syntax in Normal Ageing and Alzheimer ' s Disease. *Language and Cognitive Processes*, 10(05), 487–539. https://doi.org/10.1080/01690969508407113

Beltz, A. M., Kelly, D. P., & Berenbaum, S. A. (2020). Chapter 27 - Sex differences in brain and behavioral development. In J. Rubenstein, P. Rakic, B. Chen, & K. Y. B. T.-N. C. and C. D. (Second E. Kwan (Eds.), *Neural Circuit and Cognitive Development* (pp. 585–638). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-814411-4.00027-5

Boyer, J. T. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the nun study. *Journal of the American Geriatrics Society*, 44(10), 1279–1280. https://doi.org/10.1111/j.1532-5415.1996.tb01395.x

Burianová, H., Lee, Y., Grady, C. L., & Moscovitch, M. (2013). Age-related dedifferentiation and compensatory changes in the functional network underlying face processing. *Neurobiology of Aging*, 34(12), 2759–2767. https://doi.org/10.1016/j.neurobiolaging.2013.06.016

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philos Trans R Soc Lond B Biol Sci,* 352(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

Carp, J., Park, J., Polk, T. A., & Park, D. C. (2011). Age differences in neural distinctiveness revealed by multi-voxel pattern analysis. *Neuroimage*, 56(2), 736–743. https://doi.org/10.1016/j.neuroimage.2010.04.267.Age

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The Adolescent Brain. *Annals of the New York Academy of Sciences*, 1124(1), 111–126. https://doi.org/10.1196/annals.1440.010

Clark-Cotton, M. R., Williams, R. K., Goral, M., & Obler, Loraine, K. (2007). *Language and Communication in Aging* (pp. 1–8). https://doi.org/10.1016/B0-12-370870-2/00103-7

Cockrell, J. R., & Folstein, M. F. (2002). Mini-mental State Examination. In J. R. M. Copeland, M. T. Abou-Saleh, & D. G. Blazer (Eds.), *Principles and Practice of Geriatric Psychiatry* (pp. 140–141).

Fleck, J. I., Kuti, J., Mercurio, J., Mullen, S., Austin, K., & Pereira, O. (2017). The Impact of Age and Cognitive Reserve on Resting-State Brain Connectivity. *Frontiers in Aging Neuroscience,* 9(December), 1–13. https://doi.org/10.3389/fnagi.2017.00392

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Frid, J., Johansson, V., Johansson, R., Wengelin, Å., & Johansson, M. (2014). Developing a keystroke logging program into a writing experiment environment. *Presentation at Writing Across Borders,* WRAB, Paris February 2014.

Garrard, P., Maloney, L. M., Hodges, J. R., & Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain*, 128(2), 250–260. https://doi.org/10.1093/brain/awh341

Hox, J. J. (2002). Multilevel analysis: Techniques and applications: Second edition. *In Multilevel Analysis: Techniques and Applications.* Erlbaum. https://doi.org/10.4324/9780203852279

Juncos-Rabadán, O., Pereiro, A. X., & Rodríguez, M. S. (2005). Narrative speech in aging: Quantity, information content, and cohesion. *Brain and Language*, 95(3), 423–434. https://doi.org/10.1016/j.bandl.2005.04.001

Kavé, G., & Goral, M. (2018). Do age-related word retrieal difficulties appear (or disappear) in connected speech? *Neurospychol Dev Cogn B Aging Neuropsychol Cogn*, 24(5), 508–527. https://doi.org/10.1080/13825585.2016.1226249

Kemper, S., Greiner, L., Marquis, J., Prenovost, K., & Mitzner, T. (2001). Language decline across the life span: Findings from the nun study. *Psychology and Aging*, 16(2), 227–239. https://doi.org/10.1037/0882-7974.16.2.227

Koen, J. D., & Rugg, M. D. (2019). Neural Dedifferentiation in the Aging Brain. *In Trends in Cognitive Sciences.* https://doi.org/10.1016/j.tics.2019.04.012

Laws, K. R., Irvine, K., & Gale, T. M. (2018). Sex differences in Alzheimer ' s disease. C*urrent Opinion in Psychiatry,* 31(2), 133–139. https://doi.org/10.1097/YCO.0000000000000401

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing*, 26(4), 435–461. https://doi.org/10.1093/llc/fqr013

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication*, 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Meunier, D., Stamatakis, E. A., & Tyler, L. K. (2014). *Age-related functional reorganization , structural changes , and preserved cognition.* Neurobiology of Aging, 35, 42–54.

Min, Y. (2015). *ESL Student Handbook*. https://www.uwb.edu/wacc/for-students/eslhandbook/coherence

Mogle, J., & Sliwinski, M. (2013). Perspectives on Cognitive Aging. In J. M. Wilmoth & K. F. Ferraro (Eds.), *Gerontology: perspectives and Issues* (pp. 49–68). Springer Publishing Co.

Mu, S. H., Xu, M., Duan, J. X., Zhang, J., & Tan, L. H. (2017). Localizing Age-Related Changes in Brain Structure Using Voxel-Based Morphometry. *Neural Plasticity*, 2017. https://doi.org/10.1155/2017/6303512

Nasrabady, S. E., Rizvi, B., Goldman, J. E., & Brickman, A. M. (2018). White matter changes in Alzheimer ' s disease: a focus on myelin and oligodendrocytes. *Neuropathological Communications,* 6(22), 1–10.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment , MoCA : A Brief Screening. *Journal of the American Geriatrics Society,* 53(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Paesen, L., & Leijten, M. (2019). Name agreement and naming latencies for typed picture naming in aging adults. *Clinical Linguistics & Phonetics,* 33(10–11), 930–948. https://doi.org/10.1080/02699206.2019.1590734

Pander Maat, H., Kraf, R., Bosch, A. van den, Dekker, N., Gompel, M. van, Kleijn, S., Sanders, T. J. M., & Sloot, K. van der. (2014). T-Scan : a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.

Park, D. C., & Bischof, G. N. (2013). The aging mind: neuroplasticity in response to cognitive training. *Dialogues in Clinical Neuroscience,* 5(1), 109–119.

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain : Aging and Neurocognitive Scaffolding. *Annual Review of Psychology,* 60, 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656.The

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59(4), 413–425. https://doi.org/10.1016/j.jml.2008.02.002

Roe, C. M., Xiong, C., Miller, J. P., & Morris, J. C. (2007). Education and Alzheimer disease without dementia. *Neurology*, 68(3), 223 LP – 228. https://doi.org/10.1212/01.wnl.0000251303.50459.8a

Rosselli, M., Ardila, A., Matute, E., & Vélez-uribe, I. (2014). Language Development across the Life Span : A Neuropsychological / Neuroimaging Perspective. *Neuroscience Journal*, 2014. https://doi.org/10.1155/2014/585237

Rossor, M. N., Fox, N. C., Mummery, C. J., Schott, J. M., & Warren, J. D. (2010). The diagnosis of young-onset dementia. T*he Lancet Neurology*, 9(8), 793–806. https://doi.org/10.1016/S1474-4422(10)70159

Sampson, E. L., Warren, J. D., & Rossor, M. N. (2004). Young onset dementia. *Postgraduate Medical Journal*, 80(941), 125–139. https://doi.org/10.1136/pgmj.2003.011171

Schmid, H. J. (2007). Entrenchment, salience, and basic levels. *The Oxford Handbook of Cognitive Linguistics*, 117138. https://doi.org/10.2307/j.ctvgc6251

Schmid, H. J. (2017). *Entrenchment and the Psychology of Language Learning* (Issue May 2014). De Gruyter Mouton. https://doi.org/10.1515/9783110341423

Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the Tip-of-the-Tongue: Neural Correlates of Increased Word-finding Failures in Normal Aging. *Journal of Cognitive Neuroscience,* 19(12), 2060–2070. https://doi.org/10.1162/jocn.2007.19.12.2060

Sharma, S., Arain, Mathur, Rais, Nel, Sandhu, Haque, & Johal. (2013). Maturation of the adolescent brain. *Neuropsychiatric Disease and Treatmen*t, 449. https://doi.org/10.2147/NDT.S39776

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Healt*h, 5(1–2), 165–173. https://doi.org/10.1300/J018v05n01_09

Silagi, M., Bertolucci, P., & Ortiz, K. (2015). Naming ability in patients with mild to moderate Alzheimer's disease: what changes occur with the evolution of the disease? *Clinics*, 70(6), 423–428. https://doi.org/10.6061/clinics/2015(06)07

Sirts, K., Piguet, O., Johnson, M., & Centre, M. (2017). Idea density for predicting Alzheimer 's disease from transcribed speech. *Id*. https://doi.org/http://dx.doi.org/10.18653/v1/K17-1033

Steinkrauss, R., & Schmid, M. S. (2016). Entrenchment and language attrition. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 367–383). Mouton de Gruyter.

Wada, M., Noda, Y., Shinagawa, S., Chung, J. K., Sawada, K., Ogyu, K., Tarumi, R., Tsugawa, S., Miyazaki, T., Yamagata, B., Graff-Guerrero, A., Mimura, M., & Nakajima, S. (2018). Effect of Education on Alzheimer's Disease-Related Neuroimaging Biomarkers in Healthy Controls, and Participants with Mild Cognitive Impairment and Alzheimer's Disease: A Cross-Sectional Study. *Journal of Alzheimer's Disease* : JAD, 63(2), 861–869. https://doi.org/10.3233/JAD-171168

Wechsler, D. (2008). *Wechsler adult intelligence scale: WAIS IV; technical and interpretive manual* (Pearson (ed.)).

Williamson, J. M., & Lyons, D. A. (2018). Myelin Dynamics Throughout Life : An Ever-Changing Landscape ? *Frontiers in Cellular Neuroscience*, 12, 424. https://doi.org/10.3389/fncel.2018.00424

Will be submitted to: *European Journal of Applied Linguistics*

# Chapter four

## Clinical toolkit for the evaluation of written spontaneous speech in healthy adults

## ABSTRACT

**Purpose** – Previous studies have found that mapping initial language changes that occur in healthy ageing could provide valuable insight into the onset of several progressive neurological disorders. With this study, we set out to establish the test-retest reliability of spontaneous written language in order to lay the foundation of a sensitive, non-intrusive toolkit that can measure language changes in healthy ageing and potentially ageing with a certain progressive disorder.

**Method** – A total of 58 healthy volunteers, aged 50 and over, participated in our study thrice, with an interpose of three months. On each occasion, they completed two narrative writing tasks – tasks in which they were asked to write (type) a narrative based on given prompts. Additional cognitive screening tests, working memory tests, a typing test and a questionnaire also needed to be made. The narrative writing tasks were analysed with a focus on both the writing product and writing process.

**Results** – The interpretation of the results lead to a two-fold. The writing product was not impacted by the repeated testing, leading to a great test-rest reliability. Moreover, no significant effects of age were found. The writing process was influenced by repeated testing, especially for the first narrative writing task.

**Conclusion** – Given the test-retest reliability of spontaneous written language elicited by narrative writing tasks, these tasks form the basis of a toolkit that can be used to establish language changes in healthy ageing and could potentially even be used in differential diagnostics, provided more research is done on the topic. Further research into the generation of language (or the writing process) is needed to create a more complete picture of potential issues encountered during the production of a narrative.


## 1. INTRODUCTION

Acquired neurogenic language disorders – such as aphasia – are the result of damage to the central or peripheral nervous system. The effects may appear suddenly, mostly after an acute event such as stroke or trauma, or gradually due to a progressive neurological disorder (E.g., Alzheimer's disease) or brain tumour (Hill, 2021). Due to the fact that age is a predominant risk factor for these disorders (Franceschi et al., 2018), further research into the language changes in healthy ageing is needed to provide additional insight into the onset of the aforementioned (progressive) disorders. To that extent, researchers have often relied on language production which, in contrast to language comprehension, is prone to show signs of age-related impairments (Burke & Mackay, 1997; Shafto et al., 2007). With regard to the current study, we have decided to base ourselves on the principles found in the current test batteries that are used to diagnose neurogenic language disorders, in order to develop a new set of tasks that could be used as a sensitive, non-intrusive toolkit to measure language changes in healthy ageing.



For neurogenic language disorders to be diagnosed, researchers often rely on either a disorder focused or participation focused test. The former requires the use of more traditional tasks that focus on the deficit causing the communication disfunction, through studying general language fluency with for instance pragmatics, phonology and semantics. One such traditional task is a picture naming task, as this type of task is ideal to map semantics and word retrieval. Given that they require speeded naming of given prompts, language (either spoken or written) is elicited 'in a standardised fashion that facilitates interindividual comparison' (Hier et al., 1985). Consequently, this type of task can be found in various Dutch disorder focused diagnostic language tests, such as the CAT-NL (Visch-Brink et al., 2014) and PALPA (Bastiaanse et al., 1995).

An example of a participation focused test in Dutch is the ANELT (Amsterdam-Nijmegen Everyday Language Test). Those type of tests focus on a general communication analysis and aim at mapping the possibilities and restrictions in terms of activities and participation (Manders, 2016; Visch-brink & Wielaert, 2011). Studying general communication can be done through for instance an analysis of spontaneous discourse, as most researchers within this field argue that discourse analysis will be most sensitive to differentiate between healthy ageing and ageing with certain pathologies, such as Alzheimer disease (Choi, 2009). Evidence of the influence of AD on spoken discourse can for instance be found in the unscripted speeches of former President Ronald Reagan, whose language showed signs of impairment - with more pauses, a decrease in unique words and ample filling words (e.g., 'uhm') - long before Alzheimer's disease was diagnosed (Berisha et al., 2015). Mueller and colleagues (2018) analysed spoken discourse elicited by a picture description task of 264 participants (belonging to either a cognitively healthy cohort or cohort suffering from MCI). They found that the semantic and fluency features of discourse changed more rapidly owing to a cognitive impairment, and the proportion of pronouns and verbs increased, whereas the proportion of nouns decreased.

Nevertheless, some researchers even state that within discourse, written language tends to be even more susceptible to the effects of the progression of a neurological disorder than spoken language. At the onset of for instance AD, the patient may suffer from an impairment in the writing of irregular words (lexical dysgraphia). This will gradually become worse, until the writing of regular words is influenced (phonological agraphia) (Croisile et al., 1996; Sitek et al., 2015), whereas the disease has almost no impact on the phonological system in spoken discourse (Forbes-McKay et al., 2014; Forbes et al., 2004). Hence, studying written language output will provide a more sensitive means to measure early cognitive changes (Croisile et al., 1996; Forbes-McKay et al., 2014; Sitek et al., 2015), especially when it entails spontaneous discourse.

Evidence for the sensibility of written discourse can be found in the writings of Iris Murdoch. In order to find out how written language deteriorates, both Garrard and colleagues (2005) and Le and colleagues (2011) studied the novels of Iris Murdoch, an accomplished English author who wrote stories for most of her life and was diagnosed with Alzheimer's disease. They already found signs of linguistic changes (e.g., idea density – the number of ideas within a text) in the novels that she wrote before she and her surrounding were aware of her neurological illness. Her latest work shows a clear impairment of the lexico-semantic system, which was visible through a decrease in the number of unique words compared to her earlier novels (Garrard et al., 2005). Similarly to the aforementioned speeches of Ronald Reagan (Berisha et al., 2015), there was a decrease in the number of unique words which resulted in an increase of repeated phrases and filler words. In line with the findings of Bates and colleagues (1995), Murdoch also showed a decrease in the use of the passive voice upon ageing (Le et al., 2011). Despite the fact that the novels were reviewed by her editor, these effects were still apparent and can be corroborated by findings from for instance the nun study (Kemper et al., 2001). In this study, language samples of 180 nuns were studied, focussing on grammatical complexity and idea density. Results indicated that both measures decreased with age, language became vague and repetitive, and in line with the findings on Ronald Reagan & Iris Murdoch, and nuns who suffered from Alzheimer's disease



were revealed to show signs of impaired language – with lower idea density and grammatical complexity - even before the disease manifested itself.

Modern technology – such as certain types of computer software – has allowed researchers to study spontaneous written discourse even more thoroughly. In studies by Paesen and colleagues. (n.d.) and Paesen & Leijten (2019) the added value of studying the writing process behind a narrative text is stressed. In writing processes as well as in oral speech production, the number, length and location of pauses have proven to be important resources for identifying cognitive complexity (Olive, 2012), since they point out production flow interruptions (or 'writing fluency'; Alves et al., 2008; Hayes & Ann Chenoweth, 2007; Van Waes & Leijten, 2015). Processes that are frequently interrupted due to pauses or corrections are considered to be less fluent. In a study by Holmqvist and colleagues (2002) on personal narratives, the writing process at the beginning of the task was less fluent with pauses that proved to be lengthy and clustered, while the middle phase writing process was more fluent. These findings were attributed to the narrative story that needs to be constructed, leading to a difference in fluency in the different stages of the writing process (K. Holmqvist et al., 2002).



Moreover, specific types of pauses within the process (e.g., interkey latencies, pauses between the typing of certain keys), have been proven to provide insight into potential issues encountered in the writing process, especially upon ageing. These measures were found to be stable in test-retest situations (Paesen & Leijten, 2019); a finding corroborated by Van Waes and colleagues (2021). The fact that fluency in discourse could be used to differentiate between cognitively impaired and cognitively healthy participants was further accentuated in a picture description study by Mueller and colleagues (2018), who found a decline in discourse fluency for patients with a mild cognitive impairment and even found the change in fluency to be a predictor for cognitive health. Moreover, several studies have found issues in fluency for patients with Alzheimer's disease by increased pauses, hesitations and corrections in their discourse (Davis & Maclagan, 2009; Gayraud et al., 2011; Hoffmann et al., 2010; McNamara et al., 1992; Nicholas et al., 1985; Sajjadi et al., 2012; Tomoeda & Bayles, 1993).

The main purpose of the current study is to determine the test-retest reliability of spontaneous written language in healthy ageing, adopting both a writing product and process approach. To our knowledge there are no other studies that focus on the writing process behind narrative writing in healthy ageing to this date. The key aspects of the product and process analysis will be the (a) individual characteristics of writing, (b) general fluency measures both in the beginning of the writing process and globally throughout (e.g., by measuring active writing time, pause distribution, utterance length and number (bursts)), and (c) cohesion measures: linguistic and grammatical variables (such as idea density & density of verbs) taken from Paesen et al., (n.d.). To that extent, narrative writing tasks created by Paesen and colleagues (n.d.) will be used to allow for both a product and process approach. These tasks were created due to a need for tasks that can be used in repeated testing, and aim to allow for a participant focussed diagnosis in the future. The design of the tasks was made based on research done on The frog story and picture description tasks. The former task triggers narrative discourse with the use of 24 images, and was designed to study language acquisition in children and more specifically the syntax and discourse (Kenneth Holmqvist & Johansson, 2005; Mayer, 1969). An example of the latter task is the Boston cookie theft picture (Goodglass & Kaplan, 1983), in which a narrative is elicited based on a single image (i.e.: used in research on MCI and AD by Forbes-McKay and colleagues (2005)). Even though both tasks elicit a more or less spontaneous discourse, issues can be found when used within repeated testing. Using the same images in case of the Frog story might trigger a learning effect and using various variants of picture description tasks may trigger incomparable word use. Longitudinal research requires comparable variants of the same task: tasks that do not trigger a learning effect even though comparable results must be elicited.

In addition to measuring the test-retest reliability of written discourse with the use of narrative writing tasks, we want to standardise these tasks so as to explore their potential to map spontaneous written language in healthy ageing and ageing with a cognitive impairment. Therefore, we will try and answer the following research question: To what extent does healthy ageing affect the test-retest reliability of spontaneous written language? In order to find an answer to this question, the data of 57 healthy ageing adults was studied; these adults were tested three times over the course of six months. In accordance with Paesen and colleagues (n.d.) and Aerts and colleagues (2015), we will differentiate between sexes in the formulation of our answer. We posit the hypothesis that more pauses will be needed by the older participants due to word finding issues, even though the writing product will not be affected (e.g., in terms of total word count). Additionally, we hypothesize that no change in both product and process data should be visible over the course of the three trials.



# 2. METHOD

## 2.1 PARTICIPANTS

A total of 58 healthy volunteers aged between 50 and 79 years of age participated in this study (see Table 1 for a detailed overview). They were tested three times with an interval of three months between each test moment. With the use of the questionnaire, the MoCA and GDS, we assured that all participants met the same strict criteria. All participants were native Dutch speakers; they did not have any neurological and/or psychological illnesses, previous head trauma, developmental language issues, concentration problems and/or colour blindness. All participants had normal or corrected-to-normal vision and were experienced in working and typing on a computer. The Montreal Cognitive Assessment and GDS scores served as an exclusion criterion (see Design & procedure). Due to this strict selection, nine participants were excluded from our initial participant group of 75 participants. Another eight participants dropped out of their own accord after either test moment one or two. All participants were contacted through the network of other participants, through elderly organisations or through researcher's own network. Participants were tested individually at a location of their choice – mostly their homes. This study was approved by the Ethics Commission and required the written consent of all participants. It was performed in accordance with the Declaration of Helsinki.



## 2.2 MATERIALS

### *2.2.1 Typing test*

A typing task was administered to account for interpersonal differences in typing skills and to establish a baseline typing speed for each participant. We used the typing test developed by Inputlog (Waes et al., 2019), which measures keystrokes in seven consecutive assignments: (1) Repetition typing of two letters for a time span of 15 seconds, (2) a sentence repetition task for a time span of 30 seconds, (3 to 6) copying a combination of two/three words seven times, (7) copying four blocks of six consonants. These different tasks are controlled for specific bigram combinations, thereby allowing researchers to get insight and compare the typing behaviour of the participants. In order to measure the raw typing skills and not the working memory capacity of the participants, the target words or sentences are visible during that particular part of the typing test, allowing the cognitive load to be kept to a minimum (Leijten & Van Waes, 2013; Waes et al., 2019) A final questionnaire enquired after the handedness of the participant, using the short form of the Edinburgh Handedness Inventory (Veale, 2013).

**Table 1**

*Demographic Data Of Healthy Participants, Divided Per Decade And Sex, Presenting Their Years Of Education And MoCA Score On The Three Test Occasions.*

| | 50 – 59 | | 60 -69 | | 70 - 79 | |
|---|---|---|---|---|---|---|
| | F | M | F | M | F | M |
| N | 16 | 7 | 12 | 11 | 7 | 5 |
| Years of education | | | | | | |
| Mean | 14.50 | 16.00 | 13.92 | 15.54 | 12.43 | 14.00 |
| SD | 2.63 | 3.27 | 1.73 | 2.66 | 3.69 | 3.16 |
| Range | 8-19 | 12-22 | 11-16 | 10-21 | 8-19 | 10-18 |
| MoCA score M1 (/30) | | | | | | |
| Mean | 28.63 | 28.43 | 26.83 | 28.00 | 26.71 | 27.20 |
| SD | 1.31 | 1.40 | 2.12 | 2.00 | 1.38 | 1.30 |
| Range | 26-30 | 26-30 | 24-30 | 24-30 | 25-29 | 26-29 |
| MoCA score M2 (/30) | | | | | | |
| Mean | 28.69 | 28.71 | 28.25 | 28.27 | 28.57 | 27.60 |
| SD | 1.49 | 1.60 | 0.97 | 1.79 | 1.51 | 1.82 |
| Range | 25-30 | 26-30 | 27-30 | 25-30 | 27-30 | 25-30 |
| MoCA score M3 (/30) | | | | | | |
| Mean | 28.50 | 28.43 | 28.00 | 28.00 | 27.86 | 27.20 |
| SD | 1.51 | 1.13 | 1.41 | 1.34 | 1.68 | 1.30 |
| Range | 26-30 | 27-30 | 25-30 | 26-30 | 25-30 | 26-29 |

*Note*: M1= moment 1, M2 = moment 2, M3 = moment 3, M = male, F = female

### 2.2.2 *Narrative writing tasks*

Spontaneous written language was elicited with the use of narrative writing tasks (NWTs). Those tasks require participants to write a coherent story based on given images (see Figure 1, Figure 2). All images were retrieved from the Open Linguistic Picture Database (OLPD; www.olpd.eu) based on their name agreement, frequency, age-of-acquisition, and reaction times in Dutch (Severens et al., 2005) and were put into two categories 'easier' and 'more complex' based on these variables. For instance, a 'cat' was placed in the category 'easier' as it has a high name agreement, frequency and low age-of-acquisition; 'wheelbarrow' was considered 'more complex' based on its lower



name agreement, frequency and higher age-of-acquisition. For a full description of the selected images and the creation of the OLPD, see Paesen & Leijten, 2019 or visit www.olpd.eu. Each participant received a semi-random combination of those images; they were sampled without replacement in order to assure no learning effect could take place between the test moments. The first NWT (NWT1, see Figure 1) contains a random selection of four 'easier' images: an animal, two objects and a vehicle. The second, more complex NWT (NWT2, see Figure 2) contains six 'more complex' images: a human, two animals, two objects and a vehicle. Further differentiations are made in the instructions; NWT1 has to be written in the first person singular, NWT2 needs to be written in the third person singular and in the past tense.



**Figure 1**

*An Example Of A Possible NWT1, Consisting Of 4 Images (One Animal, One Vehicle, Two Objects)*

### 2.2.3 *Questionnaire*

A questionnaire on paper needed to be filled in, in all three test moments. The questionnaire of M1 contained questions about the background of participants: their education, job, possible duration of their retirement, language background, neurological and psychological history and information on how they perceived the tasks on the computer. The questionnaires of M2 and M3 was more concise, only inquiring after possible changes compared to the previous test moment and the perceived ease of the tasks.

### 2.2.4 *Cognitive screening test*

The Montreal Cognitive Assessment (MoCA, Nasreddine et al., 2005) was used as a cognitive screening tool during the three test moments. Healthy participants are differentiated by people with a mild cognitive impairment (MCI) based on the score obtained from this test. A score higher than 25/30 indicates that the participant is cognitively healthy. A participant who scores between 18 and 25 out of 30 can be considered to suffer from a mild cognitive impairment. The test comprises of short-term memory recall, visuospatial abilities, executive functions, attention, concentration, working memory, language and orientation to time and place (Nasreddine et al., 2005). The participants were given the 7.1 version of the test on M1. On M2 and M3 respectively versions 7.2 and 7.3 were used. Those different versions of the same test are designed for longitudinal studies with a timeframe of approximately three months in between the different tests. This higher reliability in re-testing, combined with a higher sensibility to the earlier stages of Alzheimer's disease (Nasreddine et al., 2005), is the reason we opted for the MoCA rather than the Mini-Mental State Examination (MMSE) (Folstein et al., 1975). Because this test was conducted three times, participants who showed an increase in test score, even though their score on M1 was insufficient, were included in the longitudinal study and additional reasons behind the lower score on the first moment were inquired after. Additionally, a language pathologist evaluated MoCA test and result individually to assure the supposed cognitive health of every participant.

**Figure 2**

*An Example of A Possible NWT2, Consisting of 6 images (One Person, Two Similar Animals, One Vehicle. Two Objects)*

### *2.2.5* GDS

Sheikh & Yesavage (1986) developed the Geriatric Depression Scale (GDS) to determine a possible depression in especially older people. Several versions of the test are available; for this study we used the version containing 15 questions so as to minimalize the total duration of the experiment as much as possible. Participants with a score of 6 or higher on 15 are considered to show signs of depression. Given the fact that we conducted this GDS on the three occasions, we did include participants that showed a higher score on one of the occasions.



### *2.2.6 Digit Span tasks*

In order to assess the working memory capacities of each participant, we conducted the digit span test (backward, forward and sequenced (Wechsler, 2008)). The three tests comprise of an increasing set of numbers that have to be repeated, with two sets of numbers for each 'level'. If a participant is wrong on both sets of the same level, the test is stopped and the next type of digit span task is commenced. In the forward version of the task, the numbers have to be repeated in the same order. Working memory is tested more in the backward version of the task that requires the numbers to be repeated in reversed order. Even more capacity of the working memory is needed for the sequenced digit span task in which the numbers have to be repeated in increasing order. On the three occasions the participants were given the same set of numbers.



## 2.3 DESIGN & PROCEDURE

We conducted a longitudinal study with different age groups who were tested three times (M1, M2, M3) over a period of six months. The same procedural steps were followed every time in order to assess possible changes in their cognitive, motoric and linguistic capacities. The entire process was registered with the keystroke logging tools ScriptLog and Inputlog. These tools, developed respectively at the University of Lund (Andersson et al., 2006) and Antwerp (Leijten & Van Waes, 2013), were created to log and analyse keystrokes and mouse movements made on the computer. Hence, the participant's entire writing process can be captured. The procedure consisted of three computerised tasks: (1) a typing test, (2) two narrative writing tasks and continued on paper with (3) a questionnaire, (4) the MoCA, (5) three digit span tasks and (6) the GDS. Additionally, an informed consent had to be signed every time at the beginning of the experiment and additional information was provided upon request.

### *2.3.1 Typing test*

The participants were asked to read the instructions on the screen carefully, whereafter the researcher repeated what needed to be done in order to assure the full compression of the task. Participants followed the instructions on the screen.

### 2.3.2 *Narrative writing tasks*

Each trial, the participants were presented with two NWTs: first the NWT1, followed by the more complex NWT2. Before the beginning of each task, the participants were asked to write (type) a story, using the prompted images. For NWT1, the participants were requested to write a narrative, based on the given images, in the first person singular. They received an example so as to show how (not) to incorporate the images and to indicate the desired text length. However, an exact word count limit or time limit was not given. They were asked to use the names of the images in their story, single or plural, as long as they did not merely describe the images. For the NWT2, the instructions were almost identical. Again, they were asked to write a story based on the prompted images. This time, however, the use of a past tense was required and it needed to be written in the third person singular (he / she / it). Again, the exemplifying text was shown and the same instruction was given. If a participant had written relatively few words (according to the researcher present), they were instructed for their second task to try and write as much as was in the example text and not to summarise their story into two sentences, without given explicit word count instructions. Even though some participants still remembered the instructions of the task, instructions were repeated entirely the various test moments.

Upon starting the experiment, the researcher opened the designed module in ScriptLog (Andersson et al., 2006), providing a blank screen in between the two tasks. After the instruction was given, the researcher pressed the 'next' button and the images and text box became visible. When the participant indicated that he/she was ready, the researcher used the 'next' button to go to the next blank screen.

### 2.3.3 *Questionnaire, MoCA, working memory test, GDS*

The experiment continued on paper. Participants were instructed to fill in the questionnaire, with the researcher helping out when needed. The cognitive screening test (the MoCA), the working memory tests (the digit span tasks), and the geriatric depression scale (GDS) were conducted in accordance with the official guidelines.



**Table 2**

*Selection Of Product Variables That Give Insight Into The Cohesion Within The Narratives, as replicated with permission from (Paesen et al., n.d.)*

| Variable | Explanation |
|---|---|
| **Lexicon** | |
| Cohesion | A score that indicates how cohesive a text is perceived to be. |
| Total words | The total number of words per text, including corrections if needed. |
| Word length | The average number of letters per word |
| Word frequency of nouns | The log frequency of the nouns in the text |
| Type token ratio of words | The number of different words (types) divided by the total number of words (tokens) |
| Density of adjectives | Proportion of adjectives on the total number of words (i.o.l.r.) |
| Density of nouns | Proportion of nouns on the total number of words (i.o.l.r.) |
| Density of verbs | Proportion of verbs on the total number of words (i.o.l.r.) |
| Density of concrete verbs | Proportion of concrete verbs on the total verbs (i.o.l.r.) (example verbs: to smell, to freeze) |
| Density of abstract verbs | Proportion of abstract verbs on the total verbs (i.o.l.r.) (example verbs: to contribute, to hope) |
| **Grammar** | |
| D-Level | The 'Development Level' or a measure of syntactical complexity. In a text, every sentence receives a complexity score, which results in a text average. |
| Distance subject-verb | The distance between two elements of the sentence that belong together (subject-verb). The further those two are apart, the more difficult a sentence is to interpret. The score in T-Scan reflects the average number of words between the subject and verb in the text. |
| Distance determiner-noun | The distance between two elements of the sentence that belong together (determiner-noun). The further those two are apart, the more difficult a sentence is to interpret. The score in T-Scan reflects the average number of words between the determiner and noun in the text. |
| Density of personal references | Density of references to people (personal and possessive pronouns, nouns referring to a person and names) |
| Density of conjunctions | Proportion of conjunctions on the total number of words (i.o.l.r.) |

*Note*: i.o.l.r. = indicator of lexical richness

### 2.4 DATA PREPARATION AND STATISTICAL ANALYSIS

The written narratives – the product data – were proofread and spelling or typing errors in the text were corrected. This was necessary so as to disambiguate the given words and texts for the automated processing with the computer program T-Scan: a tool for analysing Dutch texts (Pander Maat et al., 2014). T-Scan provides information on word, sentence and text level. We focussed on cohesion variables found in Paesen and colleagues (n.d.) such as word frequency, the ratio of word type occurrence, and the Type Token Ratio. The variables were divided into two categories: lexicon and grammar. With the use of Comproved (Lesterhuis et al., 2016), the texts were also assessed by 44 volunteers (all languages' students) and given a cohesion score (cf. Table 2 & 4).

**Table 3**

*Overview Of The Multilevel Models Used In The Statistical Analyses (both process & product)*

| Model | General overview of the data |
|---|---|
| 0 | Random variance |
| 1 | Effect of age |
| 2 | Effect of age and sex |
| 3 | Effect of age, sex and task |
| 4 | Effect of age, sex, task and moment |
| 5 | Effect of task and moment, interaction effect (age x sex) |
| 6 | Effect of moment, interaction effect (age x sex x task) |
| 7 | Interaction effect (age x sex x task x moment) |

Moreover, in order to obtain a process-oriented approach, we selected various process measures, with a focus on (a) individual motor characteristics of typing, (b) general fluency measures both in the beginning of the writing process and globally throughout, by measuring active writing time, pause distribution, utterance length and number (bursts), and (c) word retrieval. The obtained process data was analysed with Inputlog 7, on a general, summary, pause and word pause level, while taking the personal typing speed into consideration. From these analyses, the relevant variables were selected with the use of a correlation and factor analysis in R (a full description of the variables and their definition can be found in **Appendix D**). Due to the complex structure of the data, we relied on multilevel modelling in RStudio to analyse the variables. Our data is



hierarchically structured, with texts nested in participants. Hence, the influence of moment, sex, task and/or age could be incorporated when studying the effects of ageing on spontaneous written language generation, providing a richer basis for interpretation and a smaller opportunity of Type-I errors to occur (Hox, 2002; Quené & van den Bergh, 2008). We developed seven different models (see Table 3) that could explain the variance in the data. As can be seen in Table 3, Model 0 does not contain any explanatory variables, only random variance. Models 1, 2, 3 and 4 take into account the main effects. In models 5 to 7, interaction effects are added to the model with main effects. The models were compared based on the change in log-likelihood ratio. To determine the model with the best fit, Chi-square test statistic was used.



## 3. RESULTS

The spontaneous written texts will be analysed in light of both the writing product and the writing process in order to provide insight into the test-retest reliability of the writing. A full output of the multilevel analyses and estimates of fixed effects of the best fitting model can be found in **Appendix E. An overview of the product characteristics can be found in Table 4.**

### 3.1 WRITING PRODUCT

#### *3.1.1 Lexical*

*Cohesion, Total number of words, Type token ratio, Density of adjectives, Density of verbs, Density of abstract/concrete verbs*: as the null-model turned out to be the best model, the conclusion can be drawn that neither age, sex task nor moment can explain the variance in the data.

For *Word length* only the third model was significant, implying a main effect of task but not sex nor age: $\chi2(6) = 98.341$, $p < .001$. For *Word frequency of nouns*, the third model was the best model, $\chi2(6) = 113.237$, $p < .001$, even though the second and sixth were also significant, implying an main effect of sex and task. A significant increase in word length (est.: 0.252) and decrease in word frequency (est.: -0.254) can be noticed for NWT2; women write shorter words compared to men (est.: -0.122). The test moment did not impact the results.

*Density of nouns:* The first and second model are not signficant; hence there is no main effect of age nor sex. The third model is the best model, implying a main effect of task: $\chi2(6) = 14.000$, $p < .001$. A significant increase in noun density can be noticed for NWT2 (est.: 1.292).

**Table 4**

*Overview Of Product Characteristics Used In The Analysis Of The Tasks NWT1 & NWT2*

| Age | Sex | N | Moment | NWT1 Cohesion | NWT1 Total words | NWT1 Word length | NWT1 Word frequency | NWT2 Cohesion | NWT2 Total words | NWT2 Word length | NWT2 Word frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 50-59 | Female | 16 | M1 | -0.24 | 263.13 | 4.22 | 4.81 | -0.05 | 253.50 | 4.50 | 4.57 |
| | | | M2 | 0.04 | 226.38 | 4.24 | 4.93 | 0.03 | 214.00 | 4.47 | 4.61 |
| | | | M3 | 0.13 | 234.38 | 4.21 | 4.90 | 0.01 | 240.81 | 4.48 | 4.65 |
| | Male | 7 | M1 | -0.05 | 249.50 | 4.42 | 4.74 | 0.10 | 233.00 | 4.65 | 4.48 |
| | | | M2 | -0.10 | 220.38 | 4.27 | 4.75 | 0.44 | 241.00 | 4.60 | 4.45 |
| | | | M3 | 0.13 | 225.63 | 4.23 | 4.90 | 0.12 | 215.13 | 4.53 | 4.46 |
| 60-69 | Female | 12 | M1 | -0.10 | 194.92 | 4.24 | 4.89 | 0.15 | 197.50 | 4.56 | 4.60 |
| | | | M2 | 0.24 | 192.92 | 4.26 | 4.87 | -0.08 | 193.83 | 4.54 | 4.58 |
| | | | M3 | 0.10 | 206.83 | 4.30 | 4.85 | 0.17 | 214.67 | 4.54 | 4.53 |
| | Male | 11 | M1 | -0.14 | 219.10 | 4.35 | 4.81 | -0.17 | 225.50 | 4.59 | 4.38 |
| | | | M2 | 0.08 | 179.20 | 4.41 | 4.57 | -0.11 | 200.30 | 4.62 | 4.45 |
| | | | M3 | -0.11 | 197.30 | 4.36 | 4.69 | 0.11 | 199.90 | 4.64 | 4.51 |
| 70-79 | Female | 7 | M1 | 0.39 | 210.71 | 4.31 | 4.80 | 0.04 | 239.14 | 4.53 | 4.62 |
| | | | M2 | 0.58 | 205.71 | 4.27 | 4.80 | 0.05 | 208.14 | 4.48 | 4.71 |
| | | | M3 | 0.10 | 188.14 | 4.30 | 4.77 | -0.64 | 194.86 | 4.56 | 4.56 |
| | Male | 5 | M1 | -1.10 | 238.40 | 4.56 | 4.69 | -0.32 | 207.20 | 4.70 | 4.56 |
| | | | M2 | -0.65 | 260.80 | 4.32 | 4.67 | -1.05 | 206.00 | 4.44 | 4.52 |
| | | | M3 | 0.53 | 263.60 | 4.26 | 4.73 | -1.25 | 169.60 | 4.42 | 4.64 |

*Note*: '*Word frequency*' is the '*Word frequency of nouns*' and '*Total words*' refers to the '*Total number of words*'

4

### 3.1.2 Grammatical

*D-Level* and *Density of conjunctions*: as the null-model was the best model, the conclusion can be drawn that neither age, sex task nor moment can explain the variance in the data.

*Distance subject-verb*: The first model was not significant, the second and third were, with the third model explaing most variance in the data: $\chi 2(6) = 6.915$, $p = .009$. Therefore, age has no main effect, whereas sex and task do have a main effect. A significant increase in distance in the NWT2 (est.: 0.205) and for women (Est: 0.307). The test moment did not impact the results.

*Distance determiner-noun:* Only the fifth model was significant $\chi 2(8) = 3.946$, $p = .047$, which implies a main effect of the interaction between sex and age. There was an increase in distance in NWT2 (est.: 0.436), and a decrease in distance for women upon ageing (Est. -0.007)

*Density of personal references*: The first model was not significant; the second model was, implying there is no effect of age, but sex has a main effect: $\chi 2(5) = 8.025$, $p = .005$. Women write significantly fewer personal references (est.: -14.250). The test moment did not impact the results.

## 3.2 WRITING PROCESS

The analysis of the writing process is divided into different parts. First, the test-retest reliability of the individual hand mother characteristics is checked by measuring the typing speed over the three test moments. Afterwards, general insight into the writing process and pausing behaviour is given with a multilevel analysis. Lastly, a factor and multilevel analysis provide additional insight into the test-retest reliability of the writing process, both in the initial phase of the writing process and the middle phase, and word retrieval is checked.

### 3.2.1 Typing speed

In order to assure that the typing speed of the participants remained stable over the course of the three test moments, a correlation analysis in R-studio was run. A higher score on this test corresponds with a longer interkey interval and thus a slower typing speed. The average typing speed, within each group and participant, remained stable over the course of the experiments, indicating it is a measure that can be used in test-retest situations. Older, especially female, participants have a longer interkey interval, leading to a slower typing speed.



**Table 5**

*The Correlation And Its Significance Of The Typing Speed Over The Course Of The Three Test Moments.*

| | Typing speed M1 | Typing speed M2 | Typing speed M3 |
|---|---|---|---|
| Typing speed M1 | 1.00 | | |
| Typing speed M2 | 0.98** | 1.00 | |
| Typing speed M3 | 0.98** | 0.98** | 1.00 |

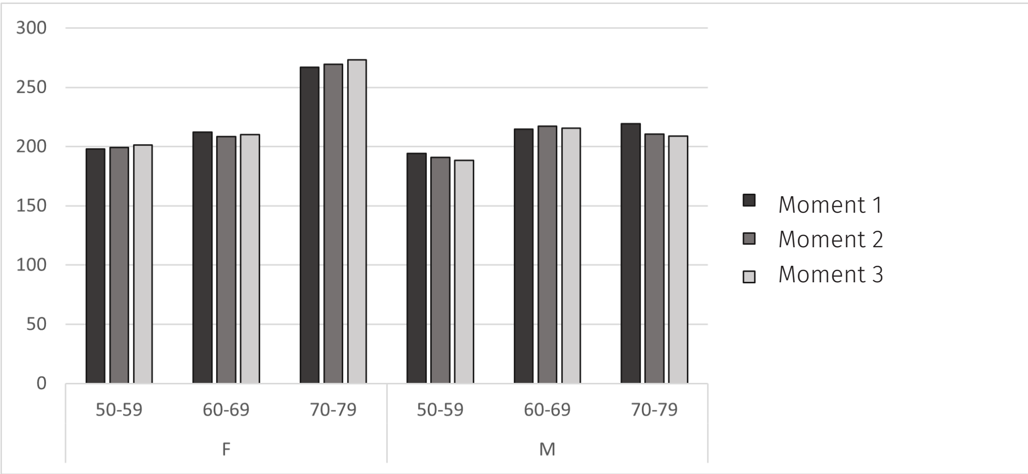
*Note*: M1= moment 1, M2 = moment 2, M3 = moment 3; **The correlation was highly significant with a score <0.01.

4

The results (see Table 5 and Figure 3) indicate that the typing speed over the three test moments remained stable, and that therefore only one measure is necessary in test-retest situations. For the continuation of the analysis, therefore, the log mean of the typing speed at M1 was used.

**Figure 3**

*The Test-Retest Reliability Of The Typing Speed At The Three Trials For The Different Age Groups*



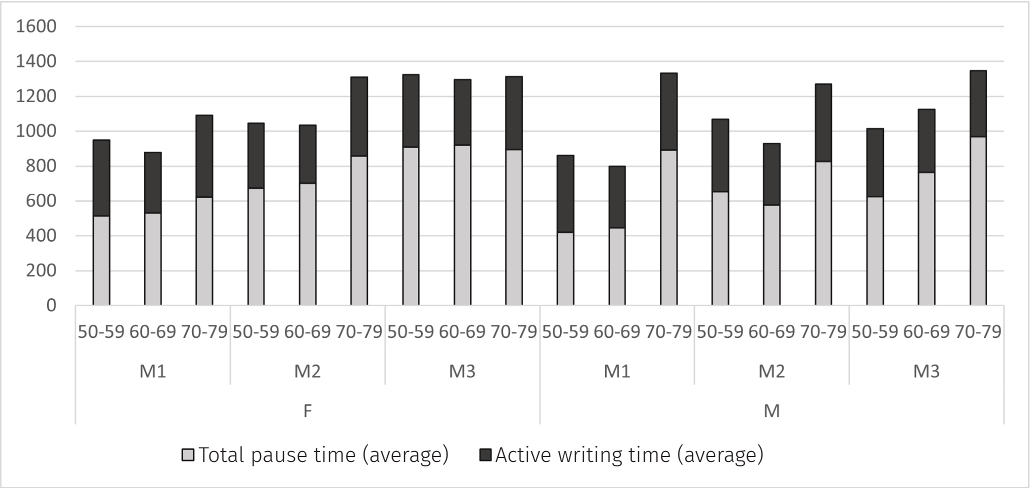

*Note*: M = male, F = female

### 3.2.2 General process information

Total time in document and Total pause time: model 4 explains most of the variance in the data. For Active writing time the null model turned out to be the best model; hence, the conclusion can be drawn that neither age, sex task nor moment can explain the variance in the data (see Figure 4).

**Figure 4**

*Total Time (S) In Document, Portrayed By The Active Writing Time And The Total Pause Time, Per Moment, Sex And Age Group*




*Note*: M1= moment 1, M2 = moment 2, M3 = moment 3, M = male, F = female

As the pause data differ interpersonally, additional analysis of the pause data provides insight into the pausing behaviour of the participants. The moments of inactivity were divided into four categories (see Figure 5).

The first type of pauses were the purely motoric pauses, each pause having a duration less than 200 milliseconds. These are the pauses that are a natural part of the typing process, both in experienced writers and in the more novice typists. The second and third type of pause is the formulation pause, the time a participant thinks about the spelling or formulation of a word or sentence; a formulation pause in this sense either takes less than 1000 ms or 2000 ms. Lastly, we

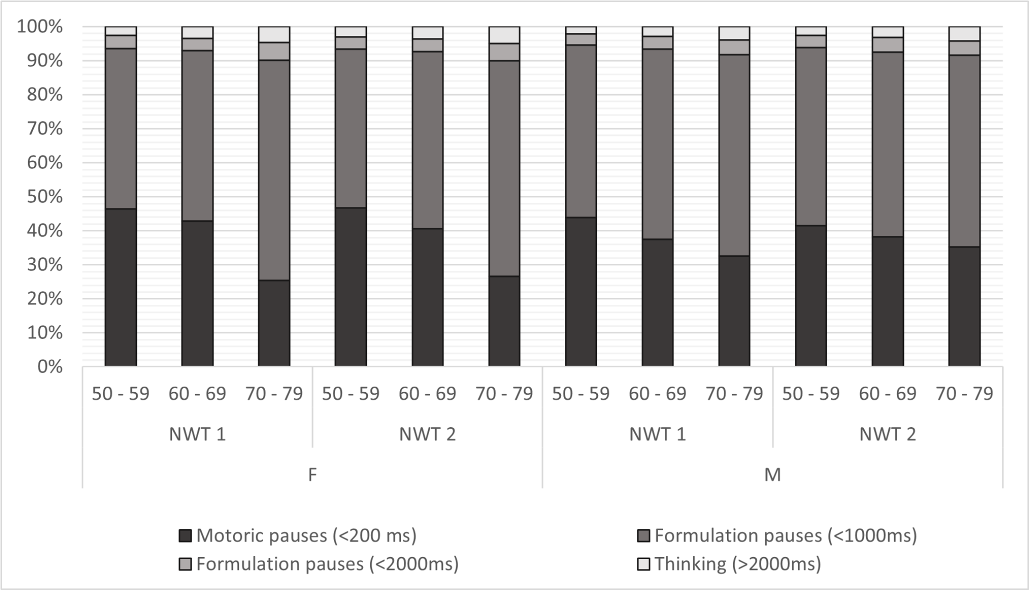


have indicated all 'thinking' pauses; the pauses longer than 2 seconds in which the participant reflects on his or her writing.

Notably, despite the fact that the typing speed of the participants was taken into consideration, the number of thinking pauses increase with age. The spontaneous language generation in younger adults is very fluent, with only few longer pauses. Older adults, even though computer literate, tend to take longer breaks throughout their writing process.

**Figure 5**

*The Division Of The Different Types Of Pauses Found Within The Writing Process*




*Note*: M = male, F = female

### 3.2.3 Test-retest reliability of process measures

After running a correlation analysis in R Studio, most process variables that had a significantly high correlation with another variable were discarded. A Factor analysis provided further insight into the amount of variance that can be explained because of certain variables. 76% of the variance in the data can be explained on the basis of the following three factors (see Table 6 and **Appendix D**).

**Table 6**

*Selected Variables And Factors, Based On The Factor Analysis*

| | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| **Factor 1: Fluency of utterance total document** | | | |
| Mean length of utterance | **0.92** | -0.59 | 0.09 |
| Keystrokes (incl. spaces) per minute | **0.95** | 0.28 | 0.02 |
| Words written per minute during the writing process | **0.96** | 0.26 | 0.02 |
| **Factor 2: Fluency of utterance at the start** | | | |
| Number of keystrokes per minute in interval 1 (the start) | 0.31 | **-0.66** | -0.28 |
| Number of utterances per minute | 0.12 | **0.92** | -0.09 |
| **Factor 3: Word retrieval** | | | |
| Ratio keystrokes product/process | 0.05 | 0.06 | **0.63** |
| Mean pause time (in s) within words | -0.06 | -0.12 | **0.76** |

*Mean length of utterance*: Both the first and the second are not significant, the third is, implying that there is a main effect of task but not of age and sex. The third model is the best model: $\chi2(6) = 11.783$, $p < .001$. A significant increase in keystrokes can be noticed for NWT2 (est.: 11.604) and a non-significant decrease with age (est.: -1.247).

For both *Keystrokes (incl. spaces) per minute* and *Words written per minute during the writing process*: The first, thrid, fourth, sixth and seventh model are significant, indicating that age, task, moment and the interaction between them have a main effect. The sixth model explains most variance in the data, respectively: $\chi2(12) = 15.822$, $p = .001$ and $\chi2(12) = 13.800$, $p = .003$, With age, the number of keystrokes per minute decrease (est.:-1.585), the NWT1 triggers fewer keystrokes per minute than NWT2 (est.: -123.972). As the trials continue, fewer keystrokes are written (est.:-7.255).



The interaction between age and task does implicate that with age, more keystrokes are entered in the NWT2 (est.: 1.485). A small decrease in words per minute can be noticed with age (est.: -0.275), and fewer words were written during NWT1 (est.: -2.959) and the following test moment (est.: -1.300). With age in the NWT2, words do significantly increase (est.: 0.255).

*Number of keystrokes per minute in interval 1 (the start)*: Both the first and third model are significant, the second is not, implying there is an effect of age and task but not sex. The third model explains most of the variance in the data; $\chi2(6) = 329.819$, $p < .001$. With age, the number of strokes per minute in that first interval significantly decreased (est.: -1.186), and in NWT2, the number of strokes significantly increased compared to NWT1 (est.: 74.866).



*Number of utterances per minute*: $\chi2(8) = 15.741$, $p < .001$, even though the first, third, fourth and seventh model were significant, indicating the main effect of age, task, moment and the interaction effect between those, the fourth model explains most variance in the data. With age, the number increases (est.: 0.026), in the NWT2 and upon retesting it decreases (respectively, Est.: -1.294 and -0.213)

For the *Ratio keystrokes product/process* and *Mean pause time (in s) within words*: the null-model turned out to be the best model, leading to the conclusion that neither age, sex task nor moment can explain the variance in the data.

## 4. DISCUSSION

With this study we aimed to find out to what extent healthy ageing affects the test-retest reliability of spontaneous written language. In order to measure spontaneous written language, 57 participants were given two narrative writing tasks (cf. Paesen et al., n.d.) on three occasions. These narrative writing tasks elicited coherent stories based on prompted coloured images and were designed based on the principles of participation focused diagnostic tests. The results were analysed with a focus on both the writing product and process, and more specifically the (a) individual characteristics of typing, (b) general fluency measures both in the beginning of the writing process and globally throughout (e.g., by measuring active writing time, pause distribution, utterance length and number (bursts)) and (c) cohesion measures: linguistic and grammatical variables (such as idea density and density of verbs) taken from Paesen et al., (n.d.). We hypothesized that more pauses would be needed by the older participants due to word finding issues, even though the writing product would not be affected (e.g., in terms of total word count). Additionally, we hypothesized that no change in both product and process data should be visible over the course of the three trials.

With regard to the **writing product** and cohesion measures, the results are rather straightforward. First of all, the multilevel analysis indicates that none of the variables were impacted by the test moment, which leads to a great test-retest reliability in longitudinal studies. Additionally, most lexical and grammatical product variables were not impacted by age, sex, task or test moment. Other effects on word length, word frequency and density of nouns can be explained by the task type, as the NWT2 triggers more and longer nouns with a lower frequency. Our results indicate that our male cohort wrote more complex sentences, with a greater distance between the subject of the sentence and the main verb and that they use fewer personal references. A study by Newman et al. (2008) indicates that the sex differences will increase with tasks that place fewer constraints on language use, supporting our findings in the fact that it remains important to distinguish between sexes in the analysis of discourse. Despite the fact that our study could not establish a clear pattern within the sex differences due to the method, our study does continue to stress the need to differentiate between sexes and suggests future studies with larger cohorts within certain age groups to study these effects even more.

When studying the changes in the **writing process**, a great test-retest reliability for general process measures such as typing speed and active writing time. Independent of the total pause



time or number of pauses required by the participant, our results indicate that the active writing time remained stable over all participants, test moments and tasks, indicating its reliability in longitudinal research. Additionally, the length of that writing process did not affect the cohesion of the stories, as this measure also remains stable. The number and length of pauses did increase with age, a phenomenon which is in line with previous studies, that can partly be attributed to the increased difficulty in lexical retrieval upon ageing (Susan Kemper & Sumner, 2001; Paesen & Leijten, 2019; Shafto & Tyler, 2014). These changes and the slowing are also consistent with the decreasing cognitive functioning. Previous studies have shown that temporal aspects of writing (such as pausing behaviour) are clearly sensitive to cognitive impairment, and therefore, possible changes in healthy ageing cannot be disregarded. Nevertheless, a test-retest reliability could not be established for several variables in the first narrative writing tasks, namely *'Words written per minute during the writing process', 'Keystrokes (incl. spaces) per minute'* and *'Number of utterances per minute'.*



This study set out to determine the **test-retest reliability** of spontaneous written language in healthy ageing and to standardise these materials so as to allow them to potentially be used in differential diagnostics. To that extent, our results clearly imply the reliability of this test with regard to retesting in a longitudinal setting due to its great intra-personal stability. This stability can be found in the writing product and in the language bursts, indicating the possibilities to potentially make a differential diagnosis and / or retest a participant / patient. Compared to other existing tests, such as the Cookie Theft Picture (Goodglass & Kaplan, 1983), the participant is presented with a new set of images upon every task and test moment, eliminated a possible learning effect, whilst maintaining the cohesion of the text and effort that the participant puts into it. Furthermore, due to its design, the test triggers spontaneous written language in a monitored setting. In other words, even though it can be seen as a more or less true representation of one's speech, this test does allow for researchers to have certain benchmarks to compare and direct the texts on both an intra-personal and interpersonal level. Nevertheless, it is important to stress that this result could not be replicated for the writing process in the first narrative writing task.

With regard to future studies, the writing process offers a wide range of opportunities. The nature of the changes in the writing process must be studied thoroughly in order to provide additional insight into the challenges the patient faces upon writing. Moreover, the narrative writing process needs to be normed and a baseline must be established. Furthermore, in line with the findings of Paesen et al. (n.d.) we would suggest future studies to reduce the task size. Currently, one test moment takes the participants around 90 minutes, leading to additional exhaustion and a lack of concentration. Based on our results and the lack of test-retest reliability of the first narrative writing task, we suggest that only the second, more challenging NWT is crucial to gain insight into the writing behaviour of the participant or patient. Reducing the task length could diminish the cognitive load, decrease potential stress triggered by the experiment and still lead to the desired results. Further studies are needed to try and replicate our findings and to validate the task when it only contains NWT2.

## 5. CONCLUSION

With this study we set out to construct the foundation of a new clinical toolkit for the evaluation of written spontaneous speech in healthy adults. The newly created narrative writing tasks contribute to the field by eliciting both spontaneous speech while still controlling for key variables. Our results were studied in light of both the writing product and writing processes. Based on the findings of the writing product, we conclude that our toolkit shows great test-retest reliability, allowing our toolkit to be used in longitudinal research. Writing process measures verified this finding and provide opportunities for future studies to gain additional insight into spontaneous written language generation. Our findings also point towards crucial sex differences. Even though the results were not conclusive, it was clear that men and women should not be treated as one healthy cohort, but as their individual sexes respectively. Our toolkit allows a non-intrusive, cost-effective measure of spontaneous written speech and provides the opportunity to establish linguistic biomarkers. If the language output changes over time, it could signify that a certain pathology is at hand and further cognitive research is needed for the wellbeing of the participant. Therefore, we believe it could be a promising digital toolkit to aid the screening process of language diagnostic in an ageing population. This test will never serve to replace other neurological testing; it aims to provide additional insight into the linguistic skill in order to support the clinical judgement on the cognitive wellbeing of an individual.



## 6. ACKNOWLEDGEMENTS

This research was supported by the research fund BOF-DOCPRO 2015, PS ID: 31661 of the University of Antwerp. We would also like to thank Eric Van Horenbeeck, the technical coordinator of Inputlog, and Luuk Van Waes, co-founder of Inputlog and the typing task. We thank Johan Frid and Victoria Johansson, respectively the technical coordinator and co-founder of Scriptlog. Further thanks go to Peter Mariën (†) and Dorien Vandenborre for their input during the design of the study. Last but not least we would like to thank the elderly who enthusiastically participated in this study.

## 7. DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

# 8. BIBLIOGRAPHY


Aerts, A., van Mierlo, P., Hartsuiker, R. J., Santens, P., & De Letter, M. (2015). Sex Differences in Neurophysiological Activation Patterns During Phonological Input Processing: An Influencing Factor for Normative Data. *Archives of Sexual Behavior*, 44(8), 2207–2218. https://doi.org/10.1007/s10508-015-0560-y

Alves, R. A., Castro, S. L., & Olive, T. (2008). Execution and pauses in writing narratives: Processing time, cognitive effort and typing skill. *International Journal of Psychology*, 43(6), 969–979. https://doi.org/10.1080/00207590701398951

Andersson, B., Dahl, J., Holmqvist, K., Holsanova, J., Johansson, V., Karlsson, H., & Wengelin, Å. (2006). Combining keystroke logging with eye tracking. *Writing and Digital Media (Studies in Writing),* 17, August 2016, 45–72.

Bastiaanse, R., Bosje, M., & Visch-Brink, E. (1995). Psycholinguïstische testbatterij voor de taalverwerking van afasiepatiënten (PALPA). *Psychology Press*.

Bates, E., Harris, C., Marchman, V., Wulfeck, B., Bates, E., Harris, C., Marchman, V., & Wulfeck, B. (1995). Production of Complex Syntax in Normal Ageing and Alzheimer ' s Disease. *Language and Cognitive Processes,* 10(05), 487–539. https://doi.org/10.1080/01690969508407113

Berisha, V., Wang, S., LaCross, A., & Liss, J. (2015). Tracking Discourse Complexity Preceding Alzheimer's Disease Diagnosis: A Case Study Comparing the Press Conferences of Presidents Ronald Reagan and George Herbert Walker Bush. *Journal of Alzheimer's Disease*, 45(3), 959–963. https://doi.org/10.3233/JAD-142763

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philos Trans R Soc Lond B Biol Sci*, 352(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

Choi, H. (2009). Performances in a Picture Description Task in Japanese Patients with Alzheimer's Disease and with Mild Cognitive Impairment. *Korean Journal of Communication Disorders*, 14, 326–337. http://www.e-csd.org/upload/14(3)_6.pdf

Croisile, B., Ska, B., Brabant, M.-J., Duchene, A., Lepage, Y., Aimard, G., & Trillet, M. (1996). Comparative Study of Oral and Written Picture Description in Patients with Alzheimer's Disease. *Brain and Language*, 53(53), 1–19. https://doi.org/10.1006/brln.1996.0033

Davis, B. H., & Maclagan, M. (2009). Examining Pauses in Alzheimer's Discourse. *American Journal of Alzheimer's Disease & Other Dementias,* 24(2), 141–154. https://doi.org/10.1177/1533317508328138

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6





Forbes-McKay, K., Shanks, M., & Venneri, A. (2014). Charting the decline in spontaneous writing in Alzheimer's disease: a longitudinal study. Acta *Neuropsychiatrica*, 26(04), 246–252. https://doi.org/10.1017/neu.2014.2

Forbes, K. E., Shanks, M. F., & Venneri, A. (2004). The evolution of dysgraphia in Alzheimer's disease. *Brain Research Bulletin*, 63(1), 19–24. https://doi.org/10.1016/j.brainresbull.2003.11.005

Franceschi, C., Garagnani, P., Morsiani, C., Conte, M., Santoro, A., Grignolio, A., Monti, D., Capri, M., & Salvioli, S. (2018). The continuum of aging and age-related diseases: Common mechanisms but different rates. *Frontiers in Medicine*, 5(MAR). https://doi.org/10.3389/fmed.2018.00061

Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language*, 93(1), 79–94. https://doi.org/10.1016/j.bandl.2004.08.003

Gayraud, F., Lee, H., & Barkat-defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics & Phonetics,* 25(March), 198–209. https://doi.org/10.3109/02699206.2010.521612

Goodglass, H., & Kaplan, E. (1983). *The Boston Diagnostic Aphasia Examination*. Lea and Febiger

Hayes, J. R., & Ann Chenoweth, N. (2007). Working Memory in an Editing Task. *Written Communication,* 24(4), 283–294. https://doi.org/10.1177/0741088307304826

Hier, D. B., Hagenlocker, K., & Shindler, A. G. (1985). Language disintegration in dementia: Effects of etiology and severity. *Brain and Language*, 25(1), 117–133. https://doi.org/10.1016/0093-934X(85)90124-5

Hill, U. C. (2021). *Adult Neurogenic Communication Disorders*. https://hsl.lib.unc.edu/speechandhearing/neurogenicsaphasia

Hoffmann, I., Nemeth, D., Dye, C. D., Pákáski, M., Irinyi, T., & Kálmán, J. (2010). Temporal parameters of spontaneous speech in Alzheimer's disease. *International Journal of Speech-Language Pathology,* 12(1), 29–34. https://doi.org/10.3109/17549500903137256

Holmqvist, K., Johansson, V., Wengelin, A., & Strömqvist, S. (2002). Analysing reading and writing online. In S. Strömqvist (Ed.), *The diversity of languages and language learning* (pp. 103–123). Lund University Press.

Holmqvist, Kenneth, & Johansson, V. (2005). Perceiving and Producing the Frog Story. *Perspectives on Language and Language Development,* January. https://doi.org/10.1007/1-4020-7911-7

Kemper, S., Greiner, L., Marquis, J., Prenovost, K., & Mitzner, T. (2001). Language decline across the life span: Findings from the nun study. *Psychology and Aging*, 16(2), 227–239. https://doi.org/10.1037/0882-7974.16.2.227

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, 16(2), 312–322. https://doi.org/10.1037//0882-7974.16.2.312

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing,* 26(4), 435–461. https://doi.org/10.1093/llc/fqr013

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication*, 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & De Maeyer, S. (2016). Comparative Judgment as a promising alternative to score competences. In E. Cano & G. Ion (Eds.), *Innovative Practices for Higher Education Assessment and Measurement* (pp. 119–138). IGI Global. https://doi.org/10.4018/978-1-5225-0531-0

Manders, E. (2016). *Handboek Neurologische communicatiestoornissen*. Garant - Uitgevers N.V.

Mayer, M. (1969). *Frog, where are you?* Dial Press.

McNamara, P., Obler, L. K., Au, R., Durso, R., & Alber, M. L. (1992). Declines in Connected Language Are Associated with Very Early Mild Cognitive Impairment: Results from the Wisconsin Registry for Alzheimer's Prevention. *Brain Language*, 42, 38–51. https://doi.org/10.1016/0093-934X(92)9 0055-J

Mueller, K. D., Koscik, R. L., Hermann, B. P., Johnson, S. C., & Turkstra, L. S. (2018). Declines in connected language are associated with very early mild cognitive impairment: Results from the Wisconsin Registry for Alzheimer's Prevention. *Frontiers in Aging Neuroscience*, 9(JAN), 1–14. https://doi.org/10.3389/fnagi.2017.00437

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment , MoCA : A Brief Screening. *Journal of the American Geriatrics Society*, 53(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes*, 45(3), 211–236. https://doi.org/10.1080/01638530802073712

Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty Speech in Alzheimer's Disease and Fluent Aphasia. *Journal of Speech, Language, and Hearing Research*, 28(3), 405–410. https://doi.org/10.1044/jshr.2803.405

Olive, T. (2012). Working Memory in Writing. In V. Beringer (Ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology* (pp. 485–506.). Psychology Press. https://doi.org/10.1007/978-94-010-0804-4_6

Paesen, L., & Leijten, M. (2019). Name agreement and naming latencies for typed picture naming in aging adults. *Clinical Linguistics & Phonetics,* 33(10–11), 930–948. https://doi.org/10.1080/02699206.2019.1590734

Paesen, L., Leijten, M., & De Letter, M. (n.d.). Aging and sex differences in the cohesion of written narratives. *Submitted*.



Pander Maat, H., Kraf, R., Bosch, A. van den, Dekker, N., Gompel, M. van, Kleijn, S., Sanders, T. J. M., & Sloot, K. van der. (2014). T-Scan : a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.

Sajjadi, S. A., Patterson, K., Tomek, M., & Nestor, P. J. (2012). Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. *Aphasiology*, 26(6), 847–866. https://doi.org/10.1080/02687038.2012.654933

Severens, E., Lommel, S. Van, Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, 119(2), 159–187. https://doi.org/10.1016/j.actpsy.2005.01.002

Shafto, M. A., Burke, D. M., Stamatakis, E. A., Tam, P. P., & Tyler, L. K. (2007). On the Tip-of-the-Tongue: Neural Correlates of Increased Word-finding Failures in Normal Aging. *Journal of Cognitive Neuroscience,* 19(12), 2060–2070. https://doi.org/10.1162/jocn.2007.19.12.2060

Shafto, M. A., & Tyler, L. K. (2014). Decline and Preservation. *Science*, 346(6209), 583–588. https://doi.org/10.1126/science.1254404

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 5(1–2), 165–173. https://doi.org/10.1300/J018v05n01_09

Sitek, E. J., Barczak, A., Kluj-Koz??owska, K., Koz??owski, M., Barcikowska, M., & S??awek, J. (2015). Is descriptive writing useful in the differential diagnosis of logopenic variant of primary progressive aphasia, alzheimer's disease and mild cognitive impairment? *Neurologia i Neurochirurgia Polska*, 49(4), 239–244. https://doi.org/10.1016/j.pjnns.2015.06.001

Tomoeda, C. K., & Bayles, K. A. (1993). Longitudinal effects of Alzheimer disease on discourse production. *Alzheimer Dis Assoc Disord*, 7, 223–236.

Van Waes, L., & Leijten, M. (2015). Fluency in Writing: A Multidimensional Perspective on Writing Fluency Applied to L1 and L2. *Computers & Composition*, 29.

Van Waes, L., Leijten, M., Roeser, J., Olive, T., & Grabowski, J. (2021). Measuring and Assessing Typing Skills in Writing Research. *Journal of Writing Research*, 13(x), 0–2. https://www.jowr.org/abstracts/vol13_x/VanWaes_et_al_2021_13_x_abstract.html

Veale, J. F. (2014). Edinburgh Handedness Inventory—Short Form: A revised version based on confirmatory factor analysis. Laterality: Asymmetries of Body, Brain and Cognition, 19(2), 164–177

Visch-Brink, E., Vandenborre, D., Jung de Smet, H., & Mariën, P. (2014). *Comprehensive Aphasia Test, CAT-NL*. Pearson.

Visch-brink, E., & Wielaert, S. (2011). Stoornisgerichte en/of functionele therapie voor gestoorde functies bij een verworven afasie? *Stem-, Spraak- En Taalpathologie*, 13(3), 153–172.

Waes, L. Van, Leijten, M., Pauwaert, T., & Horenbeeck, E. Van. (2019). A Multilingual Copy Task : Measuring Typing and Motor Skills in Writing with Inputlog. *Journal of Open Research Software,* 70(30). https://doi.org/https://doi.org/10.5334/jors.234

Wechsler, D. (2008). *Wechsler adult intelligence scale: WAIS IV; technical and interpretive manual* (Pearson (ed.)).

4

Will be submitted to: *Frontiers in Psychology*

# Chapter five

## A preliminary study into the use of narrative writing tasks for an ageing population in a clinical setting

# ABSTRACT

Introduction - This study aims to validate narrative writing tasks that elicit spontaneous written language as a tool for a differential diagnosis. These tasks attain to be non-intrusive, low-cost and time effective.

Method – Based on a longitudinal study into spontaneous language generation by Paesen et al. (submitted) within a healthy elderly population, four single-subject case studies were selected to validate the data within a clinical population. This study population entered the longitudinal study as healthy participants; however, based on their MoCA or GDS scores, a probable neurodegenerative and/or psychopathology was suspected. Using the narrative writing tasks, their entire typing process was logged up to six times with the keystroke logging tools ScriptLog and Inputlog; their results were analysed with the use of T-Scan in terms of both grammatical and lexical coherence. Additional cognitive screening tests were administered. A picture naming task was used to validate the added value of the narrative writing tasks to the current test batteries for establishing a differential diagnosis.

Results – Results indicate that the lexical variable density of abstract verbs leads to significant differences between our study group and the healthy control group. The density of personal pronouns did not lead to significant difference between our study populations.

Conclusion – The narrative writing tasks can be used to distinguish healthy elderly from elderly who suffer from a neurodegenerative and/or psychopathology. Further research into the density of abstract verbs as a language biomarker is crucial for a differential diagnosis.



Keywords – Ageing - Spontaneous language generation – Neurodegeneration – Differential diagnosis – Language biomarkers


# 1. INTRODUCTION

Due to the advancements in public health, society and economy, the world's population of people aged 60 and older is only growing. An estimate of 1 billion people were 60 or older in 2020, a number that is believed to increase to 2.1 billion in 2050 (WHO, 2021). Despite the increased life-expectancy, elderly are often plagued by various illnesses that lead to a serious reduction in quality of life. Some of most common mental and neurological issues encountered later on in life are cognitive impairment (like dementia), depression and stroke. Moreover, these issues often go hand in hand; being depressed could lead to the manifestation of cognitive disorders and increases the risk of a stroke (Hakim, 2011; Lopez et al., 2003; Salary & Moghadam, 2013). Having had multiple strokes earlier in life will make the patient more prone to cognitive impairment later on (Hakim, 2011) and feelings of depression are often symptomatic to the presence of a cognitive impairment or the aftermath of a stroke (Ashaie et al., 2019; Hakim, 2011; Salary & Moghadam, 2013). Therefore, differentiating between the language disturbances caused by these pathologies potentially poses difficulties. One possible, non-intrusive way to screen for potential pathologies is with the use of language analysis. Within cognitive linguistics, language is often viewed as a window to the mind (Tølbøll, 2019); linguistic changes are thought to be psychological markers of cognitive and emotional processes (Rosenbach & Renneberg, 2015; Tølbøll, 2019). Furthermore, in for instance Alzheimer's disease, changes in language and cohesion may be one of the primary characteristics of the disease. Nevertheless, even within healthy ageing, language and the way it is structured in the brain changes; therefore, even within the healthy control group, language production will be affected by age.



## 1.1 Language in the healthy ageing brain

Language generation is facilitated by various areas in the brain; however, the distribution of the responsible networks is subject to change in healthy ageing. Due to dedifferentiation, dedicated neural circuitry becomes less distinctive; it 'reflect[s] an impairment of neural resource allocation that compromises the precision and fidelity of neural representations and processes' (Koen & Rugg, 2019, p.547). Functional selectivity will decline and stimuli will elicit an increased activation of both the preferred and non-preferred brain regions. To compensate for this failing dedicated circuitry and to assure performance upon ageing, spontaneous neuroplasticity will take place. Function non-specific (bilateral) brain regions will be triggered to aid the otherwise faltering network connectivity (Meunier et al., 2014). Note that even though spontaneous neuroplasticity aids the faltering dedicated circuitry, performance will never be as it was with function specific brain-regions.

Nevertheless, not all peers seem to be affected by the age-related cognitive changes to the same degree due to the concept of cognitive reserve. Cognitive reserve is a protective mechanism that allows someone to maintain their cognitive functioning in case of ageing, disease or injury (Fleck et al., 2017). As a result, the effects of damage due to old age, an injury or neurodegenerative disease will not be noticeable for all peers at the same time, depending on the size of one's reserve. One of the factors impacting the reserve is their socio-economic status and sex (Fleck et al., 2017; Laws et al., 2018). The latter is hypothesized by Laws et al. (2018) who found that women have an advantage in verbal memory tasks, and therefore possibly a greater cognitive reserve to counter the effects of verbal memory decline due to Alzheimer's disease.

## 1.2 Language in an impaired ageing brain

### *1.2.1 Dementia of Alzheimer's type*

Dementia is generally described as a cognitive impairment which affects daily life. It is estimated that in 2015 over 47.5 million people worldwide were suffering from the consequences of dementia, a number that increases every year with 7.7 million people. There are several variants of dementia; the most known type is Alzheimer's disease (50% - 75%), vascular dementia (20%), fronto-temporal dementia (5%) and dementia with Lewy bodies (5%) (World Health Organization, 2015).

One of the most iconic longitudinal studies on Alzheimer disease and language is the nun-study (Kemper et al., 2001), in which diary samples of 180 nuns were analysed in terms of grammatical complexity and idea density (the number of new ideas within a text). They found a decrease in both grammatical complexity and idea density and claim that nuns who suffered from Alzheimer's disease already wrote more repetitive and vague texts (Kemper et al., 2001). In a study by Le et al. (2011), they focussed on the differences between healthy ageing and ageing with a neurodegenerative disease. In order to do so, they analysed the novels of three famous British authors: Iris Murdoch, who suffered from Alzheimer's disease, Agatha Christie whom is thought to have suffered from the disease during the writing of her later novels and Phillis Dorothy James (known as P.D. James) who had no proven pathology.



The grammatical complexity of the texts was scored with the use of the variable 'D-Level'. As the researchers expected, Iris Murdoch's novels showed a clear decrease in grammatical complexity as the disease progressed, compared to the rather stable result found in the other authors. In terms of lexicon, with the use of the type/token ratio and the word-type introduction rate, the extend of the vocabulary of the authors was studied. The type/token ratio refers to the 'number of unique lemmatized word-types' divided by the total number of tokens'; word-type introduction rate measures 'the cumulative number of unique lemmatized types computed at every 10.000-token interval' (Le et al., 2011, p. 440). Phillis James, who did not suffer from a neurodegenerative disease, did not display a change in vocabulary over the course of her writing career. Iris Murdoch's disease and its progression, on the other hand, can clearly be noticed by a decrease in vocabulary in her last novel 'Jackson's Dilemma'. Agatha Christie also showed a slight decrease; however, this was only in her last two novels. A study into lexical repetitions, designed to find possible issues with word-recall, revealed the significant increase of repetitions in the novels of Christie and Murdoch, compared to the stable results in the novels of Phillis Dorothy James . Lastly, analysis of the different word types (e.g., nouns, verbs, adjectives), indicated that Murdoch and Christie had the tendency to use an increased number of content verbs and decreased number of nouns. Again, no significant difference was found in the data from Phillis Dorothy James (Le et al., 2011).



### *1.2.2 Depression*

7% of elderly (aged 60 and older) are estimated to suffer from a unipolar depression (World Health Organization, 2017). The current Covid-pandemic is adding insult to injury, with social isolation and loneliness leading to even more depression and anxiety in older adults (Banerjee, 2020; Mukhtar, 2020). These mental health issues also affect the patients' cognitive processes, such as memory and concentration (Kircanski et al., 2012; Tølbøll, 2019), leading to an acceleration in brain ageing (Alexopoulos, 2019; Christman et al., 2020) and resulting often in a comorbidity with dementia (Alexopoulos, 2019). More specifically, almost 20% of patients with Alzheimer's disease suffer from a depressive disorder. Nevertheless, recent studies suggest that the origin of the depression differs in patients with a neurological disorder, indicating that the anatomic damage to the brain and the neurobiological changes that occur in dementia could be the cause for that depressive disorder (Alexopoulos, 2019).

Research into linguistic changes due to depression have indicated that both in verbal and written memory recall and in spontaneous speech, the usage of first person singular pronouns is elevated compared to healthy controls, and the number of third person pronouns decreases (Perna et al., 2019; Rosenbach & Renneberg, 2015). In a study by Tølbøll (2019), this finding was corroborated, through a systematic review of 57 studies into linguistic characteristics and depression. They found that there was an interaction with the frequency of first person pronouns and the presence/absence of a depression; people who suffer from a depression will use more first person pronouns compared to the healthy control; and the use of the first person pronoun increases with the severity of the disease. This is due to the high self-focussed state in which the individual is in. However, they do state that more nuance is needed when comparing smaller ranges of severity of depression (Pyszczynski & Greenberg, 1987; Tølbøll, 2019). Furthermore, individuals with a depression had the tendency to write significantly more negative emotion words and less positive emotion words compared to their healthy counterparts (Rosenbach & Renneberg, 2015; Rude et al., 2004; Tølbøll, 2019); however, within the patient group, the smaller differences between various severities of the disease is also more difficult to make (Tølbøll, 2019).

### *1.2.3 Stroke*

Approximately 70% of elderly aged 65 and over are estimated to suffer from a stroke at some part in their lives; the association between the prevalence of a stroke and age is therefore clear (Coco et al., 2016; Ellis & Urban, 2016). Unfortunately, a stroke will often result in other detrimental issues. One out of three patients who have suffered from a stroke, will experience aphasia most severely when the stroke strikes and directly afterwards. Additionally, the influence of a stroke on other cognitive deficits such as Alzheimer's disease has also been proven (Coco et al., 2016; Vijayan & Reddy, 2016), with motor and language issues that may also manifest themselves (Ginex et al., 2020). Aphasia is generally described as an acquired language disorder that affects various aspects of one's language. Depending on the type of aphasia, fluency, comprehension, repetition, naming and written language issues may arise (Ginex et al., 2020). For instance, when studying the language characteristics of the semantic variant of Primary Progressive Aphasia (also known as Semantic Dementia), issues with comprehension, irregular words and semantics are apparent (Agosta et al., 2010; Gorno-Tempini et al., 2004). More specifically, longitudinal research has indicated that conceptual deficits are prone to appear, with for instance the inability to picture certain elements of a lion even though the categorisation as it being an animal will still be possible (Patterson & Hodges, 2001). Other aspects of the language, such as fluency, phonology, syntax and working memory are relatively preserved (Agosta et al., 2010; Gorno-Tempini et al.,



2004). Nevertheless, previous research has also indicated that, despite the fluency, the speech of a patients who suffers from svPPA will become more vague and more difficult to comprehend (Patterson & Hodges, 2001).

## 1.3 Spontaneous language testing

The aforementioned research fields make use of a test battery that mainly focusses on one pathology and that often only test the cortical functions. Ample studies can be found that focus on picture naming (e.g., Boston Naming Test (Glosser & Kaplan, 1989)), even in order to distinguish between healthy ageing controls. These studies are crucial in our understanding of neural changes and how some systems remain untouched upon ageing. A deeper entrenchment for certain words leads to a better preservation of those words, for instance due to a higher frequency in the language, earlier acquisition in life and higher familiarity will lead to a higher name agreement for those words in picture naming tasks (Brysbaert, Stevens, et al., 2014; Brysbaert, Warriner, et al., 2014; Gilhooly & Logie, 1980; Kremin et al., 2001; Le et al., 2011; Rodríguez-Ferreiro et al., 2009). More recent studies pointed to the crucial role of the subcortical system; this system has often been neglected in neurolinguistic research because it was presumed to only be responsible for our primitive urges, not language. Several studies point towards the crucial role of the interplay between cortical and subcortical systems on lexico-semantic processing (Copland & Angwin, 2019; Murphy et al., 2021). Additionally, Murphy and colleagues (2021). Bridges & Van Lancker Sidtis, (2013) provided evidence for the importance of intact subcortical systems for the production of formulaic (and recital) speech. By studying individuals who suffer from selective brain damage or a neurodegenerative disease, more can be learned about the subcortical involvement language production. Nevertheless, only few studies have focused on the subcortical involvement in the generation of spontaneous language (Murphy et al., 2021), insight that could be crucial for differential diagnosis.



Furthermore, the quantitative and qualitative measurement of spontaneous language generation is lacking. The current diagnostic set of Dutch tests for assessing the patient's linguistic functioning in spontaneous speech – rather than contextualised tasks – is restricted and limited in its use for a wider set of pathologies. The ASTA (Analyse van de Spontane Taal) for instance is a complex test that is only academically in use. The test quantifies the number of specific morphosyntactic/lexical-semantic/phonemic elements in the spontaneous speech and allows for comparison with the discourse of healthy controls. However, even though the test allows for the analysis of speech in a more natural context, the patient's communicative competence cannot be thoroughly qualitatively assessed. A completely different test is the ANELT (Amsterdam-Nijmegen Everyday Language Test), that focuses on a more guided verbal information transmission, by giving the patient a real-life scenario and demanding them to describe what they would do. During analysis, the grammatical correctness of the utterances is disregarded. A third test, the Scenario test, offers a more complete analysis, by looking at a variety of skills (spoken, gestured and written) and focussing on the amount of abstract information that is being transmitted by the patient.

In a study by (Paesen et al., Submitted), a novel test was created in which spontaneous written language is generated through 'narrative writing tasks'. Patients are presented with a number of images depicting a single object / animal / vehicle / human. In combining those images into a narrative, the patient is forced to call upon his/her imagination and a more spontaneous language flow is triggered. Each patient received an easier narrative writing task and a more complex one (for more information: Paesen et al. (submitted)). Furthermore, in order to provide for a greater applicability in the field, the tasks were digitalised and administered on a computer. The writing product of both texts was analysed in light of the text cohesion, with a focus on grammatical and linguistic changes in the text as people age. Based on the data of 257 healthy controls, results indicated that in terms of grammar, age positively affected sentence and word complexity, with higher D-Level scores, greater distance between subject and verb and interestingly, a higher general density of personal references for women compared to men. In light of lexical changes, they concluded that as people age, texts and words became shorter. Nevertheless, a greater variety of words types was used. More specifically, women had the tendency to use less nouns compared to other word types upon ageing and more verbs and abstract verbs. The results of this spontaneous language study were combined in a normative dataset (Paesen et al., Submitted) which will form the basis for this study.

With this study we aim to test whether various linguistic variables derived with 'Narrative writing tasks' could serve as a biomarker for language disturbances in adults. This task, and the spontaneous written language it generates, aims to be an asset due to its versatility, non-academic applicability and digital nature. Based on the study by (Paesen et al., Submitted), we selected older adults who suffer from a potential neurological or psychological disorder for a preliminary study, in order to find an answer to our research question: In terms of spontaneous language generation, which parameters could serve as a language biomarker for language disturbances in older adults? Given the need for further study into subcortical measures and the useful yet



limited use of picture naming, 'we hypothesize that the picture naming task will not shed light on different language disturbances'. Furthermore, given the language changes that occur in healthy ageing and that have been found to distinguish healthy ageing from AD in terms of lexicon and grammar: 'We hypothesize that both lexical and grammatical measures are needed in order to make a differential diagnosis'. More specifically, in order to differentiate healthy participants from patients who suffer from a depressive disorder, 'we hypothesize that grammatical markers, such as the density of personal pronouns, could be used to make a differential diagnosis'. Lastly, 'we hypothesize that aphasic language (due to PPA) will be differentiated based on lexical markers'.

## 2. METHOD

### 2.1 PARTICIPANTS

The four cases selected for this article were found in the cohort of our longitudinal study (Paesen et al., Submitted); they were discarded for that study due to neurological/psychological issues encountered during the experiment. The longitudinal study started out with a total of 75 participants, who enrolled as cognitively healthy elderly. The participants were aged between 50 and 100 years old and were native Dutch speakers. They were recruited through the researcher's own network, senior organisations and during evening courses. All participants needed to be sufficiently computer literate, without any neurological or psychological issues or dyslexia. They were tested individually three times with interposes of three months and at a location of their choice – mostly their homes. 57 participants completed all three test moments and were considered cognitively healthy after testing; others did not complete the three trials or did not meet the criteria of the healthy cohort; the four cases were selected in this latter group (cf. Table 1). Note that this latter group of participants was not excluded immediately after administering the test; a language pathologist reviewed the entire procedure, tests and scoring after the experiments were done. Based on this evaluation, participants were placed into the healthy cohort or they were discarded.

At the start of the experiment, every participant had to sign an informed consent and additional information about the study was provided. Participants were informed of their scores upon their request, and if needed, they were encouraged to seek medical advice. The study was performed in accordance with the Declaration of Helsinki. All participants signed a written informed consent approved by the Ethics Committee.

**Table 1**

*Overview Of Participant Characteristics In The Case Studies; Scores On Multiple Occasions Were Obtained For Some Of Those Participants*

| | AD | AVS | JP | BDB |
|---|---|---|---|---|
| Sex | Female | Female | Male | Female |
| Age | 56 | 52 | 88 | 69 |
| Education | 12 | 19 | 10 | 12 |
| MoCA (/30) | | | | |
| Moment 1 (=M1) | 14 | 27 | 28 | 26 |
| Moment 2 (=M2) | / | 28 | 28 | 24 |
| Moment 3 (=M3) | / | / | 22 | 24 |
| GDS (/15) | | | | |
| Moment 1 (=M1) | 0 | 14 | 3 | 3 |
| Moment 2 (=M2) | / | 11 | 1 | 2 |
| Moment 3 (=M3) | / | / | 5 | 2 |
| Digit Span average | | | | |
| Forward | 5 | 7.5 | 6 | 7 |
| Backward | 5 | 8.5 | 5 | 6 |
| Sequence | 5 | 9.5 | 7 | 7 |
| Self-reported pathology | / | Depression | Cancer | Strokes |

*Note*: Scores in bold deviate from the standard. MoCA=Montreal Cognitive Assessment, version 7.1,7.2,7.3; GDS = Geriatric Depression Scale

## 2.2 MATERIALS

### *2.2.1 Written tests*

Typing test - The keystroke tool Inputlog (Leijten & Van Waes, 2013), developed at the University of Antwerp, provides the module 'copy task' which aims to describe the typing characteristics of the participants by logging and time stamping every keyboard and mouse activity. In seven consecutive tasks, given sentences, word or letter combinations need to be copied by typing, allowing the pausing and writing behaviour for different word, bigram and letter intervals to be examined (Leijten & Van Waes, 2013; Van Waes et al., 2017; Waes et al., 2019). The task ends with a questionnaire that enquires after the handedness of the participant, using the short form of the



Edinburgh Handedness Inventory (Veale, 2013).

Two narrative writing tasks (NWTs) were designed, in which participants are asked to type a coherent story based on the prompted images. These prompted images are selected from the Open Linguistic Picture Database (OLPD, www.olpd.eu) and depict a single object/being. Two types of images were selected, 'easier' images (e.g., 'cat' or 'bike') and more 'complex' images (such as 'wheelbarrow' and 'kangaroo'). The selection was based on the word characteristics in Dutch of the depicted object/being (for more information, cf. (Paesen & Leijten, 2019)). In order to avoid a learning effect, a different combination of those images was made for each NWT, each participant and test moment. The first and 'easier' NWT composes of four 'easier' OLPD images and requires participants to write in the first person singular perspective (for an example, see Figure 1). The second and more 'complex' NWT contains six more 'complex' OLPD images, and ought to be written in the third person singular and the past tense.



**Figure 1**

*An Example Of A Possible NWT1, Consisting Of Four Images (One animal, One vehicle, Two objects)*

Picture naming test - To accommodate for the picture naming task (PNT), 48 coloured images were taken from the OLPD were used; 19 highly recognisable images (HR; e.g., car) and 29 less recognisable images (LR; e.g., shark). Given that these images were presented in a randomised order, the images needed to meet certain criteria to enhance the interchangeability. The HR images are 'easier', being high frequent in the Dutch language, having a low age-of-acquisition and resulting in speeded naming times in previous spoken studies, etc. The LR images are more 'complex' in nature, having a lower frequency in the Dutch language and a higher age-of-acquisition. For more information about the PNT, cf. (Paesen & Leijten, 2019). The keystroke logging tool Scriptlog (Frid et al., 2014) was used to log and time stamp the entire writing process.

### 2.2.2 Questionnaire

With the use of a questionnaire, we were able to determine the participant's background; their education, their current job or the number of years retired (if applicable) and job history, their language and computer skills, possible neuro- or psychological issues and possible visual impairments.

### 2.2.3 Cognitive screening tests

**MoCA** - The cognitive screening tool Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005) was used to assess the cognitive health of our participants. This test was preferred over its famous counterpart the MMSE (Mini-Mental State Examination, (Folstein et al., 1975)), for three main reasons: a shorter administration time, the higher sensitivity to the earlier stages of AD and its high reliability in re-testing within a period of three months (Nasreddine et al., 2005). During the test, tasks are given on short-term memory recall, visuospatial abilities, executive functions, attention, concentration, working memory, language and orientation to time and place (Nasreddine et al., 2005). We administered the Dutch version 7.1 on M1, version 7.2 during M2 and version 7.3 accordingly on M3. For the purpose of this study, participants with a score of 26 out of 30 or higher were considered to be cognitively healthy. Participants with a score between 18 and 25 out of thirty were not placed in the healthy cohort since this score might indicate a possible mild cognitive impairment.

**Working memory test** - The forward, backward and sequencing digit span task (Wechsler, 2008) were used to assess the working memory capacity of the participants. During each of these tests, the participants are asked to repeat a set of numbers in the same, reversed or numeric order. There are two sets of numbers for each 'level'. If a participant cannot repeat the two sets of the same length in the demanded order, the test is stopped and the next type of digit span task is commenced.

**Geriatric depression scale** - The study was concluded with the Geriatric Depression Scale (Sheikh & Yesavage, 1986). This questionnaire, containing 15 questions, can be used as a screening tool for possible signs of depression. Participants needed a score of 5 out of 15 or lower to be considered psychologically healthy.



## 2.3 DESIGN AND PROCEDURE

During the three test moments, the procedure remained unaltered. During each trial, the participants started with the (1) computerised written tests – a typing task, two NWTs and a picture naming task; the remainder of the trial continued on paper: (2) a questionnaire and (3) the cognitive screening tests - the MoCA, three digit span tasks and the GDS. Keystroke logging tools Inputlog 7.0.0.0 (Leijten & Van Waes, 2013) and ScriptLog (Frid et al., 2014) logged and time-stamped every keyboard activity and mouse movement, providing researchers insight into the writing and pausing behaviour of the participants.

### 2.3.1 Written tests

**Typing test** - The participants were instructed to copy (by typing) the given sentences, words or letter combinations as fast as they could. They were asked not to correct their typing errors, since the typing speed is of greater importance than the typing correctness. Even though the instructions for each of the seven consecutive tasks were displayed on top of the screen, the researcher also provided the instructions orally to assure that these instructions were understood.



**Narrative writing tasks** - The participants were asked to write (type) two narrative writing tasks; first the 'easier' NWT1, followed by the more 'complex' NWT2. For the former task they were instructed to write a coherent story, based on the images visible on top of the screen. The story needed to be written in the first person singular. An exemplifying text was shown so as to clarify how the images needed to be incorporated – by explicitly naming them in the story, not mere describing - and so as to indicate the desired text length. No exact word count or time constraint was given. The participants were offered the opportunity to ask questions if necessary; if there were no further questions, the NWT1 could commence. When finished, the participant received the instructions for the NWT2, asking them to write the story in the third person singular and additionally in the past tense. Even though the instructions concerning text length and naming of the images remained unchanged, they were repeated before the start of the second NWT. Similarly, this procedure remained unchanged during the three test moments.

**Picture naming test (PNT)**- Participants were instructed to name the prompted picture by typing. They were told it was crucial that they used only one word to name the image (no adjectives, no articles) and that they had to name the image as fast as they could. Every participant was given a practise session with three images to make sure that they understood the instruction. Afterward the practise session a random combination of the 48 images selected from the OLPD followed.

### *2.3.2 Questionnaire & screening*

The questionnaire was administered on paper; participants were instructed to fill in the questionnaire themselves. The researcher read through the questions and assisted when needed. The three cognitive screening tests – the MoCA, digit span tasks and GDS – were administered in accordance with the official guidelines.

## 2.4 DATA PREPARATION AND ANALYSES

The data from the PNT was analysed with the use of the keystroke logging tool Inputlog (Leijten et al., 2013). More specifically, the tool comprises of several modules, of which the General Analysis is the primary one. This analysis allows access of information on a keystroke level (e.g., a letter or enter), and therefore allows us to code the different words that were produced. Writing errors and naming mistakes were not corrected but categorised, ranging from 'wrong' to 'synonym' and 'right with writing mistake'. These results were analysed in light of name agreement and object recognition (Table 2).

**Table 2**

*Variables Derived From The Picture Naming Tasks*

| Variable | Explanation |
|---|---|
| Name agreement | The percentage of words that are named correctly in a strict sense, not including variants or synonyms. |
| Object recognition | The percentage of words that are named correctly in a wider sense: including synonyms, morphological variants (e.g., dolphin – dolphins), more specific names (e.g., Dalmatian instead of dog) and more general names (e.g., boat instead sailboat) |



**Table 3**

*Selection Of Product Variables That Give Insight Into The Cohesion Within The Narratives, as replicated with permission from (Paesen et al., n.d.)*

| Variable | Explanation |
|---|---|
| **Lexicon** | |
| Cohesion | A score that indicates how cohesive a text is perceived to be. |
| Total words | The total number of words per text, including corrections if needed. |
| Word length | The average number of letters per word |
| Word frequency of nouns | The log frequency of the nouns in the text |
| Type token ratio of words | The number of different words (types) divided by the total number of words (tokens) |
| Density of adjectives | Proportion of adjectives on the total number of words (i.o.l.r.) |
| Density of nouns | Proportion of nouns on the total number of words (i.o.l.r.) |
| Density of verbs | Proportion of verbs on the total number of words (i.o.l.r.) |
| Density of concrete verbs | Proportion of concrete verbs on the total verbs (i.o.l.r.) (example verbs: to smell, to freeze) |
| Density of abstract verbs | Proportion of abstract verbs on the total verbs (i.o.l.r.) (example verbs: to contribute, to hope) |
| **Grammar** | |
| D-Level | The 'Development Level' or a measure of syntactical complexity. In a text, every sentence receives a complexity score, which results in a text average. |
| Distance subject-verb | The distance between two elements of the sentence that belong together (subject-verb). The further those two are apart, the more difficult a sentence is to interpret. The score in T-Scan reflects the average number of words between the subject and verb in the text. |
| Distance determiner-noun | The distance between two elements of the sentence that belong together (determiner-noun). The further those two are apart, the more difficult a sentence is to interpret. The score in T-Scan reflects the average number of words between the determiner and noun in the text. |
| Density of personal references | Density of references to people (personal and possessive pronouns, nouns referring to a person and names) |
| Density of conjunctions | Proportion of conjunctions on the total number of words (i.o.l.r.) |

*Note*: i.o.l.r. = indicator of lexical richness

With the use of the computer program T-Scan: a tool for analysing Dutch texts (Pander Maat et al., 2014), the text analyses process was automated on both a word, sentence and text level. In order to accommodate for the process, the first author proofread the texts and corrected the spelling and typing errors; missing words and inconsistencies in the text were left as they were. During the analyses, two types of cohesion measures were focused on – in accordance with the findings from the nun-study and the study by Le and colleagues (2011): lexicon and grammar; this was done by looking at variables such as word frequency, Type Token ratio, the distance between various word types, etc. (cf. Table 3). Based on a normative table of cohesive text characteristics in healthy ageing, developed in Paesen et al. (submitted), we were able to compare the cohesion data obtained by T-Scan. A comparative view of the different written texts can be found in **Appendix F,** an overview of the general case results can be found in **Appendix G**.

## 3. CASES

### 3.1 CASE 1 - AD

AD is a 56-year-old woman, who has had 12 years of education, works as a cleaning lady and is a self-reported monolingual. Even though she was considered to be healthy upon recruitment, the MoCA score of AD was low with 14/30. She decided to forfeit the experiment after M1.

**Table 4**

*AD - Scores for Name Agreement and Object Recognition*

| | Name Agreement | Object Recognition |
|---|---|---|
| Highly Relatable images (%) | 84.21 | 100 |
| Less Relatable images (%) | 72.41 | 82.76 |

*Note*: HR images = Highly recognisable images; LR = Less recognisable images.

#### *3.1.1 Picture Naming Task*

The PNT could not be used (cf. Table 4) to distinguish AD from the healthy population (cf. Paesen et al., 2019), even though the LR images did trigger more alternative names, with almost 17% percent of the words named incorrectly.



#### *3.1.2 Clinical analysis of the narrative writing task.*

Analysis of the text suggests a language deprivation due to a cognitive deterioration. The participant's age, weaker MoCA score and her clinical language profile are indicative of symptoms that align with semantic dementia. Imagining and pathological findings, together with an assessment by a neurologist and/or neuro psychologist, are necessary for a diagnosis.

In NWT1, no phonological issues were noteworthy. In terms of grammar, even though it is fluent, the structure is repetitive, short and stereotypical (subject – verb – etc). Perseverations on a word level within a sentence can also be noticed; e.g., 'kijkt rechts kijkt' (= 'looking right looking') and 'Ik zie ook een ring die ik graag zie' ( = I also see a ring that I like to see). Additionally, omissions of articles (e.g., 'handig) and reciprocal verbs (e.g., bevindt) can be noted. Semantics, too, are fluent. We note enumeration within spontaneous speech, empty speech (e.g., 'voor op reis te gaan' = 'to travel with') and desultory language use. Semantic impoverishment can be observed; e.g., 'vliegen' (to fly) appears thrice within the same sentence. 'Bij' (='at') is used instead of 'binnen' (= 'within') in the sentence 'Ik zal bij enkele weken …' (= I will go 'at' a few weeks)



In the NWT2, in terms of phonology, grapheme deletion occurred ('store' instead of 'stoere'). Grammatically, we notice short sentences (e.g., 'die grijs had rode handvaten' (= 'that grey had red handles'), that could be attributed to working memory. Semantic language use is impoverished, with irrelevant information and a perseveration on adjectives (colours). The spelling mistakes in 'cobwboy' (= 'cowboy') and 'hij vind' (= 'hij vindt') could be attributed to her educational background.

#### *2.1.3 Cohesion analysis of the narrative writing task*

The results of the cohesion analyses (cf.:Table 5) revealed that AD's texts, similar to AVS, differed only on one variable. In the NWT2, the density of abstract verbs was much higher than the norm. The other variables are unaffected.

**Table 5**

*Narrative Writing Scores Of AD On The First Test Moment*

| | NWT1 | NWT2 |
|---|---|---|
| **Lexical** | | |
| Total words | 141 | 91 |
| Word length | 3.89 | 4.23 |
| Word frequency of nouns | 5.07 | 4.31 |
| Type token ratio of words | 0.48 | 0.59 |
| Density of adjectives | 85.11 | *109.89* |
| Density of nouns | 148.94 | 208.79 |
| Density of verbs | 205.67 | 263.74 |
| Density of concrete verbs | 7.09 | 43.96 |
| Density of abstract verbs | 85.11 | **197.80** |
| **Grammatical** | | |
| D-Level | 1.92 | 2.36 |
| Distance subject-verb | 1.45 | *0.86* |
| Distance determiner-noun | 0.17 | 0.19 |
| Density of personal references | 120.57 | 153.85 |
| Density of conjunctions | 49.65 | 43.96 |

*Note*: The scores in bold diverge from the norm as defined in the normative tables; the cursive scores indicate those scores that are on the verge of the norm.



## 2.2 CASE 2 - AVS

AVS is a 52-year-old, self-reported healthy woman who has had 19 years of education and is active as an analyst. She rated her language skills of French and English to be above average. After the experiment was completed, AVS reported to have been suffering from a depression the last couple of years and as a result was taking anti-depressants. She had no issues with the tests and the MoCA; however, her psychological issues were reflected in her GDS score. She scored 14/15 the first test moment and 11/15 the second time. Due to the emotions the GDS evoked, AVS decided to forfeit the experiment after the second moment.

### *2.2.1 Picture Naming Task*

AVS was able to name most images correctly, with only a few mistakes being made on both M1 and M2 and with both the HR and LR images. Compared to the data of Paesen et al. (2019), her name agreement and object recognition scores do not differ from that of her peers, indicating that the psychological issues she has are not visible with the use of this test.



**Table 6**

*AVS - Scores for Name Agreement and Object Recognition*

| | Name Agreement | | Object Recognition | |
|---|---|---|---|---|
| | Moment 1 | Moment 2 | Moment 1 | Moment 2 |
| Highly Relatable images (%) | 89.47 | 94.74 | 94.74 | 100 |
| Less Relatable images (%) | 93.10 | 86.21 | 100 | 93.10 |

### *2.2.2 Clinical analysis of the narrative writing task*

One phonological paragraphia ('staart' instead of 'straat' – English: 'steert' instead of 'street') was obtained in M1. No other phonological, semantic or grammatical disturbances could be observed in NWT1. No clinical language deficits were registered in NWT2. Due to these finding, we conclude that AVS might suffer from probable anomic aphasia.

### 2.2.3 Cohesion analysis of the narrative writing task

Few spelling mistakes were made (cf. **Appendix F**); the only errors had to do with missing or wrong punctuation and capitalisation. Compared to the normative dataset, a clear decrease in the density of abstract verbs in the NWT2 during M2 could be noticed. All other variables are according to the norm, with noticeable longer texts and an increase in concrete verbs in M2.

**Table 7**

*Narrative Writing Scores Of AVS*

| | NWT1 | | NWT2 | |
|---|---|---|---|---|
| | Moment 1 | Moment 2 | Moment 1 | Moment 2 |
| **Lexical** | | | | |
| Total words | 106 | 173 | 96 | 102 |
| Word length | 3.88 | 4.39 | 4.75 | 4.19 |
| Word frequency of nouns | 4.67 | 4.55 | 4.33 | 4.60 |
| Type token ratio of words | 0.62 | 0.66 | 0.61 | 0.65 |
| Density of adjectives | 56.60 | 63.58 | 62.50 | 58.82 |
| Density of nouns | 226.42 | 219.65 | 187.50 | 215.69 |
| Density of verbs | 179.25 | 184.97 | 197.92 | 166.67 |
| Density of concrete verbs | 9.43 | 23.12 | *0.00* | 9.80 |
| Density of abstract verbs | 113.21 | 92.49 | 145.83 | **68.63** |
| **Grammatical** | | | | |
| D-Level | 1.10 | 2.08 | 2.22 | 1.60 |
| Distance subject-verb | 2.13 | 1.42 | 1.31 | 1.82 |
| Distance determiner-noun | 0.09 | 0.13 | 0.29 | 0.00 |
| Density of personal references | 122.64 | 132.95 | 125.00 | 166.67 |
| Density of conjunctions | 47.17 | 40.46 | 62.50 | 58.82 |

*Note*: The scores in bold diverge from the norm as defined in the normative tables; the cursive scores indicate those scores that are on the verge of the norm.



## 2.3 CASE 3 - JP

JP was a retired 88-year-old man who participated in the longitudinal study. When we started the testing, he was considered healthy, with no history of psychological or neurological illnesses, no writing disorder and Dutch as his mother tongue. He had excellent knowledge of French, having lived in the French-speaking part of Belgium for over a decade. He had had 10 years of education, followed by an active and social career, and had been retired for 23 years when entering the study. A week before the third test moment took place, JP received a cancer diagnosis. The researcher conducting the third test moment noticed clearly that JP was very confused, could not remain focused and was emotionally affected. Three weeks after the third test moment, JP died due to the illness.

### 2.3.1 Picture Naming Task

JP's lowest score was obtained during M1, before the diagnosis (cf Table 8). M2 resulted in the highest recognition of the images. The effect of the diagnosis is not noticeable in M3; there is a decrease in scores compared to the second test moment, however, results are still remarkably higher than in the first test moment. The highly related (HR) images triggered consistently a higher correct response rate compared to the less related (LR) images.



**Table 8**

*JP - Scores for Name Agreement and Object Recognition*

| | Name Agreement | | | Object Recognition | | |
|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M1 | M2 | M3 |
| Highly Relatable images (%) | 73.68 | 89.47 | 84.21 | 84.21 | 100 | 94.74 |
| Less Relatable images (%) | 72.41 | 82.76 | 79.31 | 79.31 | 89.66 | 82.76 |

*Note*: M1 = Moment 1; M2 = Moment 2; M3 = Moment 3

### 2.3.2 Clinical analysis of the narrative writing task.

Analysis of the text revealed JP's writing to be grammatically, semantically and phonologically intact. The NWT1, in the three test moments, contained negligence errors which cannot be attributed to aphasia, such as grapheme omissions and addition.

The NWT2 contained during the first test moment only negligence errors (not aphasia), such as grapheme addition and omission, reduplication (e.g., 'zeilbbot' instead of 'zeilboot') and verb omission. The M2 text contained reduplication errors ('onntelbare' instead of 'ontelbare') and syllable deletion (e.g., 'tatoeen' instead of 'tatoeëren'). The M3 text contained sloppiness mistakes ('ven' instead of 'van'), reduplication (e.g., 'gerraakt', 'gebraccht', 'kreeeg'), words were pasted together (e.g., 'geredwerd' instead of 'gered werd') and capital letters were used incorrectly (e.g., 'Kruiwagen' which should be written without a capital).

**Table 9**

*Narrative Writing Scores Of JP*

| | NWT1 | | | NWT2 | | |
|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M1 | M2 | M3 |
| **Lexical** | | | | | | |
| Total words | *136* | 74 | 82 | 112 | 75 | 117 |
| Word length | 3.96 | 4.26 | 4.07 | 4.54 | 4.65 | 4.49 |
| Word frequency of nouns | **5.29** | 4.88 | 5.23 | 4.57 | **4.90** | 4.65 |
| Type token ratio of words | 0.63 | 0.66 | 0.70 | 0.68 | **0.76** | 0.67 |
| Density of adjectives | *58.82* | **54.05** | **36.59** | 62.50 | 53.33 | 51.28 |
| Density of nouns | 183.82 | 216.22 | 182.93 | 160.71 | 200.00 | 188.03 |
| Density of verbs | 176.47 | **202.70** | **219.51** | 214.29 | **253.33** | 196.58 |
| Density of concrete verbs | 22.06 | *0.00* | 12.20 | 8.93 | 13.33 | 8.55 |
| Density of abstract verbs | 117.65 | **148.65** | 109.76 | *142.86* | **173.33** | **145.30** |
| **Grammatical** | | | | | | |
| D-Level | 2.38 | **5.33** | 1.83 | 1.71 | **3.00** | **3.00** |
| Distance subject-verb | 3.08 | 2.73 | 2.21 | **4.17** | 0.73 | 2.37 |
| Distance determiner-noun | 0.12 | 0.08 | 0.18 | 0.36 | *0.00* | 0.21 |
| Density of personal references | 95.59 | 202.70 | 146.34 | 116.07 | **160.00** | **153.85** |
| Density of conjunctions | 44.12 | 81.08 | 48.78 | 17.86 | **40.00** | **42.74** |

*Note*: The scores in bold diverge from the norm as defined in the normative tables; the cursive scores indicate those scores that are on the verge of the norm. M1 = Moment 1; M2 = Moment 2; M3 = Moment 3.



### *2.3.3 Cohesion analysis of the narrative writing task.*

JP differs from the norm (cf.: Table 9), on a variety of measures. Lexical results indicate that he refrains from using more difficult, low frequency words, in the NWT1 M1 and NWT2 M2. The NWT1 shows a clear lack of adjectives compared to the norm. Contrastingly, the density of verbs and abstract verbs was notably higher in NWT1&2, moments two and three. In terms of grammar, the density of personal references is consistently higher than the norm for the NWT2 in all three test moments; the density of conjunctions is consistently too low. D-Level increases in the NWT2, moments 2 and 3.

Even though JP had no known pathology that affected his language skills, his scores diverge from the norm for all three test moments. The emotional impact of the cancer diagnosis was noticeable in the GDS; however, the results of M3 do not stand out.



## 2.4 CASE 4 - BDB

BDB is a retired 69-year-old woman who was tested three times in the longitudinal study. She enrolled in the experiment as a healthy participant. She has had 12 years of education; was self-employed and comes from a higher socioeconomic background, leading to a self-proclaimed decent understanding and usage of English and French. She had had two cerebrovascular accidents; the last accident dated six years ago. Even though the MoCA score of M1 was sufficient, the second and third moment also indicated that some pathology could be at hand.

### *2.4.1 Picture Naming Task*

Analysis of the picture naming task (Table 10) revealed that BDB was not affected by the ease of the images (HR versus LR), naming the relatively 'more complex' LR images with greater ease than the supposedly 'easier' HR images. Nevertheless, the results were high throughout and in line with the findings from Paesen et al. (2019), leaving no room for possible pathology markers.

**Table 10**

*BDB - Scores for Name Agreement and Object Recognition*

| | Name Agreement | | | Object Recognition | | |
|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M1 | M2 | M3 |
| Highly Relatable images (%) | 84.21 | 84.21 | 84.21 | 89.47 | 100 | 89.47 |
| Less Relatable images (%) | 93.10 | 86.76 | 89.66 | 100 | 93.10 | 100 |

*Note:* M1 = Moment 1; M2 = Moment 2; M3 = Moment 3

### *2.4.2 Clinical analysis of the narrative writing task*

After analysis by the language pathologist, the NWT1 is for the three test moments considered to be linguistically inconsistent, with grapheme additions (e.g., 'aangengeven' instead of 'aangegeven'), grapheme omissions (e.g., 'honde' and 'maande' in which the letter 'n' at the end is missing; 'nar' ('naar' = 'to'); 'vakatie' ('vakantie' = 'vacation'); etc.), and grapheme substitutions (e.g., 'ben' instead of 'bed').

The NWT2 shows similar, inconsistent, language during the three test moments. No interpunction was used, determiners and nouns were written without a space in between (e.g., 'eenlandbouwer'), graphemes were both omitted and added (e.g., 'meen' instead of 'mee'); diphthongs were substituted (e.g., 'zielde' instead of 'zeilde' (= 'sailed') and spelling mistakes were made (e.g., 'vertelen' instead of 'vertellen').

### *2.4.3 Cohesion analysis of the narrative writing task.*

BDB did differ from the norm on numerous occasions (cf. Table 11), albeit never in a consistent manner. Results of NWT1, on M1, revealed an increase in type token ration (word variety) and the lack of adjectives in her narrative. The second time she wrote the NWT1, these variables were according to the norm again; she now showed signs of increased distance between subject - verb and determiner - noun. The third trial, no variables diverted from the norm. The NWT2 showed, in accordance with the previous two cases, a difference in density of abstract verbs; being lower than the norm in the second trial and higher than the norm in the third trial. Furthermore, the D-Level (measure of sentence complexity) increased in the final trial.



**Table 11**

*Narrative Writing Scores Of BDB On The First And Second Test Moment*

| | NWT1 | | | NWT2 | | |
|---|---|---|---|---|---|---|
| | M1 | M2 | M3 | M1 | M2 | M3 |
| **Lexical** | | | | | | |
| Total words | 63 | 94 | 113 | 53 | 73 | 92 |
| Word length | 4.38 | **4.81** | 4.35 | 4.43 | 4.33 | 4.62 |
| Word frequency of nouns | 4.61 | 4.96 | 4.96 | 4.15 | 4.39 | 4.19 |
| Type token ratio of words | **0.86** | 0.67 | 0.73 | 0.74 | 0.74 | 0.68 |
| Density of adjectives | **0.00** | 53.19 | 88.50 | *18.87* | 95.89 | 43.48 |
| Density of nouns | 253.97 | 255.32 | 150.44 | 226.42 | 150.69 | 141.30 |
| Density of verbs | 206.35 | 180.85 | 203.54 | 226.42 | 178.08 | 260.87 |
| Density of concrete verbs | *0.00* | 31.91 | *0.00* | *0.00* | 27.40 | *0.00* |
| Density of abstract verbs | 142.86 | 127.66 | 141.59 | 188.68 | **54.79** | **206.52** |
| **Grammatical** | | | | | | |
| D-Level | 2.00 | 3.75 | 3.00 | 1.60 | 2.50 | **5.50** |
| Distance subject-verb | 3.73 | **6.25** | 2.88 | 2.67 | 1.29 | 1.89 |
| Distance determiner-noun | 0.00 | **0.74** | 0.36 | 0.10 | *0.55* | 0.15 |
| Density of personal references | 111.11 | 53.19 | 159.29 | 113.21 | 178.08 | 184.78 |
| Density of conjunctions | 79.37 | 85.11 | 70.80 | *18.87* | 41.10 | 54.35 |

*Note*: The scores in bold diverge from the norm as defined in the normative tables; the cursive scores indicate those scores that are on the verge of the norm. M1 = Moment 1, M2 = Moment 2, M3 = Moment 3

BDB displayed various differences with the norm, even though the results are inconclusive. Given that a person who has had multiple strokes will suffer from fluctuating mental fatigue, they might exhibit inconclusiveness in testing and performance, which allows for this result makes sense. Nevertheless, in accordance with the previous cases, the density of abstract verbs was out of line, albeit it with a lower score for the second test moment and a higher score for the final test moment.

## 4. DISCUSSION

The aim of this study was to validate sex and age related data of written narratives, based on four cases with probable pathologies. With the use of two picture based narrative writing tasks we measured both lexical and grammatical characteristics of spontaneous language generation. The four cases were compared to a normative set as described in Paesen et al. (submitted). Via quantitative data-analysis of text characteristics, we answered the following research question: In terms of spontaneous language generation, which parameters are needed for a differential diagnosis between healthy and pathological?

The lexical characteristics, density of abstract verbs and density of concrete verbs, lead to remarkable findings in the NWT2; all cases diverge from the norm on the former, while the latter remains unaffected. In previous studies, researchers found that people who suffer from a depression show an increased use of concrete adjectives and words (Al-Mosaiwi & Johnstone, 2018). Consequently, the data of case 1 is in line with this finding, showing a lower density of abstract verbs; nevertheless, the density of concrete verbs was not affected. Interestingly, case 3 has a consistently higher level in density of abstract verbs; whereby the effect of the cancer diagnosis does not affect the density levels. We also hypothesized that aphasic language could be differentiated based on lexical markers. In a study by Bonner et al. (2010), a significantly greater impairment was found for concrete verbs, not for abstract verbs, in patients who suffer from semantic dementia. They contribute this issue to the comprehension difficulty for concrete verbs and link this to the degradation of the visual-perceptual feature knowledge, as represented in the anterolateral temporal lobe. The results of case 1, however, show no significant difference with the control group in terms of the density of concrete verbs. The density of abstract verbs did increase compared to the healthy controls, especially in the NWT2. Concerning the other lexical markers, it can be noted that our cases without a clear probable pathology (3, 4) write with great lexical variability over the various test moments; however, no conclusive results can be drawn from them.

In terms of grammar, we hypothesized that the density of personal pronouns could be a suitable biomarker for a depressive syndrome. However, the results indicate that there are no differences in the density of personal references, not in NWT1 – which triggers the first person pronoun – nor in the more challenging NWT2 – which elicits the use of the third person pronoun. This result stands in contrast to previous studies on depression which have found clear indications of an



elevated use of first person pronouns and a decrease in use of the third person pronoun. As the given tasks trigger those types of pronouns respectively, a discrepancy with the healthy control group was to be expected.

Even though the discrepancies with the healthy control group are rather slim, out-of-the-ordinary behaviour in NWT2 can be noticed, the task that elicited a narrative based on six images and required the use of the third person pronoun and past tense, for density of abstract verbs throughout the various case profiles. Therefore, we believe that additional research into the density of abstract verbs will shed more light on a potential differential diagnosis and its applicability as a biomarker for language issues in adults. Even though other language biomarkers already exist, such as the shorter segments of speech and responses phrased as questions that are associated with a cognitive impairment (Alhanai, 2017), the density of abstract verbs is studied through spontaneous generated written language and could provide insight into multiple pathologies. To assure the merits of spontaneous written language generation compared to the existing screening tools and methods, we replicated the more traditional picture naming studies. However, undifferentiating results over the various cases and probable pathologies indicate that the PNT could be to concise to make a differential diagnosis. Therefore, we conclude that research into spontaneous language generation is necessary to shed more light on the issue.



With regard to the future, we would suggest the study of writing behaviour and more specifically pausing times. Temporal aspects of writing (such as pausing behaviour) are clearly sensitive to cognitive impairment. If pause times required before the writing of the abstract verbs were to be studied, possible insight into the cognitive processes could be gathered. Comparing the length of the pause times to the pause times before concrete verbs could reveal the ease with which both verb types are recalled. Moreover, because this result could only be replicated in this second, more complex task, we would like to suggest shortening the total task, for instance leaving out the first narrative writing task and the picture naming task. Participants needed on average one hour to one hour-and-a-half to complete the entire experiment, which had a great impact on their concentration. Additionally, we realise it is not recommended to use the digit span task to measure verbal working memory on persons with Aphasia. Commonly, they suffer from word repetition difficulties, leading to possible mirroring of their Aphasia rather than their working memory capacity (Wright & Fergadiotis, 2012). Given that the original design of this experiment targeted healthy controls, this was not an issue. Upon replicating this study in a clinical setting, we would suggest to adapt this part of the test to fit the target group.

To conclude; studying spontaneous language generation provides a basis for establishing a differential diagnosis. The newly developed tool is versatile in nature, low-cost, provides a non-academic applicability and is non-invasive. Even though further research into the field is needed, our results indicate that the density of abstract verbs could serve as a biomarker for language disturbances in adults.

## 5. ACKNOWLEDGEMENT

This research was supported by the research fund BOF-DOCPRO 2015, PS ID: 31661 of the University of Antwerp. We would also like to thank Eric Van Horenbeeck, the technical coordinator of Inputlog, and Luuk Van Waes, co-founder of Inputlog and the typing task. We thank Johan Frid and Victoria Johansson, respectively the technical coordinator and co-founder of Scriptlog for creating a Scriptlog module especially for our research purposes. Further thanks go to Peter Mariën (†) and Dorien Vandenborre for their input during the design of the study. Last but not least we would like to thank the elderly who enthusiastically participated in this study.

## 6. DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.



## 7. BIBLIOGRAPHY

Agosta, F., Henry, R. G., Migliaccio, R., Neuhaus, J., Miller, B. L., Dronkers, N. F., Brambati, S. M., Filippi, M., Ogar, J. M., Wilson, S. M., & Gorno-Tempini, M. L. (2010). Language networks in semantic dementia. *Brain*, 133(1), 286–299. https://doi.org/10.1093/brain/awp233

Alhanai, T., Au, R., & Glass, J. (2017). Spoken language biomarkers for detecting cognitive impairment. *2017 IEEE Automatic Speech Recognition and Understanding Workshop* (ASRU), 409–416. https://doi.org/10.1109/ASRU.2017.8268965

Al-Mosaiwi, M., & Johnstone, T. (2018). In an Absolute State: Elevated Use of Absolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation. *Clinical Psychological Science*, 6(4), 529–542. https://doi.org/10.1177/2167702617747074

Alexopoulos, G. S. (2019). Mechanisms and treatment of late-life depression. *Translational Psychiatry,* 9(1). https://doi.org/10.1038/s41398-019-0514-6

Ashaie, S. A., Hurwitz, R., & Cherney, L. R. (2019). Depression and Subthreshold Depression in Stroke-Related Aphasia. *Archives of Physical Medicine and Rehabilitation*, 100(7), 1294–1299. https://doi.org/10.1016/j.apmr.2019.01.024

Banerjee, D. (2020). The impact of Covid-19 pandemic on elderly mental health. International *Journal of Geriatric Psychiatry,* 35(12), 1466–1467. https://doi.org/10.1002/gps.5320

Bridges, K. A., & Van Lancker Sidtis, D. (2013). Formulaic language in Alzheimer's disease. *Aphasiology*, 27(7), 799–810. https://doi.org/10.1080/02687038.2012.757760

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, 150, 80–84. https://doi.org/10.1016/j.actpsy.2014.04.010

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Christman, S., Bermudez, C., Hao, L., Landman, B. A., Boyd, B., Albert, K., Woodward, N., Shokouhi, S., Vega, J., Andrews, P., & Taylor, W. D. (2020). Accelerated brain aging predicts impaired cognitive performance and greater disability in geriatric but not midlife adult depression. *Translational Psychiatry*, 10(1). https://doi.org/10.1038/s41398-020-01004-z

Coco, D. Lo, Lopez, G., & Corrao, S. (2016). Cognitive impairment and stroke in elderly patients. *Vascular Health and Risk Management*, 12, 105–116. https://doi.org/10.2147/VHRM.S75306

Copland, D. A., & Angwin, A. J. (2019). Subcortical Contributions to Language. In G. I. de Zubicaray & N. O. Schiller (Eds.), *The Oxford Handbook of Neurolinguistics* (pp. 850–876). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190672027.013.33

Ellis, C., & Urban, S. (2016). Age and aphasia: a review of presence, type, recovery and clinical outcomes. *Topics in Stroke Rehabilitation*, 23(6), 430–439. https://doi.org/10.1080/10749357.2016.1150412

Fleck, J. I., Kuti, J., Mercurio, J., Mullen, S., Austin, K., & Pereira, O. (2017). The Impact of Age and Cognitive Reserve on Resting-State Brain Connectivity. *Frontiers in Aging Neuroscience,* 9(December), 1–13. https://doi.org/10.3389/fnagi.2017.00392

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research,* 12(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Frid, J., Johansson, V., Johansson, R., Wengelin, Å., & Johansson, M. (2014). Developing a keystroke logging program into a writing experiment environment. *Presentation at Writing Across Borders,* WRAB, Paris February 2014.

Gilhooly, K. J., & Logie, R. H. (1980). METHODS &DESIGNS Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation,* 12(4), 395–427. https://doi.org/10.3758/BF03201693

Ginex, V., Gilardone, G., Viganò, M., Monti, A., Judica, E., Passaro, I., Gilardone, M., Vanacore, N., & Corbo, M. (2020). Interaction Between Recovery of Motor and Language Abilities After Stroke. *Archives of Physical Medicine and Rehabilitation*, 101(8), 1367–1376. https://doi.org/10.1016/j.apmr.2020.04.010

Glosser, G., & Kaplan, E. (1989). Linguistic and nonlinguistic impairments in writing: A comparison of patients with focal and multifocal CNS disorders. *Brain and Language*, 37(3), 357–380. https://doi.org/10.1016/0093-934X(89)90025-4

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., Johnson, J. K., Weiner, M. W., & Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, 55(3), 335–346. https://doi.org/10.1002/ana.10825

Hakim, A. M. (2011). Depression, strokes and dementia: New biological insights into an unfortunate pathway. *Cardiovascular Psychiatry and Neurology,* 2011(Mci). https://doi.org/10.1155/2011/649629

Kemper, S., Greiner, L., Marquis, J., Prenovost, K., & Mitzner, T. (2001). Language decline across the life span: Findings from the nun study. *Psychology and Aging*, 16(2), 227–239. https://doi.org/10.1037/0882-7974.16.2.227

Kircanski, K., Joormann, J., & Gotlib, I. H. (2012). Cognitive Aspects of Depression. *Wiley Interdiscip Rev Cogn Sci,* 3(3), 301–313.

Koen, J. D., & Rugg, M. D. (2019). Neural Dedifferentiation in the Aging Brain. *In Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2019.04.012



Kremin, H., Perrier, D., Wilde, M. De, Dordain, M., Bayon, A. Le, Gatignol, P., Rabine, C., Corbineau, M., Lehoux, E., & Arabia, C. (2001). Factors predicting success in picture naming in Alzheimer's disease and primary progressive aphasia. *Brain and Cognition*, 46(1–2), 180–183. https://doi.org/10.1006/brcg.2000.1270

Laws, K. R., Irvine, K., & Gale, T. M. (2018). Sex differences in Alzheimer ' s disease. *Current Opinion in Psychiatry,* 31(2), 133–139. https://doi.org/10.1097/YCO.0000000000000401

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing*, 26(4), 435–461. https://doi.org/10.1093/llc/fqr013

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication,* 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Lopez, O. L., Jagust, W. J., Dulberg, C., Becker, J. T., DeKosky, S. T., Fitzpatrick, A., Breitner, J., Lyketsos, C., Jones, B., Kawas, C., Carlson, M., & Kuller, L. H. (2003). Risk Factors for Mild Cognitive Impairment in the Cardiovascular Health Study Cognition Study. *Archives of Neurology,* 60(10), 1394. https://doi.org/10.1001/archneur.60.10.1394

Meunier, D., Stamatakis, E. A., & Tyler, L. K. (2014). Age-related functional reorganization , structural changes , and preserved cognition. *Neurobiology of Aging,* 35, 42–54.

Mukhtar, S. (2020). Psychological impact of COVID-19 on older adults. *Current Medicine Research and Practice, 10, 201–202.*

Murphy, E., Hoshi, K., & Benítez-Burraco, A. (2021). *Subcortical syntax: Reconsidering the neural dynamics of language*. https://doi.org/https://doi.org/10.31234/osf.io/29cjw

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment , MoCA : A Brief Screening. *Journal of the American Geriatrics Society*, 53(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Paesen, L., & Leijten, M. (2019). Name agreement and naming latencies for typed picture naming in aging adults. *Clinical Linguistics & Phonetics,* 33(10–11), 930–948. https://doi.org/10.1080/02699206.2019.1590734

Paesen, L., Leijten, M., & De Letter, M. (n.d.). Aging and sex differences in the cohesion of written narratives. *Submitted.*

Pander Maat, H., Kraf, R., Bosch, A. van den, Dekker, N., Gompel, M. van, Kleijn, S., Sanders, T. J. M., & Sloot, K. van der. (2014). T-Scan : a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.

Patterson, K., & Hodges, J. R. (2001). Dementia , Semantic. I*nternational Encyclopedia of the Social Behavioral Sciences*, 3401–3405.

Perna, L., Wahl, H. W., Weberpals, J., Jansen, L., Mons, U., Schöttker, B., & Brenner, H. (2019). Incident depression and mortality among people with different types of dementia: results from a longitudinal cohort study. *Social Psychiatry and Psychiatric Epidemiology*, 54(7), 793–801. https://doi.org/10.1007/s00127-019-01683-0

Pyszczynski, T., & Greenberg, J. (1987). Self-regulatory perseveration and the depressive self-focusing style: A self-awareness theory of reactive depression. *Psychological Bulletin*, 102(1), 122–138.

Rodríguez-Ferreiro, J., Davies, R., González-Nosti, M., Barbón, A., & Cuetos, F. (2009). Name agreement, frequency and age of acquisition, but not grammatical class, affect object and action naming in Spanish speaking participants with Alzheimer's disease. *Journal of Neurolinguistics,* 22(1), 37–54. https://doi.org/10.1016/j.jneuroling.2008.05.003

Rosenbach, C., & Renneberg, B. (2015). Remembering rejection: Specificity and linguistic styles of autobiographical memories in borderline personality disorder and depression. *Journal of Behavior Therapy and Experimental Psychiatry,* 46, 85–92. https://doi.org/10.1016/j.jbtep.2014.09.002

Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, 18(8), 1121–1133. https://doi.org/10.1080/02699930441000030

Salary, S., & Moghadam, M. A. A. (2013). Relationship Between Depression and Cognitive Disorders in Women Affected with Dementia Disorder. *Procedia - Social and Behavioral Sciences*, 84, 1763–1769. https://doi.org/10.1016/j.sbspro.2013.07.028

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 5(1–2), 165–173. https://doi.org/10.1300/J018v05n01_09

Tølbøll, K. B. (2019). Linguistic features in depression: A meta-analysis. *Journal of Language Works - Sprogvidenskabeligt Studentertidsskrift*, 4(2), 39–59. https://tidsskrift.dk/lwo/article/view/117798

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior*, 73, 311–319. https://doi.org/10.1016/j.chb.2017.03.050

Vijayan, M., & Reddy, P. H. (2016). Stroke, Vascular Dementia, and Alzheimer's Disease: Molecular Links. *Journal of Alzheimer's Disease*, 54(2), 427–443. https://doi.org/10.3233/JAD-160527

Waes, L. Van, Leijten, M., Pauwaert, T., & Horenbeeck, E. Van. (2019). A Multilingual Copy Task : Measuring Typing and Motor Skills in Writing with Inputlog. *Journal of Open Research Software,* 70(30). https://doi.org/https://doi.org/10.5334/jors.234

Wechsler, D. (2008). *Wechsler adult intelligence scale: WAIS IV; technical and interpretive manual* (Pearson (ed.)).



WHO. (2015). *Dementia A public health priority*. http://www.who.int/mental_health/neurology/dementia/en/

WHO. (2017). *Mental health of older adults.* https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults

WHO. (2021). *Ageing and health*. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

Wright, H. H., & Fergadiotis, G. (2012). Conceptualising and measuring working memory and its relationship to aphasia. *Aphasiology*, 26(3–4), 258–278. https://doi.org/10.1080/02687038.2011.604304

# Chapter six

## General discussion

## 1. AIM

This final chapter summarises the main findings of the conducted studies in light of the research questions and hypotheses posed in chapters two to five. Moreover, the implications of our findings for future research will be addressed, together with the limitations and possibilities posed by the applied research methods in this dissertation. Additionally, the applicability of our newly created clinical toolkit for the evaluation of spontaneous written language will be discussed. This dissertation closes with a final conclusion.

Throughout this dissertation, various studies were set up to find out how spontaneous written language changes upon ageing. More specifically, we defined four main aims that needed to be achieved in order to be able to formulate an answer to our main research question: we aimed (1) to establish how age, image characteristics and repeated testing affect naming, (2) to find out how ageing and sex affect the cohesion of written narratives, (3) to measure the test-retest reliability of spontaneous written language in healthy ageing and (4) to explore if spontaneous written language could be used to make a differential diagnosis. In order to be able to reach these goals, the decision was made to construct a new clinical toolkit for the evaluation of spontaneous written language. We established three additional, technical goals that needed to be reached: (5) a toolkit needed to be designed in order to bridge the gap within differential diagnostics with regard to spontaneous written language generation, (6) within that toolkit, we needed congeneric tasks that are interchangeable on certain key characteristics, and (7) the tool kit needed to be standardised within an ageing population and a pilot needed to be run in a clinical setting. The various research questions that contribute to this goal have been examined.



In order to be able to reach the first goals of this dissertation, respectively **(1) establish how age, image characteristics and repeated testing affect naming** and **(6) create congeneric tasks,** we decided to design narrative writing tasks that elicit spontaneous written language on the basis of congeneric single-object images. The usability of these single-object images was verified in chapter two, the creation and standardisation of the narrative writing tasks was discussed in chapter three. The images used in this study were selected based on strict naming scores, so as to assure their congeneric character in test-retest situations. Therefore, with regard to naming accuracy, we hypothesised that neither word / image characteristics (e.g., word frequency) nor age would influence naming agreement and object recognition. Our results partly confirm this statement, with no influence of image characteristics on naming accuracy within the younger age group. Object recognition, i.e.: the naming correctness of images when also taking synonyms into account, was influenced in older age by image characteristics. This finding was in line with results from previous studies that point towards the increased difficulty words with a lower age-of-acquisition and lower frequency for naming accuracy. The influence of age on naming accuracy of less relatable images indicated that older people relay on synonyms and alternative names more often; a result attributed to word retrieval issues and tip-of-the-tongue mistakes in similar studies (Clark-Cotton et al., 2007).

Based on studies by for instance Boukadi and colleagues (2016) and Scaltritti and colleagues (2016) latencies were correctly hypothesised to be influenced by image characteristics, with increasing latencies for less relatable images. Age also affected latencies, with longer reaction times for the older participant group irrespective of typing speed. The test-retest reliability, with an interpose of three months, imply that naming latencies did not significantly differ upon retesting, leading to a stable measure to use in longitudinal research. Learning effects were noticed for the interkey latencies. Similarly, name agreement could be used as a predictor for naming latencies; this was not the case for interkey latencies.

The results from the image validation study had a lot of implications for the continuation of this dissertation. First of all, it was clear that even within healthy ageing, languages changes can be noticed between our different age groups. Therefore, it is crucial not to treat the healthy ageing adults as a homogeneous group, but rather as a set of individuals with multilevel analysis. Furthermore, even though this study was merely exploratory in nature, we believe that naming latencies could potentially provide a strong basis for differential diagnosis. Given that the latencies differed between both image categories and age groups, it implies that there are changes even within healthy ageing that need to be taken into consideration. Finally, based on the naming scores and latencies for the different images, the decision was made to continue this study and remove two images from our dataset.

The creation of the newly designed images allowed us to continue constructing narrative writing tasks, tasks that elicit a narrative based on prompted images. The nouns triggered by these images function as key control variables and contribute to the congeneric character of those tasks. In the third chapter of this dissertation, the narrative writing tasks were constructed and standardised within a healthy ageing population. To that extent, we posited the research question how ageing and sex affect the cohesion of written narratives. Cohesion was studied in



light of both grammatical and lexical product measures. In terms of grammatical complexity, it is important to highlight that we found that age did not negatively impact sentence complexity, a result in line with previous studies in the field. We even found sentence complexity to be increasing upon ageing in our cohort. Notably, the density of personal references did differ between tasks, presumably triggered by the different task demands between NWT1 and NWT2. Lexically, a similar result was found. Even though the NWT2 triggered longer and more words compared to NWT1, we did notice a decrease in word count and length from the age of 70 onwards. One possible explanation can be found in potential tip-of-the-tongue issues that older participants encountered, leading to the temporary disruption of semantic recall, and resulting in descriptions or substitutions of the desired word rather than the word itself (Kavé & Goral, 2018). Nevertheless, this finding did not appear to impact word difficulty and variety, with a decrease in word frequency (and thus the usage of more difficult words) and an increase of the type-token ratio. We do believe that our two task types may have contributed to an increased difference, with lower frequency & unrelated trigger words in NWT2 and a more everyday setting with easier target words in NWT1. Sex differences were found for both grammatical and lexical variables. For instance, older women write considerably shorter texts compared to their younger counterparts, and that for men, word length increases with age.



We concluded that age did not negatively affect our participants' performance. The reason for that, we believe, is twofold. First of all, we believe that our images still trigger well entrenched words. Given that the name agreement needed to be sufficiently high, the elicited words are not too complex in nature, allowing a fluent narrative. Secondly, these results are also supported by findings from other studies; even though neural circuitry becomes wide-spread upon ageing, compensatory mechanisms in the brain will support performance and allow it to be maintained (Alatorre-Cruz et al., 2018). Nevertheless, the effects of these compensatory mechanisms could affect response times. As a result, we believe insight into the writing process could reveal potential issues with the construction of the desired text that may not be apparent in the final product. Additionally, in accordance with previous studies (Aerts et al., 2015), we believe future studies could benefit from distinguishing between male and female participants, given that we found sex to be an influencing factor in our cohesion data.

These findings did not only provide an answer to our aim **(2) to find out how ageing and sex affect cohesion**, they also contribute to finalising the design **(5) of a tool that potentially bridges the gap within differential diagnosis with regard to spontaneous written language generation**.

Moreover, it allowed us to achieve the seventh goal within this dissertation, namely **(7) standardise the toolkit in a healthy population**. Based on this data, we were able to construct a normative base of cohesion measures that was used in the continuation of this dissertation. The fourth chapter of this dissertation used this normative dataset as a baseline to **(3) explore the test-retest reliability of spontaneous written language in healthy ageing**. With this study we wanted to explore to what extent spontaneous written language can be used in a test-retest setting by testing its reliability within a healthy ageing cohort. Based on the findings in the previous study, we decided to focus on both the writing product and process.

The results of this experiment clearly demonstrate the test-retest reliability of our study, due to the intra-personal stability of the writing product. None of our product measures was impacted by repeated testing, and no significant age differences were found. Image characteristics and the difference between NWT1 and NWT2 could be responsible for effects on found on word length, frequency and density. These findings indicate that the design of our template, which allows for a great number of image combinations, allowed us to create congeneric tasks that allow for retesting without receiving the same image set more than once, eliminating a potential learning effect while still requiring the same linguistic ability. Based on our initial results, we also recommend future studies into spontaneous written narrative to take the writing process into consideration. Our study indicates the variability within the writing process, while the typing speed and active writing time showed great intra-personal stability.

With the fifth and final study within this dissertation we wanted **(4) to explore whether or not spontaneous written language could be used to make a differential diagnosis** and **(7) pilot the validity of the toolkit in a clinical setting**. This pilot study focussed on four case studies and led to one clear result: we suggest future studies into narrative writing tasks to focus on the lexical characteristic 'Density of abstract verbs' as a potential language biomarker. It contributes to the current gala of language biomarkers such as 'segments of speech' and we suggest that it could be suitable to track multiple pathologies. Moreover, the comparison of the narrative writing task results and the picture naming task within our participant group reveals that more insight into potential pathologies can be gained when using spontaneous written language. It is advisable for future studies to check the writing process behind those narrative writing tasks.



## 2. METHODOLOGICAL CONSIDERATIONS

### 2.1 IMAGES & THE PICTURE NAMING TASK

The basis for the narrative writing tasks are the images that aim to elicit a narrative, presented in a semi-random combination on top of the screen. Upon starting this dissertation, we decided there was a need for a new database with images. The existing databases would not suffice for the purpose of strictly controlled narrative writing tasks. The databases we found either contained black and white images, dated images (e.g., a rotary phone), images with movement or depicting multiple things (e.g., a dolphin jumping from the water), or images that were not consistent within their category (for instance: animals looking in different directions). We believed a strict selection of images and the way they were depicted would lead to a more controllable output. Based on previous experiments, the relevant words were selected (see chapter two for a full description) and images were created.

We believe this dataset is also a strong addition to the current sets due to the way we controlled how the (non)-living things ought to be depicted. Images needed to be relatively simple line drawings, coloured, contain just one element and no movement, and per category all 'face' the same direction. The full list of images can be found in **Appendix A**. Moreover, these images were combined with the research and images developed by Meulemans C. in order to create the Open linguistic Picture Database (OLPD, www.olpd.eu) (Paesen & Meulemans, 2020). This database aims to support the research community by presenting an ever-evolving picture database. It adds to the current field by providing the opportunity for other researchers to use the images for their research and even add images to the database as long as those images follow the same strict visual rules. By adding to the current set within the OLPD, researchers will be able to use from an always increasing pool of up-to-date images and if it does not suffice for their research, only the missing images need to be created. Therefore, we hope that this database will allow for researchers to support each other and benefit from previous research.

In order to standardise the images and check their usability within narrative writing tasks, we set up a picture naming experiment and continued to check for picture naming throughout the narrative writing experiments. Based on our case study (chapter five), we hypothesized that a PNT could possibly be too concise to make a differential diagnosis between healthy controls and people with a probable pathology. Spontaneous written language triggered by narrative writing tasks would, therefore, offer the opportunity to close this gap. We suggest future studies into narrative writing tasks to analyse the writing process and more specifically the naming latencies of the target nouns. If that noun is repeated multiple times, or even synonyms are given, those processes will provide insight into working memory, sentence planning and word characteristics. Furthermore, this will enhance the link with picture naming tasks while still allowing spontaneous written language. Moreover, when doing this in a longitudinal setting, it could contribute to the sparse longitudinal research into tip-of-the-tongue issues within for instance people with a mild cognitive impairment (Campos-Magdaleno et al., 2020).

## 2.2 DESIGN OF NARRATIVE WRITING TASKS

Narrative writings tasks are newly developed tasks that aim to generate spontaneous written speech, while still controlling for certain key variables. More specifically, participants are given four to six images and are asked to write a story in which they are obliged to incorporate the nouns depicted by the images. Hence, those words can be strictly controlled, while still catering for a spontaneous language flow.

Based on the findings by Forbes-McKay and colleagues (2014), two types of narrative writing tasks were created in order to account for both floor and ceiling effects: NWT1 (a simpler version) and NWT2 (the more complex variant). The images were selected from the OLPD (Paesen & Meulemans, 2020) and divided into two categories, fitting either NWT1 or NWT2 descriptions. The NWT1 consists of a combination of four 'easier' images, the NWT2 of six more 'complex' images, that all met the criteria as described in Table 1 in the Introduction of this dissertation. The instructions between the two tasks also differed, with NWT1 texts that needed to be written in the first person singular and NWT2 texts requiring the use of the third person singular and also the use of the past tense. This discrepancy derives from the issues Alzheimer patients have with narrating about topics that divert from their reality (Hydén & Örulv, 2009).

One of the limitations of the current image set and narrative writing tasks is the replicability in other languages. Because the images were checked and designed specifically with the Dutch language in mind, there is no direct relationship with the results in other languages. The same counts for the narrative writing tasks; all results found need to be cross-referenced with new studies in other languages to assure its consistency. Nevertheless, the design of the task does allow for researchers to easily replace one of the images to another image of their choosing. It is also an advantage of the OLPD is that it allows the entry of new images. Other studies might find



the need to replace one of the images in the PNT / NWT and might, therefore, create a new one, based on the same criteria in the desired language.

## 2.3 TOOLS USED IN THIS DISSERTATION AND THEIR IMPLICATIONS

In contrast to language comprehension, previous research has demonstrated that language production is prone to be affected by age (Burke & Mackay, 1997; Shafto & Tyler, 2014). For that reason, changes in language production are often studied in order to provide insight into possible cognitive disorders and gradual impairments. Despite the large variety of studies into speech, some studies have found that written language is more susceptible to the effects of certain conditions such as Alzheimer's disease. Within the field of written language studies, we posited ourselves with a written test that triggers spontaneous language and that takes both the written product and the writing process into consideration. To that extent, online tools were used to log and analyse the data: ScriptLog, Inputlog, Comproved and T-Scan, an overview can be found in Figure 6.

**Figure 6**

*Overview Of The Different Tools Used In The Registration And Analysis Of The Data*

| | | | |
|---|---|---|---|
| REGISTRATION | inputlog | Typing test | |
| | ScriptLog | Logging NWT + PNT | |
| ANALYSIS | | | |
| 1. Writing process | ScriptLog | Turn data into readable InputLog-file | |
| | inputlog | 1. Pre-process data | |
| | | 2. Data check | General analysis<br>Summary analysis |
| | | 3. Analysis | General analysis<br>Summary analysis<br>Pause analysis<br>Word pause analysis<br>Typing test |
| 2. Writing product | W | 1. Text check<br>2. Correct spelling/typing mistakes | |
| | T-Scan | Cohesion analysis | Linguistics<br>Grammar |
| | comproved | Cohesion analysis | Comparative |

### 2.3.1 Tools to log and analyse the writing process

Two keystroke logging tools played a central role in this dissertation: ScriptLog (Frid et al., 2014) and Inputlog (Leijten & Van Waes, 2013). Both tools were used complementary and were each used for their own purpose: we wanted to (1) establishing a typing baseline for each participant, (2) log and time stamp every movement and keyboard activity, and (3) analyse the data from a writing process perspective. We will elaborate on our methodological considerations in the hopes of allowing future studies to select the right tool and analysis for their study.

#### 2.3.1.1 *Registration - Establishing a typing speed baseline*

In writing research and in studies that aim to gain insight into cognitive and linguistic changes, it is crucial to establish a baseline typing speed for the patient. Therefore, we would advise conducting a typing task, for which purpose we selected the keystroke logging tool Inputlog. This tool enables measuring someone's personal typing speed with the module 'Copy Task'. In order to establish one measure that provides insight in the typing capacity, it was essential that we selected the relevant measures. In consultation with the creators of the copy task and after a correlation analysis, we decided to opt for 'II__TB_LogMean__trimmed_', which provides insight into the log mean of 'overall synthesis Interkey Interval analysis of targeted bigrams'. This score was calculated for every participant individually and was always used in further analysis. Given that our results indicate that the typing speed remained consistent in our longitudinal setting, we conclude that it is a stable intra-personal variable and that could shed additional light on cognitive processes if it changed over a period of time. When used in a clinical setting, and with automatization of the test in mind, the relevant variable can be extracted from the analysis rather easily. However, currently, this would still need to be done by the researcher. Additional programming would be needed to extract this variable automatically.

#### 2.3.1.1 *Registration - Log and time stamp every movement and keyboard activity*

ScriptLog, developed at the University of Lund, is the keystroke logging tool that was used for the logging and time stamping for most of our data gatherings. Whereas Inputlog is the ideal keystroke logging tool to gather data in a seemingly natural writing environment, ScriptLog is designed for more experimental settings and allows the use of stimuli. When we started out with this dissertation, we realised that our type of study required a very specific experimental environment. Hence, the decision was made to work with ScriptLog and a new module within the



tool was designed for our study. With this new module, our cooperation with ScriptLog contributes to future research, as this module allows for picture naming tasks and narrative writing task or variants to be logged.

#### 2.3.1.1 *Analysis - Analyse the data from a writing process perspective*

Analyses were run with Inputlog, mainly due to the multitude of analyses and data preparation options that the tool has to offer. Several steps were needed consecutively to receive the desired output. Based on the requirements needed for this dissertation, we would suggest taking the following actions - pre-processing, data-checking, analysing and post-processing activities – for future studies into narrative writing:

**Pre-processing** - The first step in data preparation is to transform the ScriptLog files into IDFX-files, as those files are readable within Inputlog. Secondly, the clutter needs to be removed at the beginning and ending of a text. Even though the instructions were read out before the beginning of the logging of each task, due to unforeseen circumstances, a participant might still ask a question or forget to press the 'next' button. This leads to unnecessary clutter at the beginning or ending of a text that is not task related. Inputlog offers an automatized option in their 'pre-processing' module that removes clutter at both ends of the process. However, this could potentially also remove the time that the participant was rereading their text. Therefore, the clutter only needed to be removed manually with Inputlog, in exceptional circumstances and needs to be checked on an individual basis. A disadvantage is that this process requires a lot of attention and manual input from the researcher, making the automatization in a clinical setting nearly impossible.



**Data-checking** – For this dissertation, various analyses were selected to gain insight into our data. When replicating our studies, first and foremost, a general analysis should be run in order to check the different log files for possible technical issues or unforeseen circumstances. The general analysis provides a linear, event-based representation of the actions in each logging file. However, this process can be time-consuming, especially if it needs to be done for a large dataset. Therefore, it was always crucial for the researcher to pay close attention to the writing process in order to already be able to foresee such issues. As mentioned before, this clutter can be cut out. After the general analysis, the researcher can locate the issue, and use the pre-processing module to remove the issue from the file at the right location. An additional check of the logging data was run with the use of the 'Summary' analysis. This analysis provides several writing process variables that can be used to detect outliers in the process. More specifically, we mainly focussed on the 'total process time' and 'total pause time' to assure no errors in the data could be found.

**Analyses** - After the pre-processing and final data-checking, we would suggest to carefully select the right analysis in order to obtain the desired information. This dissertation used a combination of analyses, mainly focussing on the: 'summary', 'pause' and 'word pause' analysis. In future studies, we would suggest expanding the knowledge and use of the word pause analysis. For the latest studies within this dissertation, the word-pause analysis 'as was' contained some errors, leading to a loss of 50% of our data. As a result, we were forced to neglect this analysis, however promising the data could be for this type of research.

**Post-processing** - After analysing the data, the researcher obtains one file per type of analysis for each participant, task and – if applicable – test moment, containing all variables that belong to that analysis. In order to obtain a general file that combines all data, we would suggest the use of the post-process module 'Merge', which allows the researcher to combine the different analysis files into one CSV file, which allows for further analysis in statistical programs such as SPSS or RStudio. We do suggest studying the obtained CSV file beforehand, in order to select the relevant variables and possibly make some preliminary calculations with the use of pivot tables in Excel.



### *2.3.2 Tools to analyse the written product*

**T-SCAN** – T-Scan is an automated text analysis tool for Dutch texts (Pander Maat et al., 2014). It provides insight into text complexity, word use and sentence structure. In order to accurately assess the complexity of texts, the texts need to be free from errors. Therefore, each text was manually cleaned: spelling mistakes were corrected, capitalisation was checked and in certain circumstances interpunction was added (for instance if an umlaut had been forgotten or if the participant had added an 'enter' instead of a period at the end of a sentence). Afterwards, the texts were uploaded into T-Scan, analysis was run and the relevant variables were selected in the CSV output file.

**Comproved** – The comparing tool Comproved aims to provide a quick, reliable and more valid tool for comparing texts. The tool aims at supporting teachers, students and assessors by providing a trustworthy order of texts and giving them a score. For this dissertation, the tool was used to rate the narratives and compare them inter- and intra-individually. In order to analyse the texts with Comproved, the texts were cleaned. Spelling errors were corrected and capitalisation was checked, so as to provide the raters with a text that could not bias them based on spelling errors. We used a total of 44 raters; each rater was given a set of 60 comparisons. In total, 4918 comparisons were made, leading to a high inter-rater reliability score of 0.84. When a comparison within Comproved is made, each text is awarded a score, in our case a 'Cohesion' score, whereby zero is intermediate, everything above zero is more cohesive and everything below zero is less cohesive.



Even though both T-Scan and Comproved brought necessary additional insights into the texts and in our comparison of the changes in cohesion upon ageing, we do believe it will be difficult to use either tool in their current state within a clinical environment. For both analyses, the texts need to be manually checked, corrected and uploaded. T-Scan can be used to analyse both a single file and multiple documents; Comproved requires a certain number of texts and raters. Therefore, the toolkit would always require additional manual and immediate analysis by the researcher, making large-scale use of both tools in a setting without supervision impossible. Nevertheless, both Comproved and T-Scan will continue to provide necessary insights for further validation of the toolkit within an experimental setting, as that validation would entail large-scale testing.

## 3. THE FUTURE OF NARRATIVE WRITING TASKS

### 3.1 IMPROVEMENT OF THE TASK FOR CLINICAL USE

Even though the current dissertation and its research projects resulted in narrative writing tasks that can be used to detect language changes in healthy ageing and could potentially screen for several pathologies, the toolkit in itself is not yet quite functioning outside of the research context. Within this chapter, recommendations will be made to successfully turn this into a functioning screening test that can be used in everyday practice outside of the research context.

First of all, it is of vital importance to **reduce the task length** in order to assure that the tool is not too mentally constraining. The current experiment takes more than 90 minutes in its totality, including the NWTs, picture naming task, typing test, questionnaire, Montreal Cognitive Assessment and digit span tasks. Our results from chapter four clearly demonstrate the stability of the task in longitudinal settings and the similarities between the two narrative writing tasks. Even though the study by Forbes-McKay and colleagues (2014) suggested the need for two tasks in a clinical population, covering for both floor and ceiling effects, we believe a single narrative writing task will generate the same results. The main reason is that, in contrast to Forbes-McKay and colleagues (2014), our tool caters for cognitively healthy adults or adults with subjective cognitive complains (SCC). Reducing task length would, therefore, have a positive impact on cognitive load and overall fatigue after doing the test. Based on our findings, and more specifically the limited reliability of writing process variables in the NWT1 in test-retest situations, we would suggest the use of the second narrative writing task, requiring the participant to come up with a story based on six images. Retesting is necessary to try and replicate these results when using only a single narrative writing task and to validate the reliability of the data.

Additionally, more research is needed to study the effects of ageing on the **writing process** triggered by narrative writing tasks. Our second chapter provided evidence for the strength of naming latencies (referred to as 'Before word pauses' in Inputlog), proving to be a stable measure in possibly differentiating healthy adults and adults with impaired cognitive skills. The nature of the changes in interkey latencies (or 'Within word pauses' in Inputlog) and the various writing process variables discussed in chapter four must be thoroughly studied in order to gain additional insight in potential linguistic challenges a person faces when writing. One method to do so, would be by comparing the naming/interkey latencies from the picture naming task with the naming/interkey latencies of the images used in the narrative writing tasks, and comparing



these latencies when multiple references to that same image can be found within the narrative. In order to turn this task into a fully functioning screening tool, a normative base of the relevant writing process variables is required.

Thirdly, the **functionality** of the screening tool needs to be improved. Currently, a researcher is required to supervise and guide the entire experiment due to the different tools and interfaces that make up our toolkit. When we started out with this project, we wanted to lay the foundation of a tool that people could access from their homes, by logging in on their computer and opening our screening tool, without needing to install additional programs (e.g., Inputlog & ScriptLog) and minimise / eliminate the guidance from a researcher. When they use this tool, it would simply serve as a first indicator on whether or not that person would need to see their GP, inquiring after additional testing into a certain pathology. To that extent, a user-friendly interface is needed that includes the various subparts of this toolkit (e.g., the typing task, NWT2, questionnaire), eliminating the need for a supervisor. A future project should enable linguists, psycholinguists, speech therapists, writing researchers and tool developers to come together and turn the foundation laid by this dissertation into the development of a fully functioning test that can be used outside the academic world too.



Note, however, that we cannot call this tool a '**test**' yet. In order to be allowed to do so, the 'test' needs to meet certain criteria. First of all, a test needs to be efficient. It needs to be constructed in such a way that it triggers the necessary information in the most optimal way possible. Secondly, the test needs to be standardised; multiple people should have been subjected to the 'test'. A third criteria is the required construction of a norm, based on the test. In other words, a large group of participants should have done the test and should be able to be ranked based on their score compared to others. The fourth demand is the need for objectivity. No matter who conducts the research, the obtained results should be the same. Reliability is the fifth demand, and this can be achieved through verifying the test-retest stability, without creating a learning effect. Last but not least, the test should be validated. With 'validated', a whole range of additional targets need to be met: the aim of the test should be clear, with clear tasks that are representative of the domains that need to be checked, the construct (e.g., intelligence, reading skills) should be clearly defined, it should be clear which predictions can be made on the basis of the findings, the test should correlate with a test that already exists and the test should actually test what it set out to test. In order for us to turn this toolkit into a 'test', the validity of the toolkit needs to be ameliorated. Further research is needed to clearly define which predictions can be made based on certain scores. Currently, we know that the 'density of abstract' verbs could point towards inconsistencies in linguistic skills compared to the control group. However, how and the extent to which someone has to deviate from the norm in order to make a certain prediction has not yet been defined. Additionally, a manual should also be constructed that aids the understanding of other researchers into the interpretation of the results.

Furthermore, within this dissertation we focussed on **ages** 50 and onwards. The main reason for this choice was the cognitive changes that happen from this age onwards. Research by Alenius et al. (2019) indicates that while a decrease in linguistic skill is noticeable by advanced age, differences started to appear from the age of 50 onwards. However, in order to gain a general view of the population as a whole, additional testing within the age groups 30-39 and 40-49 would benefit this tool even more. That way, once our toolkit becomes a fully functioning test, it could serve as a diagnostic aid than can be administered repeatedly upon ageing, allowing the smallest of changes in the languages to be noticed from the onset onwards, and, therefore, be able to notify the participant of a possible pathology from the onset.

Last but not least, a **study in the field** would be needed to test this 'improved' narrative writing screening tool, to see whether or not it could be functioning on its own and alliances with GP's would need to be constructed to establish a trustworthy basis for the test.



### 2.2 LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The current dissertation has several limitations. The first and main issue relates to the **socioeconomic background** that might have an impact on someone's narrative skills, reasoning and general wellbeing. It is impossible to control for all factors that might have had an influence of somebody's health and current state. For instance, for the studies presented in the current research, we deliberately opted not to select participants based on their educational level. Participants from all walks of life were included, independent of the amount of schooling they had had. Nevertheless, ample studies indicate that education plays a crucial role in protecting the total brain volume and thus reserve, preventing the onset of dementia (Anderson et al., 2020; Wada et al., 2018). We believe that a normative dataset with subdivisions based on educational background can and should be made in future studies. Similarly, research has shown that a patient's fitness levels impacts their cognitive changes. Language has been found to decrease more in people with a lower general physique (Segaert et al., 2018). The impact of stress cannot be underestimated either. Some participants in the longitudinal study clearly stated to be under a lot of stress, resulting in remarkably lower scores for their MoCA or digit span tasks at that test moment compared to their other moments. Several studies have shown that chronic stress leads to speeded cognitive decline and a greater vulnerability to a cognitive disorder (Aggarwal et al., 2014; Scott et al., 2015; R. S. Wilson et al., 2005; Robert S. Wilson et al., 2007). We acknowledge that the socioeconomic status cannot be underestimated. Within this current dissertation, we have taken the most relevant intra-personal variables into consideration; future studies with larger study population could benefit from using additional variables such as education into consideration.

A second limitation to the current dissertation is the use of **computerised** tests and required typing skills within an elderly population. Within the current group of elderly, most people did not have to use a computer in daily life. Our participants indicated that they learned how to type, sometimes at a young age, on a typewriter, or that they started to use a computer only later on in life for leisure activities. Computer literacy and typing, therefore, do not necessarily come naturally to those people. However, this is only a temporary limitation. In the coming years, computer literacy amongst elderly will grow, with people retiring who are used to work with a computer on a daily basis. This will not only benefit research into writing processes and cognitive decline, also other clinical studies have already proven that computerised training could aid healthy elderly in everyday functioning and emotional wellbeing (Gates et al., 2019; Gordon et al., 2013; Rebok et al., 2014).

Thirdly, due to the strict control of the selected images, our studies might **not be easily replicated** in other countries. Given that the images were checked for Dutch naming agreement, frequency, naming speed, and ease of spelling, we cannot ascertain that those results can be replicated for other languages. New and additional research is needed to guarantee that the selected images still have the right requirements in other languages and new images might need to be created. Those images could supplement the OLPD, as we want to accommodate for both new images in Dutch and other languages, contributing to the research community. In consequence, the narrative writing tasks will also have to be standardised anew, and the normative dataset replicated.

## 2.3 CLINICAL IMPLICATIONS

When we set out with this dissertation, we sought to design a new clinical tool for the evaluation of spontaneous written language that was non-intrusive, cost and time effective. Our computerised screening tool aims to distinguish healthy ageing from ageing with a beginning pathology, based on linguistic biomarkers. Even though this research was mainly theoretical in nature, the developed tool and our findings do have important clinical implications.

**Spontaneous though controlled writing** - A first and important implication of this study is the creation of the narrative writing tasks that allow for spontaneous written language generation while still controlling for certain target nouns. We thereby add to the current battery of test for acquired neurogenic disorders; to our knowledge, this is one of the few studies that was able to generate free written speech while still controlling predetermined linguistic variables. The advantage of free speech is that it is less affected by test/retesting effects, which allows for gathering multiple language sample over the course of a certain period (Rofes et al., 2018). Additionally, it replicates the natural speech of a person, diminishing the experimental context of the test. The target nouns allow this study to still control and automate language analyse. Moreover, previous studies into AD and HC, indicate the need to check for idea generation, sentence structure, empty words (Forbes-McKay et al., 2014; Forbes-McKay & Venneri, 2005; Hier et al., 1985), which is possible with our test. The fact that spontaneous speech can be logged allows for even further insight into potential issues encountered during language generation.

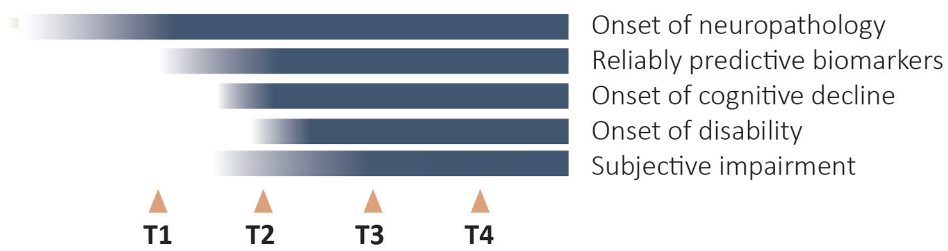


**Early screening of cognitive decline** - In 2020 more than a billion people were aged 60 or over, a number that is expected to double by 2050. Social and economical advancements, together with progress in healthcare and technology, enable a greater functional ability and thus a greater well-being for elderly (WHO, 2020). Nevertheless, an ageing population is also more prone to certain disorders, with neurological disorders being the main cause of disability and second leading cause of death (Avan & Hachinski, 2021; Feigin et al., 2019). Stroke accounts for 69.8% of neurological deaths and dementia 17.3% (Avan & Hachinski, 2021). An early diagnosis is crucial as it allows for a timely treatment, and allows a slowing of the onset of cognitive decline and the subjective impairment (Duboisa et al., 2015; Francis & Pandian, 2021; Li & Liu, 2018). Given that our tasks can be administered repeatedly throughout someone's life, we would be able to monitor the slightest of changes in someone's language. As a result, repeated testing with our NWTs for AD would allow for an early screening and referral to the GP for diagnosis, around T1 and T2 as exemplified by Figure 2, at the onset of cognitive decline and subjective impairment, but before the onset of the disability.

**Figure 2**

*Visualisation Of AD Progression And Diagnosis Points (Adapted From Prince Et Al. (2011)).*






**T1** Earliest possible diagnosis in the event that we develop reliably predictive biomarkers
**T2** Earliest possible diagnosis using currently available technology
**T3** 'Timely' diagnosis responding to patients and caregivers concerns rather than proactively screening for the disease
**T4** Current 'late-stage' diagnosis

**Linguistic biomarker** - The results from this dissertation also imply that the linguistic biomarker 'Density of abstract verbs' is crucial for future studies, as it is the one variable in our case studies that continuously diverged from the norm. This finding is in line with previous studies, who established that a 'concreteness effect' could be found for both healthy controls and people with an acquired language impairment. Abstract words are affected more by brain damage, which coincides with slower recognition times, and a weaker connectivity pattern was found for processing abstract nouns (Fahimi Hnazaee, 2020; Fahimi Hnazaee et al., 2020). More research within a clinical setting is needed to establish what the relationship is between various pathologies and the absence or abundance of abstract verbs compared to other word types in a text. The promise of 'density of abstract verbs' as a linguistic biomarker could also contribute to other studies outside the field of narrative writing tasks.

**Computerised test** - Additionally, the use of this computerised test and the finding of the digital linguistic biomarkers is advantages to the field. Provided the changes suggested in 'Improvement of the task for clinical use' are made, the test proves to be non-intrusive, cost and time effective. By having a computer interface, the test can be administered in people's home. This allows them to generate a spontaneous narrative in their natural habitat, which gives the test a non-intrusive character. The computerised processes allow the analysis to be automated, eliminating manual analysis and possible human error thereupon, and speed the generation of the output. Furthermore, results from our computerised screening tool prove that writing processes provide additional insight into spontaneous language generation, and could shed light on the cognitive status of a patient (Van Waes et al., 2017).

**Individuality** - Another important finding of our research for future studies is that healthy adults cannot be placed within one cohort; differentiations between ages and sexes must be carefully made. For the purpose of these studies, we differentiated between sex and placed age on a continuum in our analyses. Linguistic differences between men and women were found, resulting in a differentiation between the two in our normative database. We suggest future studies into writing and cognitive disorders to continue differentiating between the two sexes.



## 4. GENERAL CONCLUSION

This research project has constructed a new clinical toolkit for the evaluation of spontaneous written language. We have laid the basis of a test that is non-intrusive, low-cost and time effective. The different studies within this dissertation contributed to finding out how spontaneous written language changes upon ageing, from a linguistic perspective. Our findings clearly indicate that:

1. age and image characteristics affect both naming correctness and speed, whereas repeated testing (of naming latencies) has no effect;
2. age and sex affect the cohesion of written narratives;
3. spontaneous written language is reliable in test-retest situations;
4. spontaneous written language can be used to make a differential diagnosis.

Additionally, we were able to lay the basis of a test that can be used in test-retest settings and could be used for a differential diagnosis. More specifically, we created a toolkit that:

5. bridges the gap within differential diagnostics with regard to spontaneous written language generation;
6. contains congeneric tasks, tasks that are interchangeable on certain key characteristics;
7. was standardised within a healthy ageing population and piloted its validity within a clinical setting.

By developing a new database with images that are strictly controlled for both visual and linguistic characteristics, we were able to build the foundation of congeneric writing tasks in our first study. In the second study, 'narrative writing tasks' were constructed and standardised in both a cross-sectional and longitudinal study. To our knowledge, the narrative writing task is one of the only tasks designed to allow for both spontaneous language generation and still control for certain key variables, the target nouns triggered by the images. The images are interchangeable per category, thus allowing for test-retest situations while still controlling for similar characteristics.

This dissertation's second contribution to the field is the validation of a normative database, that can be used as reference in clinical research. The focus for this normative base was cohesion in written narratives, wherefore we selected both linguistic and grammatical variables. Additionally, in our longitudinal study, we verified the test-retest reliability of our task. This finding allows our toolkit to be used in a longitudinal design. This normative base also allowed us to verify if a potential linguistic biomarker could be established. Our case studies pointed towards the vulnerability of '*Density of abstract verbs*'. More research is needed as to how this biomarker relates to different pathologies. Only a few steps are still needed to turn this toolkit into test that can be used within the field.

Given these findings and the foundation of a test into spontaneous language generation, we hope to inspire future studies into healthy ageing and ageing with an acquired neurologic language disorder. We believe to have laid the basis of a strong addition to the current set of screening tools in ageing of language.



## 5. BIBLIOGRAPHY

Aerts, A., van Mierlo, P., Hartsuiker, R. J., Santens, P., & De Letter, M. (2015). Sex Differences in Neurophysiological Activation Patterns During Phonological Input Processing: An Influencing Factor for Normative Data. *Archives of Sexual Behavior*, 44(8), 2207–2218. https://doi.org/10.1007/s10508-015-0560-y

Aggarwal, N. T., Wilson, R. S., Beck, T. L., Rajan, K. B., Mendes de Leon, C. F., Evans, D. A., & Everson-Rose, S. A. (2014). Perceived Stress and Change in Cognitive Function Among Adults 65 Years and Older. *Psychosomatic Medicine*, 76(1), 80–85. https://doi.org/10.1097/PSY.0000000000000016

Alatorre-Cruz, G. C., Silva-Pereyra, J., Fernández, T., Rodríguez-Camacho, M. A., Castro-Chavira, S. A., & Sanchez-Lopez, J. (2018). Effects of age and working memory load on syntactic processing: An event-related potential study. *Frontiers in Human Neuroscience,* 12(May), 1–11. https://doi.org/10.3389/fnhum.2018.00185

Alenius, M., Koskinen, S., Hallikainen, I., Ngandu, T., Lipsanen, J., Sainio, P., Tuulio-Henriksson, A., & Hänninen, T. (2019). Cognitive Performance among Cognitively Healthy Adults Aged 30-100 Years. *Dementia and Geriatric Cognitive Disorders Extra*, 9(1), 11–23. https://doi.org/10.1159/000495657

Anderson, E. L., Howe, L. D., Wade, K. H., Ben-Shlomo, Y., Hill, W. D., Deary, I. J., Sanderson, E. C., Zheng, J., Korologou-Linden, R., Stergiakouli, E., Davey Smith, G., Davies, N. M., & Hemani, G. (2020). Education, intelligence and Alzheimer's disease: Evidence from a multivariable two-sample Mendelian randomization study. *International Journal of Epidemiology*, 49(4), 1163–1172. https://doi.org/10.1093/ije/dyz280

Avan, A., & Hachinski, V. (2021). Stroke and dementia, leading causes of neurological disability and death, potential for prevention. *Alzheimer's and Dementia*, 17(6), 1072–1076. https://doi.org/10.1002/alz.12340

Boukadi, M., Zouaidi, C., & Wilson, M. A. (2016). Norms for name agreement, familiarity, subjective frequency, and imageability for 348 object names in Tunisian Arabic. *Behavior Research Methods,* 48(2), 585–599. https://doi.org/10.3758/s13428-015-0602-3

Burke, D. M., & Mackay, D. G. (1997). Memory, language, and ageing. *Philos Trans R Soc Lond B Biol Sci,* 352(1363), 1845–1856. https://doi.org/10.1098/rstb.1997.0170

Campos-Magdaleno, M., Leiva, D., Pereiro, A. X., Lojo-Seoane, C., Mallo, S. C., Nieto-Vieites, A., Juncos-Rabadán, O., & Facal, D. (2020). Longitudinal Patterns of the Tip-of-the-Tongue Phenomenon in People With Subjective Cognitive Complaints and Mild Cognitive Impairment. *Frontiers in Psychology,* 11(March), 1–14. https://doi.org/10.3389/fpsyg.2020.00425

Clark-Cotton, M. R., Williams, R. K., Goral, M., & Obler, Loraine, K. (2007). *Language and Communication in Aging* (pp. 1–8). https://doi.org/10.1016/B0-12-370870-2/00103-7

Duboisa, B., Padovanib, A., Scheltensc, P., Rossid, A., & Agnello, G. D. (2015). Timely diagnosis for alzheimer's disease: A literature review on benefits and challenges. *Journal of Alzheimer's Disease*, 49(3), 617–631. https://doi.org/10.3233/JAD-150692

Fahimi Hnazaee, M. (2020). *Spatiotemporal Models of Invasive and Noninvasive Eeg Recordings During Abstract and Concrete Word Processing in the [KU Leuven]*. https://lirias.kuleuven.be/retrieve/591134$$DPhDThesis_Final_MFH.pdf

Fahimi Hnazaee, M., Khachatryan, E., Chehrazad, S., Kotarcic, A., De Letter, M., & Van Hulle, M. M. (2020). *Overlapping connectivity patterns during semantic processing of abstract and concrete words revealed with multivariate Granger Causality analysis. Scientific Reports*, 10(1), 1–19. https://doi.org/10.1038/s41598-020-59473-7

Feigin, V. L., Nichols, E., Alam, T., Bannick, M. S., Beghi, E., Blake, N., Culpepper, W. J., Dorsey, E. R., Elbaz, A., Ellenbogen, R. G., Fisher, J. L., Fitzmaurice, C., Giussani, G., Glennie, L., James, S. L., Johnson, C. O., Kassebaum, N. J., Logroscino, G., Marin, B., … Vos, T. (2019). Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology,* 18(5), 459–480. https://doi.org/10.1016/S1474-4422(18)30499-X

Forbes-McKay, K. E., & Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurological Sciences,* 26(4), 243–254. https://doi.org/10.1007/s10072-005-0467-9

Forbes-McKay, K., Shanks, M., & Venneri, A. (2014). Charting the decline in spontaneous writing in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatrica*, 26(04), 246–252. https://doi.org/10.1017/neu.2014.2

Francis, A., & Pandian, I. A. (2021). Early detection of Alzheimer's disease using local binary pattern and convolutional neural network. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-021-11161-y

Frid, J., Johansson, V., Johansson, R., Wengelin, Å., & Johansson, M. (2014). Developing a keystroke logging program into a writing experiment environment. *Presentation at Writing Across Borders,* WRAB, Paris February 2014.

Gates, N., Rutjes, A., Di Nisio, M., Karim, S., Chong, L., March, E., Martínez, G., & Vernooij, R. (2019). in cognitively healthy people in late life. *Cochrane Database of Systematic Reviews*, 3. https://doi.org/10.1002/14651858.CD012277.pub2.www.cochranelibrary.com

Gordon, E., Palmer, D. M., Liu, H., Rekshan, W., & DeVarney, S. (2013). Online Cognitive Brain Training Associated With Measurable Improvements in Cognition and Emotional Well-Being. *Technology & Innovation*, 15(1), 53–62. https://doi.org/10.3727/194982413x13608676060574

Hier, D. B., Hagenlocker, K., & Shindler, A. G. (1985). Language disintegration in dementia: Effects of etiology and severity. *Brain and Language*, 25(1), 117–133. https://doi.org/10.1016/0093-934X(85)90124-5



Hydén, L. C., & Örulv, L. (2009). Narrative and identity in Alzheimer's disease: A case study. *Journal of Aging Studies*, 23(4), 205–214. https://doi.org/10.1016/j.jaging.2008.01.001

Kavé, G., & Goral, M. (2018). Do age-related word retrieal difficulties appear (or disappear) in connected speech? *Neurospychol Dev Cogn B Aging Neuropsychol Cogn*, 24(5), 508–527. https://doi.org/10.1080/13825585.2016.1226249

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication,* 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Li, F., & Liu, M. (2018). Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Computerized Medical Imaging and Graphics*, 70, 101–110. https://doi.org/10.1016/j.compmedimag.2018.09.009

Paesen, L., & Meulemans, C. (2020). *Open Linguistic Picture Database [Data set].* Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.3738213

Pander Maat, H., Kraf, R., Bosch, A. van den, Dekker, N., Gompel, M. van, Kleijn, S., Sanders, T. J. M., & Sloot, K. van der. (2014). T-Scan : a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.

Prince, M., Bryce, R., & Ferri, C. (2011). World Alzheimer Report 2011: The benefits of early diagnosis and intervention Alzheimer's Disease International World Alzheimer Report 2011 The benefits of early diagnosis and intervention. *In World Disease International*. https://www.alz.co.uk/research/WorldAlzheimerReport2011.pdf

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., Marsiske, M., Morris, J. N., Tennstedt, S. L., Unverzagt, F. W., & Willis, S. L. (2014). Ten-Year Effects of the Advanced Cognitive Training for Independent and Vital Elderly Cognitive Training Trial on Cognition and Everyday Functioning in Older Adults. *Journal of the American Geriatrics Society*, 62(1), 16–24. https://doi.org/10.1111/jgs.12607

Rofes, A., Talacchi, A., Santini, B., Pinna, G., Nickels, L., Bastiaanse, R., & Miceli, G. (2018). Language in individuals with left hemisphere tumors: Is spontaneous speech analysis comparable to formal testing? *Journal of Clinical and Experimental Neuropsychology*, 40(7), 722–732. https://doi.org/10.1080/13803395.2018.1426734

Scaltritti, M., Arfé, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition*, 156, 16–29. https://doi.org/10.1016/j.cognition.2016.07.006

Scott, S. B., Graham-Engeland, J. E., Engeland, C. G., Smyth, J. M., Almeida, D. M., Katz, M. J., Lipton, R. B., Mogle, J. A., Munoz, E., Ram, N., & Sliwinski, M. J. (2015). The Effects of Stress on Cognitive Aging, Physiology and Emotion (ESCAPE) Project. *BMC Psychiatry,* 15(1), 1–14. https://doi.org/10.1186/s12888-015-0497-7

Segaert, K., Lucas, S. J. E., Burley, C. V., Segaert, P., Milner, A. E., Ryan, M., & Wheeldon, L. (2018). Higher physical fitness levels are associated with less language decline in healthy ageing. *Scientific Reports*, 8(1), 1–10. https://doi.org/10.1038/s41598-018-24972-1

Shafto, M. A., & Tyler, L. K. (2014). Decline and Preservation. *Science*, 346(6209), 583–588. https://doi.org/10.1126/science.1254404

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior,* 73, 311–319. https://doi.org/10.1016/j.chb.2017.03.050

Wada, M., Noda, Y., Shinagawa, S., Chung, J. K., Sawada, K., Ogyu, K., Tarumi, R., Tsugawa, S., Miyazaki, T., Yamagata, B., Graff-Guerrero, A., Mimura, M., & Nakajima, S. (2018). Effect of Education on Alzheimer's Disease-Related Neuroimaging Biomarkers in Healthy Controls, and Participants with Mild Cognitive Impairment and Alzheimer's Disease: A Cross-Sectional Study. *Journal of Alzheimer's Disease* : JAD, 63(2), 861–869. https://doi.org/10.3233/JAD-171168

WHO. (2020). Decade of Healthy Ageing. *World Health Organization,* 1–24. https://www.who.int/initiatives/decade-of-healthy-ageing%0Ahttps://www.who.int/initiatives/decade-of-healthy-ageing/connection-series%0Ahttps://www.who.int/ageing/decade-of-healthy-ageing

Wilson, R. S., Barnes, L. L., Bennett, D. A., Li, Y., Bienias, J. L., de Leon, C. F. M., & Evans, D. A. (2005). Proneness to psychological distress and risk of Alzheimer disease in a biracial community. *Neurology*, 64(2), 380–382. https://doi.org/10.1212/01.WNL.0000149525.53525.E7

Wilson, Robert S., Arnold, S. E., Schneider, J. A., Li, Y., & Bennett, D. A. (2007). Chronic Distress, Age-Related Neuropathology, and Late-Life Dementia. *Psychosomatic Medicine*, 69(1), 47–53. https://doi.org/10.1097/01.psy.0000250264.25017.21

# Chapter seven

## Appendices

# APPENDIX A

**Images used to construct the narrative writing tasks and picture naming task. Dutch translations were provided where necessary. (Chapter two)**

Anchor (anker) ©

Car (auto) ©

Ball (bal) ©

Bed ©

Book (boek) ©

Glasses (bril) ©

Bus ©

Diamond (diamant) ©

Dolphin (dolfijn) ©

Box (doos) ©

Squirrel (eekhoorn) ©

Elk (eland) ©

Eskimo ©

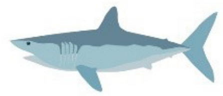
Bycicle (fiets) ©

Bottle (fles) ©

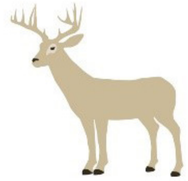
Guitar (gitaar) ©



Shark (haai) ©



Deer (hert) ©



Dog (hond) ©

kangaroo (kangoeroe) ©

Cat (kat) ©

Chicken (kip) ©

Cow (koe) ©

Crab (krab) ©

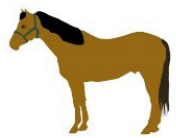
Lobster (kreeft) ©

Wheelbarrow (kruiwagen) ©

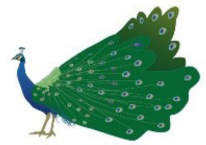
Lamp ©

Medal (medaille) ©

Horse (paard) ©



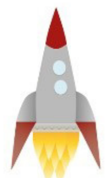
Peacock (pauw) ©



Piano ©

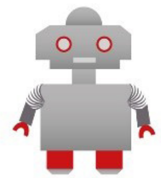
Pirate (piraat) ©

Rocket (raket) ©



Ring ©

Robot ©



Roller skate (rolschaats) ©

Step ©

Chair
(stoel) ©

Ironing board
(strijkplank) ©

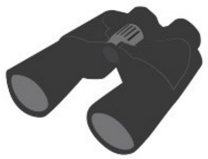
Wizard
(tovenaar) ©

Tractor ©

Train
(trein) ©

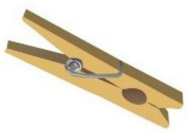
Trumpet
(trompet) ©

Binoculars
(verrekijker) ©



Plane
(vliegtuig) ©

Clothespin
(wasknijper) ©



Sailing boat
(zeilboot) ©

Swan
(zwaan) ©





# APPENDIX B

**Linear mixed models - Effects of age, sex, and task on the product variables (Chapter 3)**

| Model | General overview of the data |
|---|---|
| 0 | Random variance |
| 1 | Effect of age |
| 2 | Effect of age and sex |
| 3 | Effect of age, sex and task |
| 4 | Effect of task and interaction effect (age x sex) |
| 5 | Interaction effect (age x sex x task) |

*General overview - Lexical*

| Total number of words | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 6190.700 | 3 | | | |
| Model 1 | 6144.420 | 4 | 46.280 | 1 | 0.000 |
| Model 2 | 6144.230 | 5 | 0.190 | 1 | 0.663 |
| Model 3 | 6138.241 | 6 | 5.989 | 1 | 0.014 |
| Model 4 | 6138.000 | 7 | 0.241 | 1 | 0.624 |
| Model 5 | 6129.491 | 10 | 8.509 | 3 | 0.037 |
| **Word length** | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 291.396 | 3 | | | |
| Model 1 | 289.806 | 4 | 1.590 | 1 | 0.207 |
| Model 2 | 281.537 | 5 | 8.268 | 1 | 0.004 |
| Model 3 | 146.635 | 6 | 134.902 | 1 | 0.000 |
| Model 4 | 145.213 | 7 | 1.422 | 1 | 0.233 |
| Model 5 | 143.806 | 10 | 1.407 | 3 | 0.704 |

| Word frequency nouns | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 242.505 | 3 | | | |
| Model 1 | 226.206 | 4 | 16.300 | 1 | 0.000 |
| Model 2 | 222.164 | 5 | 4.042 | 1 | 0.044 |
| Model 3 | 90.888 | 6 | 131.275 | 1 | 0.000 |
| Model 4 | 89.724 | 7 | 1.164 | 1 | 0.281 |
| Model 5 | 87.184 | 10 | 2.540 | 3 | 0.468 |

| Type token ratio words | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | -1175.194 | 3 | | | |
| Model 1 | -1228.691 | 4 | 53.497 | 1 | 0.000 |
| Model 2 | -1229.044 | 5 | 0.352 | 1 | 0.553 |
| Model 3 | -1231.319 | 6 | 2.275 | 1 | 0.131 |
| Model 4 | -1231.347 | 7 | 0.028 | 1 | 0.866 |
| Model 5 | -1234.144 | 10 | 2.796 | 3 | 0.424 |

| Density of adjectives | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 5089.871 | 3 | | | |
| Model 1 | 5088.941 | 4 | 0.931 | 1 | 0.335 |
| Model 2 | 5088.938 | 5 | 0.002 | 1 | 0.961 |
| Model 3 | 5087.838 | 6 | 1.101 | 1 | 0.294 |
| Model 4 | 5087.826 | 7 | 0.011 | 1 | 0.915 |
| Model 5 | 5087.191 | 10 | 0.635 | 3 | 0.888 |



| Density of nouns | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 5384.950 | 3 | | | |
| Model 1 | 5381.857 | 4 | 3.093 | 1 | 0.079 |
| Model 2 | 5372.199 | 5 | 9.658 | 1 | 0.002 |
| Model 3 | 5362.555 | 6 | 9.644 | 1 | 0.002 |
| Model 4 | 5362.491 | 7 | 0.064 | 1 | 0.800 |
| Model 5 | 5358.797 | 10 | 3.694 | 3 | 0.296 |

| Density of verbs | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 5233.230 | 3 | | | |
| Model 1 | 5233.143 | 4 | 0.087 | 1 | 0.768 |
| Model 2 | 5222.401 | 5 | 10.742 | 1 | 0.001 |
| Model 3 | 5222.195 | 6 | 0.206 | 1 | 0.650 |
| Model 4 | 5222.134 | 7 | 0.061 | 1 | 0.805 |
| Model 5 | 5219.914 | 10 | 2.220 | 3 | 0.528 |

| Density of concrete verbs | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 4447.164 | 3 | | | |
| Model 1 | 4447.163 | 4 | 0.001 | 1 | 0.975 |
| Model 2 | 4446.905 | 5 | 0.258 | 1 | 0.611 |
| Model 3 | 4445.158 | 6 | 1.747 | 1 | 0.186 |
| Model 4 | 4445.131 | 7 | 0.026 | 1 | 0.871 |
| Model 5 | 4444.248 | 10 | 0.883 | 3 | 0.830 |

| Density of abstract verbs | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 5268.506 | 3 | | | |
| Model 1 | 5268.080 | 4 | 0.426 | 1 | 0.514 |
| Model 2 | 5263.692 | 5 | 4.388 | 1 | 0.036 |
| Model 3 | 5260.296 | 6 | 3.396 | 1 | 0.065 |
| Model 4 | 5259.631 | 7 | 0.665 | 1 | 0.415 |
| Model 5 | 5257.270 | 10 | 2.361 | 3 | 0.501 |

*Grammatical*

| D-Level | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | p |
| Model 0 | 1818.192 | 3 | | | |
| Model 1 | 1812.972 | 4 | 5.219 | 1 | 0.022 |
| Model 2 | 1812.966 | 5 | 0.007 | 1 | 0.934 |
| Model 3 | 1811.675 | 6 | 1.291 | 1 | 0.256 |
| Model 4 | 1810.625 | 7 | 1.051 | 1 | 0.305 |
| Model 5 | 1809.081 | 10 | 1.543 | 3 | 0.672 |
| **Distance subject-verb** | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 1432.189 | 3 | | | |
| Model 1 | 1430.045 | 4 | 2.144 | 1 | 0.143 |
| Model 2 | 1423.426 | 5 | 6.619 | 1 | 0.010 |
| Model 3 | 1419.255 | 6 | 4.171 | 1 | 0.041 |
| Model 4 | 1419.137 | 7 | 0.118 | 1 | 0.731 |
| Model 5 | 1410.925 | 10 | 8.212 | 3 | 0.042 |



| Distance determiner-noun | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | -464.734 | 3 | | | |
| Model 1 | -468.084 | 4 | 3.350 | 1 | 0.067 |
| Model 2 | -468.145 | 5 | 0.060 | 1 | 0.806 |
| Model 3 | -469.822 | 6 | 1.677 | 1 | 0.195 |
| Model 4 | -470.123 | 7 | 0.302 | 1 | 0.583 |
| Model 5 | -472.218 | 10 | 2.094 | 3 | 0.553 |
| **Density of personal references** | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 5394.041 | 3 | | | |
| Model 1 | 5385.721 | 4 | 8.320 | 1 | 0.004 |
| Model 2 | 5384.157 | 5 | 1.564 | 1 | 0.211 |
| Model 3 | 5378.453 | 6 | 5.704 | 1 | 0.017 |
| Model 4 | 5377.669 | 7 | 0.784 | 1 | 0.376 |
| Model 5 | 5376.624 | 10 | 1.045 | 3 | 0.790 |
| **Density of conjunctions** | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 4783.353 | 3 | | | |
| Model 1 | 4781.796 | 4 | 1.557 | 1 | 0.212 |
| Model 2 | 4769.565 | 5 | 12.231 | 1 | 0.000 |
| Model 3 | 4766.840 | 6 | 2.724 | 1 | 0.099 |
| Model 4 | 4766.185 | 7 | 0.656 | 1 | 0.418 |
| Model 5 | 4765.390 | 10 | 0.795 | 3 | 0.851 |

**Linear mixed models - Estimates of fixed effects (Chapter 3)**

| Total words | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 155.920 | 5.250 | .000 | 273.370 | 17.212 | .000 | 276.895 | 19.016 | .000 | 27.994 | 19.179 | .000 | 259.381 | 3.451 | .000 | 254.120 | 34.269 | .000 |
| Age | | | | -1.969 | .277 | .000 | -1.986 | .280 | .000 | -1.977 | .280 | .000 | -1.791 | .471 | .000 | -1.639 | .529 | .002 |
| Sex | | | | | | | -4.230 | 9.702 | .663 | -3.701 | 9.708 | .703 | 13.629 | 36.613 | .710 | -2.784 | 4.941 | .946 |
| Task | | | | | | | | | | 9.966 | 4.046 | .014 | 1.008 | 4.047 | .014 | 19.615 | 26.985 | .468 |
| Age*sex | | | | | | | | | | | | | -.286 | .584 | .624 | -.122 | .648 | .850 |
| Age*task | | | | | | | | | | | | | | | | -.292 | .413 | .480 |
| Sex*task | | | | | | | | | | | | | | | | 33.813 | 32.501 | .299 |
| Age*sex*task | | | | | | | | | | | | | | | | -.343 | .509 | .501 |

| Word length | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 4.360 | .015 | .000 | 4.424 | .053 | .000 | 4.496 | .058 | .000 | 4.344 | .058 | .000 | 4.258 | .093 | .000 | 4.231 | .120 | .000 |
| Age | | | | -.001 | .001 | .208 | -.001 | .001 | .089 | -.001 | .001 | .137 | .000 | .001 | .932 | .000 | .002 | .855 |
| Sex | | | | | | | -.085 | .029 | .004 | -.075 | .029 | .011 | .052 | .111 | .636 | .069 | .142 | .627 |
| Task | | | | | | | | | | .267 | .020 | .000 | .267 | .020 | .000 | .312 | .135 | .021 |
| Age*sex | | | | | | | | | | | | | -.002 | .002 | .234 | -.002 | .002 | .377 |
| Age*task | | | | | | | | | | | | | | | | -.000 | .002 | .884 |
| Sex*task | | | | | | | | | | | | | | | | -.024 | .162 | .882 |
| Age*sex*task | | | | | | | | | | | | | | | | -.000 | .003 | .881 |

**Word frequency of nouns**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 4.552 | .013 | .000 | 4.735 | .047 | .000 | 4.690 | .052 | .000 | 4.839 | .052 | .000 | 4.909 | .083 | .000 | 4.941 | .112 | .000 |
| Age | | | | -.003 | .001 | .000 | -.003 | .001 | .000 | -.003 | .001 | .000 | -.004 | .001 | .001 | -.004 | .002 | .012 |
| Sex | | | | | | | .053 | .026 | .044 | .044 | .026 | .095 | -.060 | .099 | .547 | -.134 | .133 | .315 |
| Task | | | | | | | | | | -.265 | .020 | .000 | -.266 | .020 | .000 | -.319 | .135 | .019 |
| Age*sex | | | | | | | | | | | | | .002 | .002 | .281 | .002 | .002 | .244 |
| Age*task | | | | | | | | | | | | | | | | .000 | .002 | .891 |
| Sex*task | | | | | | | | | | | | | | | | .138 | .163 | .398 |
| Age*sex*task | | | | | | | | | | | | | | | | -.001 | .003 | .604 |

**Type token ratio of words**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | .642 | .005 | .000 | .526 | .016 | .000 | .522 | .017 | .000 | .526 | .018 | .000 | .530 | .028 | .000 | .507 | .034 | .000 |
| Age | | | | .002 | .000 | .000 | .002 | .000 | .000 | .002 | .000 | .000 | .002 | .000 | .000 | .002 | .001 | .000 |
| Sex | | | | | | | .005 | .009 | .553 | .005 | .009 | .573 | -.000 | .034 | .989 | .024 | .041 | .558 |
| Task | | | | | | | | | | -.008 | .005 | .132 | -.008 | .005 | .131 | .032 | .034 | .359 |
| Age*sex | | | | | | | | | | | | | .000 | .001 | .866 | -.000 | .001 | .789 |
| Age*task | | | | | | | | | | | | | | | | -.000 | .001 | .357 |
| Sex*task | | | | | | | | | | | | | | | | -.042 | .041 | .314 |
| Age*sex*task | | | | | | | | | | | | | | | | .000 | .001 | .512 |

**Density of adjectives**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 71.036 | 1.285 | .000 | 75.346 | 4.644 | .000 | 75.238 | 5.150 | .000 | 76.482 | 5.282 | .000 | 75.783 | 8.432 | .000 | 75.586 | 11.553 | .000 |
| Age | | | | -.072 | .075 | .335 | -.071 | .075 | .344 | -.073 | .075 | .335 | -.062 | .128 | .632 | -.072 | .177 | .685 |
| Sex | | | | | | | .128 | 2.630 | .961 | .072 | 2.630 | .978 | 1.104 | 1.042 | .913 | 3.981 | 13.689 | .771 |
| Task | | | | | | | | | | -2.251 | 2.144 | .295 | -2.246 | 2.145 | .296 | -2.025 | 14.252 | .887 |
| Age*sex | | | | | | | | | | | | | -.017 | .159 | .915 | -.043 | .214 | .840 |
| Age*task | | | | | | | | | | | | | | | | .022 | .219 | .918 |
| Sex*task | | | | | | | | | | | | | | | | -5.590 | 17.139 | .745 |
| Age*sex*task | | | | | | | | | | | | | | | | .050 | .269 | .852 |

**Density of nouns**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 185.950 | 1.810 | .000 | 174.941 | 6.493 | .000 | 184.474 | 7.066 | .000 | 179.935 | 7.202 | .000 | 177.667 | 11.496 | .000 | 174.082 | 15.022 | .000 |
| Age | | | | .184 | .104 | .079 | .136 | .104 | .189 | .141 | .103 | .172 | .177 | .176 | .314 | .186 | .230 | .419 |
| Sex | | | | | | | -11.338 | 3.614 | .002 | -11.107 | 3.609 | .002 | -7.753 | 13.733 | .573 | -3.204 | 17.827 | .857 |
| Task | | | | | | | | | | 8.147 | 2.603 | .002 | 8.161 | 2.604 | .002 | 13.988 | 17.276 | .419 |
| Age*sex | | | | | | | | | | | | | -.055 | .217 | .800 | -.044 | .279 | .874 |
| Age*task | | | | | | | | | | | | | | | | .002 | .265 | .995 |
| Sex*task | | | | | | | | | | | | | | | | -7.694 | 2.788 | .712 |
| Age*sex*task | | | | | | | | | | | | | | | | -.041 | .326 | .900 |

**Density of verbs**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 189.112 | 1.545 | .000 | 187.531 | 5.581 | .000 | 178.907 | 6.073 | .000 | 178.317 | 6.210 | .000 | 18.222 | 9.912 | .000 | 167.360 | 13.156 | .000 |
| Age | | | | .026 | .090 | .768 | .069 | .089 | .436 | .070 | .089 | .431 | .040 | .151 | .792 | .238 | .201 | .238 |
| Sex | | | | | | | 1.265 | 3.105 | .001 | 1.293 | 3.106 | .001 | 7.480 | 11.830 | .528 | 2.959 | 15.605 | .180 |
| Task | | | | | | | | | | 1.059 | 2.331 | .650 | 1.047 | 2.332 | .654 | 23.793 | 15.466 | .125 |
| Age*sex | | | | | | | | | | | | | .046 | .187 | .805 | -.160 | .244 | .513 |
| Age*task | | | | | | | | | | | | | | | | -.353 | .237 | .138 |
| Sex*task | | | | | | | | | | | | | | | | -23.965 | 18.607 | .199 |
| Age*sex*task | | | | | | | | | | | | | | | | .369 | .292 | .207 |

**Density of concrete verbs**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 18.542 | .663 | .000 | 18.470 | 2.406 | .000 | 17.880 | 2.669 | .000 | 18.790 | 2.756 | .000 | 18.235 | 4.399 | .000 | 17.107 | 6.331 | .007 |
| Age | | | | .001 | .039 | .976 | .004 | .039 | .914 | .003 | .039 | .933 | .012 | .067 | .857 | .030 | .097 | .760 |
| Sex | | | | | | | .692 | 1.361 | .612 | .657 | 1.361 | .630 | 1.473 | 5.219 | .778 | .224 | 7.492 | .976 |
| Task | | | | | | | | | | -1.661 | 1.255 | .187 | -1.657 | 1.255 | .188 | .372 | 8.294 | .964 |
| Age*sex | | | | | | | | | | | | | -.013 | .082 | .871 | .010 | .117 | .930 |
| Age*task | | | | | | | | | | | | | | | | -.032 | .127 | .803 |
| Sex*task | | | | | | | | | | | | | | | | 2.736 | 9.965 | .784 |
| Age*sex*task | | | | | | | | | | | | | | | | -.051 | .157 | .746 |

| **Density of abstract verbs** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 114.712 | 1.553 | .000 | 118.234 | 5.611 | .000 | 112.635 | 6.174 | .000 | 11.110 | 6.323 | .000 | 103.697 | 1.083 | .000 | 92.856 | 13.573 | .000 |
| Age | | | | -.059 | .090 | .514 | -.031 | .090 | .735 | -.028 | .090 | .759 | .074 | .154 | .632 | .233 | .208 | .262 |
| Sex | | | | | | | 6.632 | 3.155 | .036 | 6.748 | 3.156 | .033 | 16.216 | 12.023 | .179 | 22.569 | 16.092 | .161 |
| Task | | | | | | | | | | 4.552 | 2.462 | .066 | 4.593 | 2.462 | .063 | 23.772 | 16.317 | .146 |
| Age*sex | | | | | | | | | | | | | -.155 | .190 | .415 | -.229 | .252 | .364 |
| Age*task | | | | | | | | | | | | | | | | -.283 | .251 | .260 |
| Sex*task | | | | | | | | | | | | | | | | -1.147 | 19.627 | .606 |
| Age*sex*task | | | | | | | | | | | | | | | | .112 | .308 | .718 |

| **D-level** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 2.685 | .072 | .000 | 2.120 | .257 | .000 | 2.130 | .284 | .000 | 2.072 | .289 | .000 | 2.440 | .459 | .000 | 2.563 | .570 | .000 |
| Age | | | | .009 | .004 | .023 | .009 | .004 | .025 | .009 | .004 | .024 | .004 | .007 | .606 | .003 | .009 | .760 |
| Sex | | | | | | | -.012 | .145 | .934 | -.009 | .145 | .953 | -.553 | .550 | .315 | -.627 | .678 | .356 |
| Task | | | | | | | | | | .101 | .089 | .256 | .099 | .089 | .266 | -.107 | .596 | .858 |
| Age*sex | | | | | | | | | | | | | .009 | .009 | .306 | .008 | .011 | .434 |
| Age*task | | | | | | | | | | | | | | | | .001 | .009 | .883 |
| Sex*task | | | | | | | | | | | | | | | | .102 | .717 | .887 |
| Age*sex*task | | | | | | | | | | | | | | | | .002 | .011 | .867 |

| **Distance subject-verb** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 1.996 | .044 | .000 | 2.220 | .159 | .000 | 2.415 | .174 | .000 | 2.335 | .178 | .000 | 2.411 | .285 | .000 | 1.710 | .383 | .000 |
| Age | | | | -.004 | .003 | .143 | -.005 | .003 | .065 | -.005 | .003 | .070 | -.006 | .004 | .180 | .005 | .006 | .361 |
| Sex | | | | | | | -.230 | .089 | .010 | -.227 | .089 | .011 | -.339 | .340 | .318 | .354 | .454 | .436 |
| Task | | | | | | | | | | .144 | .070 | .041 | .143 | .070 | .042 | 1.395 | .461 | .003 |
| Age*sex | | | | | | | | | | | | | .002 | .005 | .731 | -.009 | .007 | .187 |
| Age*task | | | | | | | | | | | | | | | | -.020 | .007 | .005 |
| Sex*task | | | | | | | | | | | | | | | | -1.235 | .554 | .027 |
| Age*sex*task | | | | | | | | | | | | | | | | .020 | .009 | .020 |

**Distance determiner-noun**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | .224 | .008 | .000 | .271 | .027 | .000 | .275 | .030 | .000 | .266 | .031 | .000 | .245 | .049 | .000 | .230 | .067 | .001 |
| Age | | | | -.001 | .000 | .068 | -.001 | .000 | .065 | -.001 | .000 | .068 | -.000 | .001 | .530 | -.000 | .001 | .765 |
| Sex | | | | | | | -.004 | .015 | .806 | -.003 | .015 | .826 | .028 | .058 | .637 | .010 | .079 | .896 |
| Task | | | | | | | | | | .016 | .012 | .196 | .016 | .012 | .192 | .041 | .081 | .611 |
| Age*sex | | | | | | | | | | | | | -.001 | .001 | .583 | -.000 | .001 | .949 |
| Age*task | | | | | | | | | | | | | | | | -.000 | .001 | .823 |
| Sex*task | | | | | | | | | | | | | | | | .039 | .098 | .693 |
| Age*sex*task | | | | | | | | | | | | | | | | -.001 | .002 | .552 |

**Density of personal references**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 124.191 | 1.713 | .000 | 141.256 | 6.110 | .000 | 137.600 | 6.757 | .000 | 141.313 | 6.934 | .000 | 133.674 | 11.053 | .000 | 143.100 | 15.120 | .000 |
| Age | | | | -.285 | .098 | .004 | -.266 | .099 | .008 | -.270 | .099 | .007 | -.150 | .169 | .376 | -.289 | .231 | .212 |
| Sex | | | | | | | 4.322 | 3.451 | .212 | 4.164 | 3.454 | .229 | 15.426 | 13.165 | .242 | 3.018 | 17.910 | .866 |
| Task | | | | | | | | | | -6.726 | 2.800 | .017 | -6.673 | 2.801 | .018 | -23.430 | 18.580 | .208 |
| Age*sex | | | | | | | | | | | | | -.184 | .208 | .376 | -.000 | .280 | 1.000 |
| Age*task | | | | | | | | | | | | | | | | .248 | .285 | .385 |
| Sex*task | | | | | | | | | | | | | | | | 22.710 | 22.350 | .310 |
| Age*sex*task | | | | | | | | | | | | | | | | -.338 | .351 | .336 |

**Density of conjunctions**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 54.692 | 1.052 | .000 | 5.150 | 3.785 | .000 | 43.925 | 4.098 | .000 | 42.550 | 4.181 | .000 | 46.761 | 6.666 | .000 | 44.202 | 8.677 | .000 |
| Age | | | | .076 | .061 | .213 | .107 | .060 | .076 | .109 | .060 | .072 | .042 | .102 | .683 | .073 | .133 | .583 |
| Sex | | | | | | | 7.421 | 2.096 | .000 | 7.488 | 2.097 | .000 | 1.261 | 7.966 | .874 | 5.916 | 1.299 | .566 |
| Task | | | | | | | | | | 2.455 | 1.483 | .099 | 2.430 | 1.484 | .103 | 6.862 | 9.894 | .489 |
| Age*sex | | | | | | | | | | | | | .102 | .126 | .418 | .041 | .161 | .798 |
| Age*task | | | | | | | | | | | | | | | | -.053 | .152 | .729 |
| Sex*task | | | | | | | | | | | | | | | | -8.631 | 11.906 | .469 |
| Age*sex*task | | | | | | | | | | | | | | | | .112 | .187 | .551 |

## APPENDIX C

**Normative table of lexical measures - part 1**

| | | | Male | | | | | | Female | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NWT1 | | | NWT2 | | | NWT1 | | | NWT2 | | |
| | | | M | SD | Range | M | SD | Range | M | SD | Range | M | SD | Range |
| Baseline | 20 – 35 | Total words | 189 | 70 | 109-331 | 246 | 73 | 131-363 | 188 | 50 | 121-317 | 206 | 70 | 108-385 |
| | | Word length | 4.23 | 0.21 | 3.88-4.65 | 4.44 | 0.22 | 4.25-4.95 | 4.28 | 0.19 | 3.87-4.63 | 4.46 | 0.25 | 4.09-4.99 |
| | | Word frequency | 4.99 | 0.18 | 4.67-5.12 | 4.61 | 0.24 | 4.17-4.98 | 4.79 | 0.19 | 4.35-5.10 | 4.66 | 0.25 | 4.02-5.03 |
| | | Type token ratio | 0.53 | 0.04 | 0.50-0.62 | 0.55 | 0.06 | 0.51-0.71 | 0.59 | 0.06 | 0.47-0.70 | 0.55 | 0.07 | 0.41-0.66 |
| Target | 50 - 59 | Total words | 146 | 74 | 29-309 | 160 | 59 | 31-267 | 145 | 85 | 43-391 | 144 | 83 | 22-359 |
| | | Word length | 4.27 | 0.19 | 3.91-4.71 | 4.48 | 0.27 | 4.30-5.27 | 4.24 | 0.22 | 3.84-4.77 | 4.45 | 0.25 | 4.02-5.09 |
| | | Word frequency | 4.71 | 0.23 | 4.22-5.11 | 4.37 | 0.22 | 3.96-4.84 | 4.72 | 0.22 | 4.26-5.08 | 4.46 | 0.24 | 3.89-5.03 |
| | | Type token ratio | 0.64 | 0.08 | 0.46-0.79 | 0.63 | 0.05 | 0.50-0.68 | 0.6 | 0.09 | 0.43-0.81 | 0.6 | 0.08 | 0.41-0.77 |
| | 60 - 69 | Total words | 127 | 65 | 25-291 | 124 | 49 | 43-260 | 118 | 58 | 35-271 | 130 | 53 | 34-252 |
| | | Word length | 4.28 | 0.33 | 3.53-4.88 | 4.61 | 0.28 | 4.02-5.12 | 4.25 | 0.27 | 3.65-4.77 | 4.45 | 0.21 | 3.96-5.00 |
| | | Word frequency | 4.65 | 0.23 | 4.27-5.22 | 4.43 | 0.24 | 3.91-4.84 | 4.66 | 0.23 | 4.26-5.26 | 4.49 | 0.22 | 4.00-4.86 |
| | | Type token ratio | 0.63 | 0.08 | 0.50-0.84 | 0.65 | 0.06 | 0.54-0.77 | 0.65 | 0.08 | 0.48-0.83 | 0.65 | 0.08 | 0.50-0.82 |
| | 70 - 79 | Total words | 104 | 54 | 34-203 | 129 | 61 | 38-253 | 97 | 67 | 26-250 | 112 | 74 | 31-286 |
| | | Word length | 4.2 | 0.39 | 3.40-5.02 | 4.58 | 0.34 | 3.95-5.31 | 4.11 | 0.28 | 3.67-4.80 | 4.4 | 0.25 | 3.91-4.79 |
| | | Word frequency | 4.69 | 0.27 | 4.14-5.15 | 4.36 | 0.27 | 3.66-4.84 | 4.73 | 0.26 | 4.22-5.33 | 4.41 | 0.22 | 4.07-4.99 |
| | | Type token ratio | 0.67 | 0.1 | 0.46-0.90 | 0.67 | 0.08 | 0.51-0.82 | 0.69 | 0.09 | 0.51-0.82 | 0.68 | 0.06 | 0.57-0.80 |
| | 80+ | Total words | 94 | 33 | 40-136 | 100 | 61 | 38-239 | 71 | 49 | 21-170 | 88 | 57 | 42-203 |
| | | Word length | 4.32 | 0.27 | 3.66-4.60 | 4.5 | 0.19 | 4.13-4.66 | 4.08 | 0.35 | 3.58-4.70 | 4.43 | 0.34 | 3.83-4.80 |
| | | Word frequency | 4.69 | 0.28 | 4.34-5.25 | 4.39 | 0.27 | 3.84-4.78 | 4.68 | 0.29 | 4.08-4.92 | 4.47 | 0.28 | 4.02-4.97 |
| | | Type token ratio | 0.66 | 0.09 | 0.58-0.81 | 0.69 | 0.04 | 0.61-0.74 | 0.71 | 0.14 | 0.54-0.93 | 0.64 | 0.11 | 0.47-0.86 |

Note: M = Mean, 'Total words' = Total number of words', 'Word frequency' = Word frequency of nouns', 'Type token ratio' = 'Type token ratio of words'

**Normative table of lexical measures – part 2**

| | | | Male | | | | | | Female | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NWT1 | | | NWT2 | | | NWT1 | | | NWT2 | | |
| | | | M | SD | Range | M | SD | Range | M | SD | Range | M | SD | range |
| Baseline | 20 - 35 | Adjectives | 75.53 | 12.49 | 57.59-97.17 | 44.72 | 10.6 | 33.56-66.20 | 86.42 | 22.35 | 29.41-125.93 | 49.31 | 15.57 | 18.52-76.57 |
| | | Nouns | 161.94 | 38.92 | 117.24-236.11 | 195.12 | 25.1 | 151.22-220.00 | 172.97 | 25.5 | 117.28-222.80 | 176.13 | 34.39 | 125.86-259.26 |
| | | Verbs (v) | 178.14 | 28.12 | 138.97-235.60 | 182.32 | 12.33 | 156.00-199.19 | 189.83 | 11.81 | 155.69-206.70 | 188.5 | 23.64 | 146.67-231.48 |
| | | Concrete v | 20.24 | 5.02 | 11.43-26.46 | 16.53 | 7.4 | 0.00-26.32 | 17.52 | 12.49 | 0.00-52.63 | 17.56 | 9.11 | 4.52-36.36 |
| | | Abstract v | 110.09 | 21.07 | 72.87-136.20 | 114.33 | 10.96 | 94.34-131.87 | 118.01 | 20.46 | 73.86-143.54 | 118.93 | 26.2 | 73.31-177.36 |
| Target | 50 - 59 | Adjectives | 72.73 | 25.24 | 32.26-120.57 | 48.69 | 13.73 | 29.85-80.36 | 72.46 | 22.29 | 23.26-129.41 | 53.82 | 17.54 | 13.70-90.91 |
| | | Nouns | 178.57 | 29.09 | 137.93-242.86 | 202.25 | 30.9 | 147.93-258.07 | 180.69 | 28.87 | 125.00-240.74 | 183.76 | 32.71 | 142.86-269.23 |
| | | Verbs (v) | 181.82 | 26.68 | 123.19-226.67 | 189.57 | 23.16 | 145.46-237.11 | 189.98 | 24.32 | 146.67-248.28 | 196.64 | 29.96 | 136.84-263.74 |
| | | Concrete v | 14.29 | 12.51 | 0.00-41.42 | 20.83 | 11.41 | 0.00-44.78 | 15.09 | 13.05 | 0.00-45.45 | 13.73 | 13.18 | 0.00-49.65 |
| | | Abstract v | 98.36 | 30.18 | 50.72-151.52 | 119.26 | 22.71 | 78.01-163.12 | 111.5 | 22.61 | 69.77-164.89 | 127.56 | 21.12 | 83.33-173.91 |
| | 60 - 69 | Adjectives | 60.92 | 19.52 | 33.33-113.21 | 58.28 | 22.53 | 0.00-103.09 | 73.17 | 28.15 | 8.33-137.93 | 54.79 | 23.56 | 18.87-117.65 |
| | | Nouns | 188.09 | 34.02 | 125.00-269.23 | 199.01 | 33.77 | 142.86-270.83 | 180.33 | 31.03 | 107.14-256.88 | 172.04 | 33.7 | 117.65-258.93 |
| | | Verbs (v) | 187.92 | 30.83 | 120.00-228.26 | 192.31 | 28.75 | 117.12-244.44 | 190.35 | 22.56 | 139.18-243.90 | 196.24 | 32.29 | 128.21-272.73 |
| | | Concrete v | 17.26 | 11.73 | 0.00-40.00 | 14.39 | 11.44 | 0.00-46.51 | 15.15 | 14.36 | 0.00-48.78 | 15.87 | 12.84 | 0.00-46.51 |
| | | Abstract v | 110.96 | 34.48 | 31.25-173.08 | 118.18 | 30.99 | 46.88-188.57 | 125 | 26.65 | 57.14-173.91 | 125.75 | 34.42 | 61.67-205.88 |
| | 70 - 79 | Adjectives | 66.88 | 29.13 | 0.00-128.08 | 51.09 | 19.54 | 14.93-105.26 | 72.58 | 35.8 | 0.00-142.86 | 60 | 22.91 | 14.08-100.00 |
| | | Nouns | 185.69 | 36.74 | 127.45-256.41 | 206.51 | 35 | 131.87-263.16 | 169.78 | 33.69 | 104.84-250.00 | 185.19 | 32.84 | 137.62-254.72 |
| | | Verbs (v) | 186.22 | 39.91 | 118.64-270.83 | 184.96 | 24.49 | 126.32-214.29 | 194.15 | 26.76 | 145.70-246.64 | 183.75 | 29.62 | 120.00-235.29 |
| | | Concrete v | 11.46 | 12.02 | 0.00-47.17 | 14.41 | 13.02 | 0.00-46.88 | 24.19 | 16.98 | 0.00-59.70 | 15.63 | 13.5 | 0.00-45.45 |
| | | Abstract v | 117.24 | 38.51 | 33.33-196.97 | 123.72 | 31.01 | 48.78-166.67 | 109.49 | 37.94 | 25.64-171.72 | 117.65 | 40.52 | 32.26-185.19 |
| | 80+ | Adjectives | 74.34 | 9.13 | 58.82-81.97 | 47.06 | 19.64 | 17.24-78.95 | 88.06 | 37.94 | 28.99-142.86 | 65.49 | 25.46 | 19.61-109.09 |
| | | Nouns | 194.69 | 51.38 | 104.17-250.00 | 166.56 | 28.95 | 149.43-235.29 | 184.85 | 50.88 | 100.67-238.10 | 191.22 | 56.87 | 97.40-254.90 |
| | | Verbs (v) | 171.43 | 18.34 | 133.66-190.00 | 179.97 | 22.42 | 152.94-215.69 | 189.81 | 52.05 | 142.86-304.35 | 200.92 | 36.87 | 137.26-241.38 |
| | | Concrete v | 17.7 | 8.5 | 0.00-25.00 | 10.74 | 10.37 | 0.00-30.00 | 25.13 | 18.77 | 0.00-50.00 | 9.85 | 10.33 | 0.00-25.97 |
| | | Abstract v | 113.64 | 18.5 | 71.43-130.00 | 105.82 | 19.49 | 80.00-142.86 | 101.31 | 52.73 | 42.55-202.90 | 106.06 | 30.05 | 58.82-161.07 |

*Note*: The density of the various variables is displayed.

## Normative table of grammatical measures

| | | | Male | | | | | | Female | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NT1 | | | NT2 | | | NT1 | | | NT2 | | |
| | | | M | SD | Range | M | SD | Range | M | SD | Range | M | SD | Range |
| Baseline | 20 - 35 | D-level | 2.74 | 1.01 | 0.20-3.83 | 2.33 | 0.84 | 1.20-3.55 | 2.18 | 0.81 | 0.74-4.00 | 2.38 | 0.84 | 0.95-4.29 |
| | | Distance 1 | 1.76 | 0.78 | 0.62-3.04 | 2.48 | 0.60 | 1.70-3.81 | 1.87 | 0.73 | 0.94-3.83 | 1.97 | 0.48 | 0.97-2.97 |
| | | Distance 2 | 0.18 | 0.09 | 0.05-0.33 | 0.25 | 0.10 | 0.11-0.45 | 0.21 | 0.14 | 0.00-0.63 | 0.24 | 0.14 | 0.00-0.56 |
| | | Density pers | 132.93 | 27.21 | 91.74-176.03 | 73.17 | 18.66 | 53.69-114.50 | 129.11 | 35.85 | 74.07-216.05 | 73.50 | 17.59 | 42.86-113.21 |
| | | Density conj | 48.28 | 20.20 | 9.17-74.29 | 124.54 | 34.40 | 76.00-190.84 | 56.10 | 17.24 | 20.73-94.74 | 133.33 | 25.29 | 67.71-193.33 |
| Target | 50 - 59 | D-level | 2.25 | 1.03 | 0.81-4.63 | 2.43 | 1.24 | 0.94-5.00 | 1.91 | 0.85 | 0.50-3.71 | 2.45 | 0.87 | 0.54-4.00 |
| | | Distance 1 | 2.00 | 0.65 | 1.13-3.15 | 2.21 | 0.43 | 1.25-2.66 | 1.75 | 0.68 | 0.33-3.17 | 1.83 | 0.80 | 0.86-3.86 |
| | | Distance 2 | 0.22 | 0.12 | 0.00-0.43 | 0.21 | 0.12 | 0.00-0.45 | 0.22 | 0.11 | 0.00-0.52 | 0.19 | 0.13 | 0.00-0.50 |
| | | Density pers | 125.00 | 38.75 | 65.22-186.67 | 56.74 | 24.50 | 0.00-97.90 | 139.07 | 28.80 | 61.73-191.92 | 63.26 | 20.21 | 15.87-109.89 |
| | | Density conj | 51.99 | 16.47 | 16.39-75.47 | 112.43 | 29.18 | 62.50-167.83 | 46.98 | 16.82 | 19.61-81.82 | 126.23 | 30.35 | 47.62-178.95 |
| | 60 - 69 | D-level | 2.92 | 1.98 | 0.67-7.00 | 2.67 | 1.02 | 0.83-5.00 | 2.56 | 1.10 | 0.67-5.75 | 2.62 | 0.66 | 1.13-4.00 |
| | | Distance 1 | 2.00 | 1.01 | 0.11-3.87 | 2.00 | 0.87 | 0.81-4.38 | 1.78 | 0.69 | 0.53-3.57 | 1.84 | 0.58 | 0.92-3.38 |
| | | Distance 2 | 0.15 | 0.15 | 0.00-0.57 | 0.22 | 0.16 | 0.05-0.64 | 0.17 | 0.13 | 0.00-0.50 | 0.21 | 0.16 | 0.00-0.55 |
| | | Density pers | 114.29 | 47.30 | 31.25-230.77 | 63.25 | 26.58 | 24.10-145.63 | 116.67 | 35.30 | 43.48-205.67 | 67.46 | 25.22 | 18.87-122.30 |
| | | Density conj | 46.45 | 13.90 | 25.32-76.34 | 115.04 | 29.29 | 61.35-170.73 | 55.79 | 21.48 | 9.26-108.70 | 123.66 | 26.35 | 68.38-181.10 |
| | 70 - 79 | D-level | 2.29 | 0.91 | 0.50-3.67 | 2.63 | 1.13 | 0.50-5.33 | 2.50 | 1.26 | 0.00-4.50 | 2.34 | 1.25 | 0.24-4.80 |
| | | Distance 1 | 1.93 | 0.77 | 0.40-3.36 | 1.77 | 0.94 | 0.71-4.27 | 1.65 | 0.81 | 0.00-2.67 | 1.57 | 0.74 | 0.13-3.21 |
| | | Distance 2 | 0.18 | 0.12 | 0.00-0.46 | 0.19 | 0.12 | 0.00-0.44 | 0.20 | 0.16 | 0.00-0.55 | 0.23 | 0.14 | 0.00-0.50 |
| | | Density pers | 124.19 | 43.38 | 16.39-191.08 | 75.85 | 26.86 | 24.39-125.00 | 128.21 | 38.21 | 65.22-201.61 | 81.19 | 24.63 | 22.22-126.26 |
| | | Density conj | 50.07 | 14.97 | 25.00-84.91 | 118.62 | 43.76 | 52.63-240.88 | 55.56 | 30.59 | 0.00-130.44 | 120.00 | 31.71 | 65.57-171.43 |
| | 80+ | D-level | 1.99 | 1.24 | 0.40-4.50 | 2.00 | 0.70 | 0.50-2.83 | 2.00 | 0.84 | 0.67-3.33 | 2.93 | 1.64 | 0.67-5.71 |
| | | Distance 1 | 2.00 | 1.04 | 0.14-3.12 | 1.64 | 0.94 | 0.00-2.80 | 1.63 | 0.57 | 0.82-2.67 | 1.80 | 0.93 | 0.22-3.13 |
| | | Distance 2 | 0.10 | 0.12 | 0.00-0.36 | 0.18 | 0.22 | 0.00-0.64 | 0.21 | 0.21 | 0.07-0.71 | 0.09 | 0.12 | 0.00-0.33 |
| | | Density pers | 125.00 | 49.04 | 21.98-206.35 | 62.86 | 12.76 | 39.22-79.14 | 103.43 | 23.93 | 80.54-148.15 | 55.54 | 26.75 | 18.18-98.04 |
| | | Density conj | 45.45 | 19.40 | 14.29-75.00 | 122.65 | 10.91 | 103.45-131.58 | 60.90 | 19.06 | 37.04-93.96 | 111.47 | 38.21 | 47.62-167.79 |

*Note*: 'Distance 1' = Distance subject-verb', 'Distance 2' = 'Distance determiner-noun',
'Density pers'= ' Density of personal references', 'Density conj' = 'Density of cojunctions'

# APPENDIX D

**Process Variables Used In The Analysis Of The Data (CHAPTER 4)**

| Variables | Definition |
|---|---|
| **Factor 1: Fluency of utterance total document** | |
| Mean length of utterance | The average number of characters per utterance |
| Keystrokes (incl. spaces) per minute | The number of keystrokes produced per minute including spaces, excluding corrections |
| Words written per minute during the writing process | The average number of words written per minute during the writing process |
| **Factor 2: Fluency of utterance at the start** | |
| Number of keystrokes per minute in interval 1 (the start) | The number of keystrokes produced per minute at the beginning of the text |
| Number of utterances per minute | The average number of utterances produced per minute |
| **Factor 3: Correctness of words** | |
| Ratio keystrokes product/process | This variable provides insight into the revising behaviour of the participants. It represents the number of keystrokes that can be found in the final text versus the total number of keystrokes that were typed |
| Mean pause time (in s) within words | The average pause time required to type a word |

*Note*: Utterances are in writing research known as p-bursts.

## APPENDIX E

**Linear mixed models - Effects of age, sex, task and moment on the product variables (Chapter 4)**

| Model | Effects |
|---|---|
| Model 0 | Random variance |
| Model 1 | Effect of age |
| Model 2 | Effect of age and sex |
| Model 3 | Effect of age, sex and task |
| Model 4 | Effect of age, sex, task and moment |
| Model 5 | Effect of task, moment and interaction effect between age and sex |
| Model 6 | Effect of moment and the interaction effect between age, sex and task |
| Model 7 | Interaction effect between age, sex, task and moment |

*Lexical*

| Cohesion | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 1.043.923 | 3 | | | |
| Model 1 | 1.042.333 | 4 | 1.590 | 1 | 0.207 |
| Model 2 | 1.042.028 | 5 | 0.305 | 1 | 0.581 |
| Model 3 | 1.041.549 | 6 | 0.479 | 1 | 0.489 |
| Model 4 | 1.040.958 | 7 | 0.591 | 1 | 0.442 |
| Model 5 | 1.039.932 | 8 | 1.026 | 1 | 0.311 |
| Model 6 | 1.036.848 | 11 | 3.084 | 3 | 0.379 |
| Model 7 | 1.029.843 | 18 | 7.006 | 7 | 0.428 |



| Total words | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.982.907 | 3 | | | |
| Model 1 | 3.981.292 | 4 | 1.615 | 1 | 0.204 |
| Model 2 | 3.981.280 | 5 | 0.011 | 1 | 0.915 |
| Model 3 | 3.980.197 | 6 | 1.083 | 1 | 0.298 |
| Model 4 | 3.977.754 | 7 | 2.443 | 1 | 0.118 |
| Model 5 | 3.977.738 | 8 | 0.017 | 1 | 0.897 |
| Model 6 | 3.974.017 | 11 | 3.720 | 3 | 0.293 |
| Model 7 | 3.972.189 | 18 | 1.829 | 7 | 0.969 |
| **Word length** | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 109.500 | 3 | | | |
| Model 1 | 109.232 | 4 | 0.268 | 1 | 0.604 |
| Model 2 | 106.242 | 5 | 2.990 | 1 | 0.084 |
| Model 3 | 7.901 | 6 | 98.341 | 1 | 0.000 |
| Model 4 | 6.003 | 7 | 1.898 | 1 | 0.168 |
| Model 5 | 5.992 | 8 | 0.011 | 1 | 0.915 |
| Model 6 | 2.889 | 11 | 3.103 | 3 | 0.376 |
| Model 7 | -1.527 | 18 | 4.416 | 7 | 0.731 |
| **Word frequency of nouns** | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 79.595 | 3 | | | |
| Model 1 | 79.135 | 4 | 0.460 | 1 | 0.497 |
| Model 2 | 72.875 | 5 | 6.260 | 1 | 0.012 |
| Model 3 | -40.362 | 6 | 113.237 | 1 | 0.000 |
| Model 4 | -40.992 | 7 | 0.630 | 1 | 0.427 |
| Model 5 | -41.157 | 8 | 0.165 | 1 | 0.684 |
| Model 6 | -49.148 | 11 | 7.991 | 3 | 0.046 |
| Model 7 | -55.199 | 18 | 6.051 | 7 | 0.534 |

| Type token ratio of words | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | -883.696 | 3 | | | |
| Model 1 | -885.785 | 4 | 2.089 | 1 | 0.148 |
| Model 2 | -885.850 | 5 | 0.066 | 1 | 0.798 |
| Model 3 | -886.797 | 6 | 0.947 | 1 | 0.331 |
| Model 4 | -888.113 | 7 | 1.316 | 1 | 0.251 |
| Model 5 | -888.658 | 8 | 0.545 | 1 | 0.460 |
| Model 6 | -890.169 | 11 | 1.511 | 3 | 0.680 |
| Model 7 | -892.185 | 18 | 2.016 | 7 | 0.959 |

| Density of adjectives | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.117.461 | 3 | | | |
| Model 1 | 3.116.923 | 4 | 0.538 | 1 | 0.463 |
| Model 2 | 3.116.247 | 5 | 0.677 | 1 | 0.411 |
| Model 3 | 3.115.654 | 6 | 0.593 | 1 | 0.441 |
| Model 4 | 3.115.126 | 7 | 0.527 | 1 | 0.468 |
| Model 5 | 3.114.932 | 8 | 0.195 | 1 | 0.659 |
| Model 6 | 3.107.706 | 11 | 7.226 | 3 | 0.065 |
| Model 7 | 3.092.880 | 18 | 14.825 | 7 | 0.038 |

| Density of nouns | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.308.812 | 3 | | | |
| Model 1 | 3.307.778 | 4 | 1.034 | 1 | 0.309 |
| Model 2 | 3.306.748 | 5 | 1.030 | 1 | 0.310 |
| Model 3 | 3.292.748 | 6 | 14.000 | 1 | 0.000 |
| Model 4 | 3.292.652 | 7 | 0.095 | 1 | 0.757 |
| Model 5 | 3.291.405 | 8 | 1.247 | 1 | 0.264 |
| Model 6 | 3.288.442 | 11 | 2.963 | 3 | 0.397 |
| Model 7 | 3.276.722 | 18 | 11.720 | 7 | 0.110 |



| Density of verbs | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.242.827 | 3 | | | |
| Model 1 | 3.240.539 | 4 | 2.288 | 1 | 0.130 |
| Model 2 | 3.239.237 | 5 | 1.301 | 1 | 0.254 |
| Model 3 | 3.238.702 | 6 | 0.535 | 1 | 0.464 |
| Model 4 | 3.238.673 | 7 | 0.029 | 1 | 0.864 |
| Model 5 | 3.238.315 | 8 | 0.357 | 1 | 0.550 |
| Model 6 | 3.236.863 | 11 | 1.453 | 3 | 0.693 |
| Model 7 | 3.231.402 | 18 | 5.461 | 7 | 0.604 |

| Density of concrete verbs | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 2.764.742 | 3 | | | |
| Model 1 | 2.764.686 | 4 | 0.056 | 1 | 0.813 |
| Model 2 | 2.764.602 | 5 | 0.084 | 1 | 0.771 |
| Model 3 | 2.763.610 | 6 | 0.992 | 1 | 0.319 |
| Model 4 | 2.763.545 | 7 | 0.065 | 1 | 0.799 |
| Model 5 | 2.763.537 | 8 | 0.008 | 1 | 0.929 |
| Model 6 | 2.760.197 | 11 | 3.341 | 3 | 0.342 |
| Model 7 | 2.757.577 | 18 | 2.620 | 7 | 0.918 |

| Density of abstract verbs | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.247.999 | 3 | | | |
| Model 1 | 3.246.396 | 4 | 1.603 | 1 | 0.205 |
| Model 2 | 3.243.966 | 5 | 2.431 | 1 | 0.119 |
| Model 3 | 3.243.945 | 6 | 0.021 | 1 | 0.886 |
| Model 4 | 3.243.888 | 7 | 0.057 | 1 | 0.811 |
| Model 5 | 3.243.823 | 8 | 0.064 | 1 | 0.800 |
| Model 6 | 3.242.051 | 11 | 1.772 | 3 | 0.621 |
| Model 7 | 3.237.389 | 18 | 4.662 | 7 | 0.701 |

*Grammatical*

| D-Level | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 943.708 | 3 | NA | NA | NA |
| Model 1 | 942.618 | 4 | 1.090 | 1 | 0.296 |
| Model 2 | 941.579 | 5 | 1.039 | 1 | 0.308 |
| Model 3 | 941.539 | 6 | 0.039 | 1 | 0.842 |
| Model 4 | 941.272 | 7 | 0.267 | 1 | 0.605 |
| Model 5 | 940.367 | 8 | 0.905 | 1 | 0.342 |
| Model 6 | 934.698 | 11 | 5.669 | 3 | 0.129 |
| Model 7 | 931.066 | 18 | 3.632 | 7 | 0.821 |

| Distance subject-verb | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 829.423 | 3 | NA | NA | NA |
| Model 1 | 828.916 | 4 | 0.507 | 1 | 0.476 |
| Model 2 | 823.287 | 5 | 5.630 | 1 | 0.018 |
| Model 3 | 816.371 | 6 | 6.915 | 1 | 0.009 |
| Model 4 | 814.106 | 7 | 2.265 | 1 | 0.132 |
| Model 5 | 814.096 | 8 | 0.010 | 1 | 0.921 |
| Model 6 | 813.727 | 11 | 0.369 | 3 | 0.947 |
| Model 7 | 799.895 | 18 | 13.832 | 7 | 0.054 |



| Distance determiner-noun | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | -286.088 | 3 | NA | NA | NA |
| Model 1 | -286.490 | 4 | 0.402 | 1 | 0.526 |
| Model 2 | -287.134 | 5 | 0.644 | 1 | 0.422 |
| Model 3 | -288.675 | 6 | 1.541 | 1 | 0.215 |
| Model 4 | -289.376 | 7 | 0.701 | 1 | 0.402 |
| Model 5 | -293.322 | 8 | 3.946 | 1 | 0.047 |
| Model 6 | -298.638 | 11 | 5.316 | 3 | 0.150 |
| Model 7 | -303.445 | 18 | 4.807 | 7 | 0.683 |

| Density of personal references | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.368.256 | 3 | NA | NA | NA |
| Model 1 | 3.367.368 | 4 | 0.889 | 1 | 0.346 |
| Model 2 | 3.359.343 | 5 | 8.025 | 1 | 0.005 |
| Model 3 | 3.359.039 | 6 | 0.304 | 1 | 0.582 |
| Model 4 | 3.359.011 | 7 | 0.028 | 1 | 0.866 |
| Model 5 | 3.357.105 | 8 | 1.906 | 1 | 0.167 |
| Model 6 | 3.354.235 | 11 | 2.869 | 3 | 0.412 |
| Model 7 | 3.353.075 | 18 | 1.160 | 7 | 0.992 |

| Density of conjunctions | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | $\chi^2_{change}$ | $df_{change}$ | *p* |
| Model 0 | 3.010.360 | 3 | NA | NA | NA |
| Model 1 | 3.010.197 | 4 | 0.163 | 1 | 0.686 |
| Model 2 | 3.009.859 | 5 | 0.338 | 1 | 0.561 |
| Model 3 | 3.007.112 | 6 | 2.747 | 1 | 0.097 |
| Model 4 | 3.006.737 | 7 | 0.375 | 1 | 0.540 |
| Model 5 | 3.003.789 | 8 | 2.947 | 1 | 0.086 |
| Model 6 | 3.002.769 | 11 | 1.020 | 3 | 0.796 |
| Model 7 | 2.995.674 | 18 | 7.096 | 7 | 0.419 |

## Linear mixed models - Product variables - Estimates of fixed effects (Chapter 4)

| Cohesion | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | -.027 | .129 | .833 | 1.391 | 1.124 | .221 | 1.410 | 1.122 | .214 | 1.444 | 1.123 | .204 | 1.351 | 1.129 | .236 | .416 | 1.449 | .775 | -.111 | 1.553 | .943 | -2.501 | 2.467 | .312 |
| Age | | | | -.023 | .018 | .209 | -.022 | .018 | .221 | -.022 | .018 | .221 | -.022 | .018 | .221 | -.007 | .023 | .762 | .002 | .025 | .946 | .039 | .039 | .327 |
| Sex | | | | | | | -.144 | .261 | .582 | -.144 | .261 | .582 | -.144 | .261 | .582 | 2.149 | 2.269 | .348 | 1.691 | 2.435 | .489 | 6.766 | 3.880 | .082 |
| Task | | | | | | | | | | -.069 | .099 | .489 | -.069 | .099 | .489 | -.069 | .099 | .489 | .984 | 1.124 | .382 | 1.946 | 2.937 | .508 |
| Moment | | | | | | | | | | | | | .047 | .061 | .442 | .047 | .061 | .442 | .047 | .060 | .440 | 1.242 | .961 | .198 |
| Age*sex | | | | | | | | | | | | | | | | -.036 | .036 | .313 | -.030 | .038 | .438 | -.114 | .061 | .063 |
| Age*task | | | | | | | | | | | | | | | | | | | -.017 | .018 | .333 | -.028 | .047 | .547 |
| Sex*task | | | | | | | | | | | | | | | | | | | .915 | 1.767 | .605 | -3.730 | 4.620 | .420 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.013 | .028 | .642 | -.018 | .015 | .229 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -2.537 | 1.512 | .094 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -.481 | 1.360 | .724 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .064 | .073 | .381 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .042 | .024 | .077 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .005 | .022 | .802 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 2.323 | 2.138 | .278 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.038 | .034 | .255 |

**Total words**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 204.968 | 9.399 | .000 | 308.709 | 81.581 | .000 | 308.440 | 81.612 | .000 | 311.908 | 81.680 | .000 | 324.632 | 82.082 | .000 | 333.345 | 106.248 | .003 | 354.636 | 112.669 | .002 | 452.029 | 171.827 | .009 |
| Age | | | | -1.651 | 1.290 | .206 | -1.659 | 1.292 | .204 | -1.659 | 1.292 | .204 | -1.659 | 1.292 | .204 | -1.799 | 1.684 | .290 | -2.172 | 1.787 | .228 | -3.734 | 2.732 | .173 |
| Sex | | | | | | | 2.023 | 18.982 | .915 | 2.023 | 18.982 | .915 | 2.023 | 18.982 | .915 | -19.340 | 166.532 | .908 | -112.936 | 176.747 | .525 | -73.001 | 27.244 | .787 |
| Task | | | | | | | | | | -6.937 | 6.661 | .299 | -6.937 | 6.633 | .296 | -6.937 | 6.633 | .296 | -49.519 | 75.302 | .511 | -217.762 | 198.604 | .274 |
| Moment | | | | | | | | | | | | | -6.362 | 4.062 | .118 | -6.362 | 4.062 | .118 | -6.362 | 4.036 | .116 | -55.059 | 65.008 | .398 |
| Age*sex | | | | | | | | | | | | | | | | .339 | 2.625 | .898 | 1.903 | 2.786 | .497 | 1.269 | 4.260 | .766 |
| Age*task | | | | | | | | | | | | | | | | | | | .747 | 1.197 | .533 | 3.392 | 3.158 | .284 |
| Sex*task | | | | | | | | | | | | | | | | | | | 187.193 | 118.432 | .115 | 226.619 | 312.356 | .469 |
| Age*moment | | | | | | | | | | | | | | | | | | | -3.128 | 1.867 | .095 | .781 | 1.034 | .451 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -19.968 | 102.243 | .845 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 84.122 | 91.936 | .361 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -3.603 | 4.924 | .465 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .317 | 1.612 | .844 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -1.323 | 1.462 | .366 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -19.713 | 144.593 | .892 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .238 | 2.280 | .917 |

**Word length**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 4.417 | .024 | .000 | 4.307 | .214 | .000 | 4.296 | .209 | .000 | 4.170 | .209 | .000 | 4.209 | .211 | .000 | 4.227 | .273 | .000 | 4.189 | .303 | .000 | 4.366 | .545 | .000 |
| Age | | | | .002 | .003 | .606 | .001 | .003 | .674 | .001 | .003 | .674 | .001 | .003 | .674 | .001 | .004 | .798 | .002 | .005 | .734 | -.002 | .009 | .810 |
| Sex | | | | | | | .085 | .049 | .085 | .085 | .049 | .085 | .085 | .049 | .085 | .040 | .427 | .925 | -.181 | .474 | .704 | -.342 | .857 | .690 |
| Task | | | | | | | | | | .252 | .023 | .000 | .252 | .023 | .000 | .252 | .023 | .000 | .328 | .264 | .215 | .395 | .693 | .569 |
| Moment | | | | | | | | | | | | | -.020 | .014 | .169 | -.020 | .014 | .169 | -.020 | .014 | .166 | -.108 | .227 | .634 |
| Age*sex | | | | | | | | | | | | | | | | .001 | .007 | .916 | .004 | .007 | .559 | .009 | .014 | .495 |
| Age*task | | | | | | | | | | | | | | | | | | | -.001 | .004 | .801 | -.002 | .011 | .878 |
| Sex*task | | | | | | | | | | | | | | | | | | | .442 | .415 | .288 | .371 | 1.090 | .734 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.007 | .007 | .263 | .002 | .004 | .606 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | .080 | .357 | .822 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -.034 | .321 | .917 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -.007 | .017 | .667 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.002 | .006 | .667 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .000 | .005 | .951 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .036 | .505 | .944 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .000 | .008 | .996 |

**Word frequency of nouns**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 4.680 | .025 | .000 | 4.825 | .215 | .000 | 4.842 | .204 | .000 | 4.968 | .204 | .000 | 4.948 | .206 | .000 | 4.879 | .266 | .000 | 5.000 | .292 | .000 | 4.361 | .507 | .000 |
| Age | | | | -.002 | .003 | .499 | -.002 | .003 | .579 | -.002 | .003 | .579 | -.002 | .003 | .579 | -.001 | .004 | .867 | -.003 | .005 | .572 | .008 | .008 | .346 |
| Sex | | | | | | | -.122 | .047 | .013 | -.122 | .047 | .013 | -.122 | .047 | .013 | .046 | .416 | .912 | .311 | .458 | .498 | .454 | .798 | .569 |
| Task | | | | | | | | | | -.254 | .022 | .000 | -.254 | .022 | .000 | -.254 | .022 | .000 | -.496 | .243 | .042 | -.391 | .635 | .538 |
| Moment | | | | | | | | | | | | | .010 | .013 | .428 | .010 | .013 | .428 | .010 | .013 | .421 | .330 | .208 | .114 |
| Age*sex | | | | | | | | | | | | | | | | -.003 | .007 | .686 | -.007 | .007 | .342 | -.009 | .013 | .468 |
| Age*task | | | | | | | | | | | | | | | | | | | .004 | .004 | .321 | .002 | .010 | .806 |
| Sex*task | | | | | | | | | | | | | | | | | | | -.531 | .381 | .165 | -.047 | .999 | .963 |
| Age*moment | | | | | | | | | | | | | | | | | | | .008 | .006 | .162 | -.005 | .003 | .123 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -.071 | .327 | .827 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -.053 | .294 | .858 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -.000 | .016 | .982 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .001 | .005 | .828 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .001 | .005 | .885 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.242 | .462 | .601 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .004 | .007 | .546 |

**Type token ratio of words**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | .604 | .008 | .000 | .498 | .073 | .000 | .499 | .073 | .000 | .496 | .073 | .000 | .504 | .074 | .000 | .460 | .095 | .000 | .451 | .101 | .000 | .476 | .157 | .003 |
| Age | | | | .002 | .001 | .150 | .002 | .001 | .146 | .002 | .001 | .146 | .002 | .001 | .146 | .002 | .002 | .113 | .003 | .002 | .111 | .002 | .002 | .393 |
| Sex | | | | | | | -.004 | .017 | .799 | -.004 | .017 | .799 | -.004 | .017 | .799 | .105 | .149 | .483 | .155 | .158 | .331 | .055 | .247 | .824 |
| Task | | | | | | | | | | .006 | .006 | .331 | .006 | .006 | .330 | .006 | .006 | .330 | .024 | .070 | .733 | .076 | .184 | .682 |
| Moment | | | | | | | | | | | | | -.004 | .004 | .252 | -.004 | .004 | .252 | -.004 | .004 | .250 | -.017 | .060 | .781 |
| Age*sex | | | | | | | | | | | | | | | | -.002 | .002 | .462 | -.003 | .002 | .304 | -.001 | .004 | .825 |
| Age*task | | | | | | | | | | | | | | | | | | | -.000 | .001 | .764 | -.001 | .003 | .739 |
| Sex*task | | | | | | | | | | | | | | | | | | | -.101 | .110 | .361 | -.081 | .290 | .781 |
| Age*moment | | | | | | | | | | | | | | | | | | | .002 | .002 | .325 | .000 | .001 | .817 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | .050 | .095 | .598 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -.026 | .085 | .762 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .001 | .005 | .844 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.001 | .001 | .565 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .000 | .001 | .812 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.010 | .134 | .941 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .000 | .002 | .849 |

**Density of adjectives**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 72.086 | 1.567 | .000 | 62.058 | 13.730 | .000 | 62.406 | 13.656 | .000 | 63.232 | 13.698 | .000 | 65.139 | 13.947 | .000 | 6.160 | 17.920 | .001 | 79.628 | 21.585 | .000 | 99.829 | 46.035 | .031 |
| Age | | | | .160 | .217 | .465 | .171 | .216 | .433 | .171 | .216 | .433 | .171 | .216 | .433 | .250 | .281 | .377 | -.044 | .341 | .898 | -.500 | .732 | .495 |
| Sex | | | | | | | -2.620 | 3.176 | .413 | -2.620 | 3.176 | .413 | -2.620 | 3.176 | .413 | 9.588 | 27.829 | .732 | -38.342 | 33.702 | .258 | -86.576 | 72.402 | .233 |
| Task | | | | | | | | | | -1.653 | 2.145 | .442 | -1.653 | 2.143 | .441 | -1.653 | 2.143 | .441 | -4.588 | 24.185 | .094 | -111.388 | 62.372 | .075 |
| Moment | | | | | | | | | | | | | -.954 | 1.312 | .468 | -.954 | 1.312 | .468 | -.953 | 1.296 | .463 | -11.054 | 2.416 | .589 |
| Age*sex | | | | | | | | | | | | | | | | -.194 | .439 | .660 | .523 | .531 | .326 | 1.515 | 1.141 | .185 |
| Age*task | | | | | | | | | | | | | | | | | | | .588 | .385 | .127 | 1.910 | .992 | .055 |
| Sex*task | | | | | | | | | | | | | | | | | | | 95.859 | 38.037 | .012 | 61.418 | 98.097 | .532 |
| Age*moment | | | | | | | | | | | | | | | | | | | -1.434 | .600 | .017 | .228 | .325 | .483 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 24.117 | 32.110 | .453 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 35.400 | 28.873 | .221 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -1.166 | 1.547 | .452 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.496 | .506 | .328 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.661 | .459 | .151 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 17.221 | 45.410 | .705 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.134 | .716 | .851 |

**Density of nouns**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 179.948 | 2.227 | .000 | 199.660 | 19.427 | .000 | 199.053 | 19.265 | .000 | 193.908 | 19.313 | .000 | 192.880 | 19.597 | .000 | 21.576 | 24.990 | .000 | 21.918 | 29.344 | .000 | 186.109 | 59.911 | .002 |
| Age | | | | -.314 | .307 | .311 | -.333 | .305 | .280 | -.333 | .305 | .280 | -.333 | .305 | .280 | -.616 | .393 | .123 | -.652 | .464 | .163 | -.177 | .953 | .853 |
| Sex | | | | | | | 4.569 | 4.481 | .312 | 4.569 | 4.481 | .312 | 4.569 | 4.481 | .312 | -38.821 | 38.895 | .322 | -38.910 | 45.856 | .398 | -7.393 | 94.226 | .456 |
| Task | | | | | | | | | | 1.292 | 2.717 | .000 | 1.292 | 2.717 | .000 | 1.292 | 2.717 | .000 | 9.608 | 3.887 | .756 | 49.037 | 8.085 | .541 |
| Moment | | | | | | | | | | | | | .514 | 1.664 | .758 | .514 | 1.664 | .758 | .514 | 1.655 | .756 | 12.919 | 26.214 | .623 |
| Age*sex | | | | | | | | | | | | | | | | .689 | .613 | .266 | .765 | .723 | .292 | 1.311 | 1.485 | .378 |
| Age*task | | | | | | | | | | | | | | | | | | | .071 | .491 | .885 | -.888 | 1.274 | .486 |
| Sex*task | | | | | | | | | | | | | | | | | | | .178 | 48.578 | .997 | 19.039 | 125.955 | .880 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.153 | .766 | .842 | -.237 | .417 | .570 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 15.741 | 41.229 | .703 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -19.715 | 37.072 | .595 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -.118 | 1.986 | .953 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.273 | .650 | .675 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .480 | .590 | .417 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -9.431 | 58.306 | .872 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.018 | .919 | .985 |

**Density of verbs**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 191.371 | 1.906 | .000 | 166.400 | 16.452 | .000 | 166.977 | 16.277 | .000 | 167.914 | 16.327 | .000 | 168.452 | 16.625 | .000 | 176.486 | 21.326 | .000 | 165.101 | 25.808 | .000 | 194.626 | 56.210 | .001 |
| Age | | | | .397 | .260 | .132 | .416 | .258 | .112 | .416 | .258 | .112 | .416 | .258 | .112 | .287 | .335 | .395 | .456 | .407 | .266 | -.062 | .894 | .945 |
| Sex | | | | | | | -4.343 | 3.786 | .256 | -4.343 | 3.786 | .256 | -4.343 | 3.786 | .256 | -24.042 | 33.116 | .471 | -2.161 | 4.290 | .618 | -65.480 | 88.405 | .459 |
| Task | | | | | | | | | | -1.874 | 2.561 | .465 | -1.874 | 2.560 | .465 | -1.874 | 2.560 | .465 | 2.895 | 29.183 | .475 | -45.720 | 76.488 | .550 |
| Moment | | | | | | | | | | | | | -.269 | 1.568 | .864 | -.269 | 1.568 | .864 | -.269 | 1.564 | .864 | -15.031 | 25.036 | .549 |
| Age*sex | | | | | | | | | | | | | | | | .313 | .522 | .552 | .283 | .635 | .657 | 1.028 | 1.394 | .461 |
| Age*task | | | | | | | | | | | | | | | | | | | -.337 | .464 | .468 | .917 | 1.216 | .452 |
| Sex*task | | | | | | | | | | | | | | | | | | | -7.762 | 45.898 | .866 | 81.467 | 12.297 | .499 |
| Age*moment | | | | | | | | | | | | | | | | | | | .059 | .724 | .935 | .259 | .398 | .516 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 22.660 | 39.376 | .565 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 33.308 | 35.407 | .348 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -1.656 | 1.896 | .383 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.372 | .621 | .549 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.627 | .563 | .266 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -44.615 | 55.687 | .424 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .857 | .878 | .330 |

**Density of concrete verbs**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 17.010 | .810 | .000 | 15.341 | 7.126 | .036 | 15.277 | 7.124 | .036 | 15.941 | 7.155 | .030 | 16.356 | 7.339 | .029 | 16.880 | 9.410 | .078 | 18.610 | 12.053 | .125 | 11.270 | 28.676 | .695 |
| Age | | | | .027 | .113 | .814 | .025 | .113 | .828 | .025 | .113 | .828 | .025 | .113 | .828 | .016 | .147 | .913 | .002 | .190 | .993 | .096 | .456 | .833 |
| Sex | | | | | | | .482 | 1.657 | .772 | .482 | 1.657 | .772 | .482 | 1.657 | .772 | -.804 | 14.538 | .956 | -12.829 | 18.784 | .496 | -36.437 | 45.101 | .420 |
| Task | | | | | | | | | | -1.327 | 1.332 | .320 | -1.327 | 1.331 | .320 | -1.327 | 1.331 | .320 | -4.786 | 15.126 | .752 | 27.636 | 39.838 | .488 |
| Moment | | | | | | | | | | | | | -.208 | .815 | .799 | -.208 | .815 | .799 | -.208 | .811 | .798 | 3.462 | 13.040 | .791 |
| Age*sex | | | | | | | | | | | | | | | | .020 | .229 | .929 | .178 | .296 | .550 | .571 | .711 | .422 |
| Age*task | | | | | | | | | | | | | | | | | | | .029 | .241 | .904 | -.440 | .634 | .488 |
| Sex*task | | | | | | | | | | | | | | | | | | | 24.051 | 23.789 | .313 | 7.545 | 62.656 | .904 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.314 | .375 | .403 | -.047 | .207 | .820 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 11.804 | 2.509 | .565 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -16.211 | 18.442 | .380 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -.103 | .988 | .917 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.197 | .323 | .543 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .234 | .293 | .425 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 8.253 | 29.004 | .776 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.106 | .457 | .817 |

**Density of abstract verbs**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 123.735 | 1.693 | .000 | 105.114 | 14.701 | .000 | 105.815 | 14.403 | .000 | 106.005 | 14.463 | .000 | 105.231 | 14.822 | .000 | 102.208 | 19.004 | .000 | 83.534 | 24.223 | .001 | 93.012 | 57.078 | .104 |
| Age | | | | .296 | .232 | .207 | .318 | .228 | .168 | .318 | .228 | .168 | .318 | .228 | .168 | .367 | .297 | .222 | .659 | .382 | .087 | .488 | .908 | .591 |
| Sex | | | | | | | -5.278 | 3.350 | .121 | -5.278 | 3.350 | .121 | -5.278 | 3.350 | .121 | 2.135 | 29.377 | .942 | 15.124 | 37.756 | .689 | -2.884 | 89.770 | .816 |
| Task | | | | | | | | | | -.380 | 2.648 | .886 | -.380 | 2.648 | .886 | -.380 | 2.648 | .886 | 36.967 | 3.159 | .221 | -61.826 | 79.155 | .435 |
| Moment | | | | | | | | | | | | | .387 | 1.621 | .811 | .387 | 1.621 | .811 | .387 | 1.616 | .811 | -4.352 | 25.910 | .867 |
| Age*sex | | | | | | | | | | | | | | | | -.118 | .463 | .800 | -.309 | .595 | .605 | .355 | 1.415 | .802 |
| Age*task | | | | | | | | | | | | | | | | | | | -.583 | .480 | .225 | 1.044 | 1.259 | .408 |
| Sex*task | | | | | | | | | | | | | | | | | | | -25.978 | 47.433 | .584 | 113.463 | 124.492 | .363 |
| Age*moment | | | | | | | | | | | | | | | | | | | .382 | .748 | .609 | .085 | .412 | .837 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 18.004 | 4.750 | .659 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 49.396 | 36.642 | .179 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -2.050 | 1.963 | .297 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.332 | .642 | .606 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.814 | .583 | .164 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -69.721 | 57.629 | .227 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 1.216 | .909 | .182 |

**D-Level**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 2.636 | .100 | .000 | 1.727 | .872 | .053 | 1.699 | .865 | .054 | 1.708 | .866 | .053 | 1.653 | .873 | .063 | .975 | 1.119 | .387 | .439 | 1.224 | .720 | -1.161 | 2.098 | .580 |
| Age | | | | .014 | .014 | .299 | .014 | .014 | .325 | .014 | .014 | .325 | .014 | .014 | .325 | .024 | .018 | .173 | .032 | .019 | .105 | .060 | .033 | .073 |
| Sex | | | | | | | .206 | .201 | .310 | .206 | .201 | .310 | .206 | .201 | .310 | 1.867 | 1.751 | .291 | 2.671 | 1.917 | .167 | 4.862 | 3.299 | .142 |
| Task | | | | | | | | | | -.017 | .088 | .843 | -.017 | .088 | .843 | -.017 | .088 | .843 | 1.054 | .992 | .289 | 2.108 | 2.609 | .420 |
| Moment | | | | | | | | | | | | | .028 | .054 | .606 | .028 | .054 | .606 | .028 | .053 | .602 | .828 | .854 | .333 |
| Age*sex | | | | | | | | | | | | | | | | -.026 | .028 | .344 | -.036 | .030 | .236 | -.072 | .052 | .170 |
| Age*task | | | | | | | | | | | | | | | | | | | -.015 | .016 | .353 | -.037 | .041 | .376 |
| Sex*task | | | | | | | | | | | | | | | | | | | -1.607 | 1.561 | .304 | -4.504 | 4.103 | .273 |
| Age*moment | | | | | | | | | | | | | | | | | | | .019 | .025 | .431 | -.014 | .014 | .298 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -1.096 | 1.343 | .415 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -.527 | 1.208 | .663 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .066 | .065 | .307 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .018 | .021 | .403 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .011 | .019 | .565 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 1.449 | 1.899 | .446 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.023 | .030 | .436 |

**Distance subject-verb**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 2.173 | .065 | .000 | 2.577 | .570 | .000 | 2.537 | .543 | .000 | 2.434 | .545 | .000 | 2.291 | .553 | .000 | 2.336 | .712 | .002 | 2.381 | .837 | .005 | 1.952 | 1.706 | .253 |
| Age | | | | -.006 | .009 | .478 | -.008 | .009 | .373 | -.008 | .009 | .373 | -.008 | .009 | .373 | -.008 | .011 | .455 | -.009 | .013 | .476 | -.005 | .027 | .861 |
| Sex | | | | | | | .307 | .126 | .018 | .307 | .126 | .018 | .307 | .126 | .018 | .197 | 1.109 | .859 | .154 | 1.308 | .906 | -3.870 | 2.683 | .150 |
| Task | | | | | | | | | | .205 | .078 | .009 | .205 | .077 | .008 | .205 | .077 | .008 | .114 | .883 | .897 | -.893 | 2.280 | .696 |
| Moment | | | | | | | | | | | | | .071 | .047 | .133 | .071 | .047 | .133 | .071 | .047 | .132 | .286 | .746 | .702 |
| Age*sex | | | | | | | | | | | | | | | | .002 | .017 | .921 | .003 | .021 | .878 | .075 | .042 | .077 |
| Age*task | | | | | | | | | | | | | | | | | | | .002 | .014 | .884 | .019 | .036 | .608 |
| Sex*task | | | | | | | | | | | | | | | | | | | .086 | 1.388 | .951 | 3.560 | 3.586 | .322 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.003 | .022 | .896 | -.002 | .012 | .843 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 2.012 | 1.174 | .088 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | .504 | 1.055 | .634 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -.065 | .057 | .249 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.036 | .019 | .053 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.008 | .017 | .621 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -1.737 | 1.660 | .296 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .031 | .026 | .234 |

**Distance determiner-noun**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | .247 | .012 | .000 | .313 | .106 | .004 | .311 | .105 | .005 | .301 | .105 | .006 | .284 | .107 | .010 | .114 | .134 | .395 | .134 | .161 | .408 | .234 | .350 | .505 |
| Age | | | | -.001 | .002 | .528 | -.001 | .002 | .495 | -.001 | .002 | .495 | -.001 | .002 | .495 | .002 | .002 | .456 | .001 | .003 | .624 | -.000 | .006 | .976 |
| Sex | | | | | | | .020 | .024 | .424 | .020 | .024 | .424 | .020 | .024 | .424 | .436 | .207 | .040 | .167 | .252 | .509 | .273 | .550 | .620 |
| Task | | | | | | | | | | .020 | .016 | .215 | .020 | .016 | .214 | .020 | .016 | .214 | -.019 | .181 | .918 | -.192 | .476 | .687 |
| Moment | | | | | | | | | | | | | .008 | .010 | .403 | .008 | .010 | .403 | .008 | .010 | .399 | -.042 | .156 | .789 |
| Age*sex | | | | | | | | | | | | | | | | -.007 | .003 | .048 | -.002 | .004 | .561 | -.004 | .009 | .607 |
| Age*task | | | | | | | | | | | | | | | | | | | .001 | .003 | .822 | .003 | .008 | .654 |
| Sex*task | | | | | | | | | | | | | | | | | | | .539 | .285 | .060 | -.244 | .748 | .744 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.009 | .004 | .057 | .001 | .002 | .775 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -.053 | .245 | .828 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | .087 | .220 | .695 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .004 | .012 | .743 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .001 | .004 | .781 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.001 | .004 | .695 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .392 | .346 | .259 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.006 | .005 | .255 |

| Density of personal references | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 121.223 | 2.561 | .000 | 10.191 | 22.369 | .000 | 102.085 | 2.884 | .000 | 102.911 | 2.937 | .000 | 102.291 | 21.257 | .000 | 125.937 | 26.950 | .000 | 135.088 | 31.849 | .000 | 111.458 | 66.864 | .096 |
| Age | | | | .335 | .354 | .348 | .394 | .331 | .238 | .394 | .331 | .238 | .394 | .331 | .238 | .016 | .424 | .970 | -.155 | .503 | .758 | .233 | 1.063 | .827 |
| Sex | | | | | | | -14.250 | 4.857 | .005 | -14.250 | 4.857 | .005 | -14.250 | 4.857 | .005 | -72.228 | 41.925 | .090 | -6.083 | 49.760 | .230 | -42.108 | 105.161 | .689 |
| Task | | | | | | | | | | -1.652 | 2.998 | .582 | -1.652 | 2.998 | .582 | -1.652 | 2.998 | .582 | -19.955 | 34.082 | .559 | 5.689 | 89.993 | .574 |
| Moment | | | | | | | | | | | | | .310 | 1.836 | .866 | .310 | 1.836 | .866 | .310 | 1.827 | .865 | 12.125 | 29.457 | .681 |
| Age*sex | | | | | | | | | | | | | | | | .920 | .661 | .169 | .792 | .785 | .315 | .461 | 1.658 | .781 |
| Age*task | | | | | | | | | | | | | | | | | | | .342 | .542 | .528 | -.810 | 1.431 | .572 |
| Sex*task | | | | | | | | | | | | | | | | | | | -24.291 | 53.603 | .651 | -58.728 | 141.538 | .679 |
| Age*moment | | | | | | | | | | | | | | | | | | | .255 | .845 | .763 | -.194 | .468 | .679 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -8.987 | 46.329 | .846 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -35.322 | 41.659 | .397 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .905 | 2.231 | .685 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .166 | .730 | .821 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .576 | .662 | .385 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 17.218 | 65.519 | .793 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.325 | 1.033 | .753 |

| Density of conjunctions | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 56.615 | 1.279 | .000 | 52.100 | 11.245 | .000 | 52.302 | 11.218 | .000 | 5.765 | 11.256 | .000 | 52.152 | 11.481 | .000 | 36.427 | 14.401 | .014 | 34.378 | 17.825 | .056 | 65.249 | 4.153 | .105 |
| Age | | | | .072 | .178 | .688 | .078 | .178 | .661 | .078 | .178 | .661 | .078 | .178 | .661 | .330 | .226 | .149 | .360 | .281 | .203 | -.143 | .639 | .823 |
| Sex | | | | | | | -1.519 | 2.609 | .563 | -1.519 | 2.609 | .563 | -1.519 | 2.609 | .563 | 37.038 | 22.320 | .102 | 27.069 | 27.808 | .332 | 22.371 | 63.151 | .723 |
| Task | | | | | | | | | | 3.074 | 1.850 | .098 | 3.074 | 1.849 | .098 | 3.074 | 1.849 | .098 | 7.173 | 21.092 | .734 | -13.477 | 55.125 | .807 |
| Moment | | | | | | | | | | | | | -.694 | 1.132 | .541 | -.694 | 1.132 | .541 | -.694 | 1.130 | .540 | -16.129 | 18.044 | .372 |
| Age*sex | | | | | | | | | | | | | | | | -.612 | .352 | .087 | -.448 | .438 | .309 | -.445 | .996 | .655 |
| Age*task | | | | | | | | | | | | | | | | | | | -.060 | .335 | .858 | .358 | .877 | .683 |
| Sex*task | | | | | | | | | | | | | | | | | | | 19.938 | 33.173 | .548 | -63.643 | 86.699 | .464 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.328 | .523 | .531 | .251 | .287 | .382 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 2.349 | 28.379 | .934 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 1.325 | 25.518 | .686 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .968 | 1.367 | .479 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.002 | .447 | .997 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.209 | .406 | .607 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 41.790 | 4.134 | .299 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.648 | .633 | .307 |

**Linear mixed models - Effects of age, sex, task and moment on the process variables (Chapter 4)**

| Model | Effects |
|---|---|
| Model 0 | Random variance |
| Model 1 | Effect of age |
| Model 2 | Effect of age and sex |
| Model 3 | Effect of age, sex and task |
| Model 4 | Effect of age, sex, task and moment |
| Model 5 | Effect of task, moment and interaction effect between age and sex |
| Model 6 | Effect of moment and the interaction effect between age, sex and task |
| Model 7 | Interaction effect between age, sex, task and moment |

| Mean length of Utterance | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 3.460.402 | 3 | | | |
| Model 1 | 3.457.409 | 4 | 2.993 | 1 | 0.084 |
| Model 2 | 3.457.362 | 5 | 0.048 | 1 | 0.827 |
| Model 3 | 3.445.579 | 6 | 11.783 | 1 | 0.001 |
| Model 4 | 3.442.147 | 8 | 3.431 | 2 | 0.180 |
| Model 5 | 3.442.136 | 9 | 0.011 | 1 | 0.916 |
| Model 6 | 3.441.467 | 12 | 0.669 | 3 | 0.880 |
| Model 7 | 3.428.987 | 26 | 12.481 | 14 | 0.568 |



| Keystrokes (incl. spaces) per minute | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 3.489.123 | 3 | | | |
| Model 1 | 3.485.216 | 4 | 3.907 | 1 | 0.048 |
| Model 2 | 3.485.203 | 5 | 0.013 | 1 | 0.908 |
| Model 3 | 3.395.971 | 6 | 89.231 | 1 | 0.000 |
| Model 4 | 3.372.911 | 8 | 23.061 | 2 | 0.000 |
| Model 5 | 3.372.759 | 9 | 0.151 | 1 | 0.697 |
| Model 6 | 3.356.937 | 12 | 15.822 | 3 | 0.001 |
| Model 7 | 3.255.145 | 26 | 101.792 | 14 | 0.000 |
| Words written per minute during the writing process | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 2.272.365 | 3 | | | |
| Model 1 | 2.268.496 | 4 | 3.870 | 1 | 0.049 |
| Model 2 | 2.268.404 | 5 | 0.092 | 1 | 0.762 |
| Model 3 | 2.196.031 | 6 | 72.373 | 1 | 0.000 |
| Model 4 | 2.173.552 | 8 | 22.479 | 2 | 0.000 |
| Model 5 | 2.173.395 | 9 | 0.157 | 1 | 0.692 |
| Model 6 | 2.159.595 | 12 | 13.800 | 3 | 0.003 |
| Model 7 | 2.054.604 | 26 | 104.990 | 14 | 0.000 |
| Number of keystrokes per minute in interval 1 (the start) | | | | | |
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 3.673.903 | 3 | | | |
| Model 1 | 3.667.557 | 4 | 6.346 | 1 | 0.012 |
| Model 2 | 3.667.557 | 5 | 0.000 | 1 | 0.985 |
| Model 3 | 3.337.737 | 6 | 329.819 | 1 | 0.000 |
| Model 4 | 3.333.971 | 8 | 3.766 | 2 | 0.152 |
| Model 5 | 3.333.938 | 9 | 0.033 | 1 | 0.856 |
| Model 6 | 3.326.847 | 12 | 7.091 | 3 | 0.069 |
| Model 7 | 3.317.470 | 26 | 9.377 | 14 | 0.806 |

| Number of utterances per minute | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 1.084.029 | 3 | | | |
| Model 1 | 1.080.064 | 4 | 3.965 | 1 | 0.046 |
| Model 2 | 1.077.290 | 5 | 2.774 | 1 | 0.096 |
| Model 3 | 926.406 | 6 | 150.885 | 1 | 0.000 |
| Model 4 | 910.665 | 8 | 15.741 | 2 | 0.000 |
| Model 5 | 910.208 | 9 | 0.457 | 1 | 0.499 |
| Model 6 | 909.300 | 12 | 0.908 | 3 | 0.823 |
| Model 7 | 835.779 | 26 | 73.521 | 14 | 0.000 |

| Ratio keystrokes product/process | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 2.915.324 | 3 | | | |
| Model 1 | 2.914.866 | 4 | 0.458 | 1 | 0.498 |
| Model 2 | 2.914.859 | 5 | 0.008 | 1 | 0.930 |
| Model 3 | 2.914.814 | 6 | 0.044 | 1 | 0.834 |
| Model 4 | 2.912.632 | 8 | 2.182 | 2 | 0.336 |
| Model 5 | 2.912.627 | 9 | 0.005 | 1 | 0.946 |
| Model 6 | 2.910.384 | 12 | 2.244 | 3 | 0.523 |
| Model 7 | 2.898.338 | 26 | 12.045 | 14 | 0.603 |

| Mean pause time (in s) within words | | | | | |
|---|---|---|---|---|---|
| | Model Fit | | Model Comparison | | |
| | -2LL | df | χ2change | dfchange | p |
| Model 0 | 1.178.991 | 3 | | | |
| Model 1 | 1.178.988 | 4 | 0.003 | 1 | 0.955 |
| Model 2 | 1.178.625 | 5 | 0.363 | 1 | 0.547 |
| Model 3 | 1.177.795 | 6 | 0.830 | 1 | 0.362 |
| Model 4 | 1.177.352 | 8 | 0.443 | 2 | 0.801 |
| Model 5 | 1.176.837 | 9 | 0.515 | 1 | 0.473 |
| Model 6 | 1.173.599 | 12 | 3.238 | 3 | 0.356 |
| Model 7 | 1.160.622 | 26 | 12.976 | 14 | 0.528 |

## Linear mixed models - Process variables - Estimates of fixed effects (Chapter 4)

**Mean length of utterance**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 43.260 | 5.193 | .000 | 12.999 | 44.633 | .009 | 12.712 | 44.634 | .009 | 115.018 | 44.715 | .013 | 111.268 | 44.888 | .016 | 115.115 | 58.115 | .052 | 10.660 | 61.273 | .105 | 11.621 | 89.462 | .217 |
| Age | | | | -1.236 | .705 | .085 | -1.246 | .706 | .083 | -1.247 | .707 | .083 | -1.251 | .707 | .082 | -1.313 | .922 | .160 | -1.084 | .972 | .269 | -1.173 | 1.421 | .410 |
| Sex | | | | | | | 2.261 | 1.356 | .828 | 2.279 | 1.368 | .827 | 2.226 | 1.371 | .831 | -7.217 | 91.232 | .937 | 15.089 | 96.089 | .876 | -58.546 | 14.394 | .677 |
| Task | | | | | | | | | | 11.604 | 3.345 | .001 | 11.604 | 3.339 | .001 | 11.604 | 3.339 | .001 | 4.371 | 38.688 | .298 | 5.583 | 99.909 | .955 |
| Moment | | | | | | | | | | | | | 2.009 | 2.048 | .327 | 2.010 | 2.048 | .327 | 2.018 | 2.046 | .325 | -3.190 | 33.255 | .924 |
| Age*sex | | | | | | | | | | | | | | | | .150 | 1.437 | .917 | -.201 | 1.514 | .895 | .890 | 2.212 | .688 |
| Age*task | | | | | | | | | | | | | | | | | | | -.456 | .613 | .458 | .013 | 1.587 | .993 |
| Sex*task | | | | | | | | | | | | | | | | | | | -44.549 | 6.146 | .459 | -111.177 | 156.826 | .479 |
| Age*moment | | | | | | | | | | | | | | | | | | | .701 | .946 | .460 | .048 | .527 | .927 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 37.045 | 51.583 | .473 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 17.877 | 46.905 | .703 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | 1.748 | 2.468 | .479 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.549 | .812 | .500 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.242 | .744 | .745 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 32.525 | 72.966 | .656 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.513 | 1.148 | .655 |

**Keystrokes (incl. spaces) per minute**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 81.048 | 3.728 | <2e-16 | 144.503 | 31.771 | .000 | 144.605 | 31.779 | .000 | 159.621 | 31.813 | .000 | 174.131 | 32.003 | .000 | 163.946 | 41.304 | .000 | 21.916 | 44.221 | .000 | 113.269 | 64.954 | .082 |
| Age | | | | -1.009 | .502 | .049 | -1.006 | .503 | .050 | -1.006 | .503 | .050 | -1.006 | .503 | .050 | -.843 | .654 | .203 | -1.585 | .701 | .027 | -.466 | 1.033 | .652 |
| Sex | | | | | | | -.852 | 7.383 | .909 | -.852 | 7.383 | .909 | -.852 | 7.383 | .909 | 24.224 | 64.837 | .710 | 27.209 | 69.502 | .697 | 132.291 | 102.391 | .198 |
| Task | | | | | | | | | | -3.033 | 2.938 | <2e-16 | -3.033 | 2.855 | <2e-16 | -3.033 | 2.855 | <2e-16 | -123.972 | 31.760 | .000 | -1.507 | 73.779 | .887 |
| Moment | | | | | | | | | | | | | -7.255 | 1.748 | .000 | -7.255 | 1.748 | .000 | -7.255 | 1.702 | .000 | 41.569 | 24.150 | .086 |
| Age*sex | | | | | | | | | | | | | | | | -.398 | 1.022 | .698 | -.456 | 1.095 | .678 | -2.112 | 1.614 | .192 |
| Age*task | | | | | | | | | | | | | | | | | | | 1.485 | .505 | .004 | .573 | 1.173 | .626 |
| Sex*task | | | | | | | | | | | | | | | | | | | -5.970 | 5.066 | .905 | -71.302 | 116.303 | .540 |
| Age*moment | | | | | | | | | | | | | | | | | | | .117 | .789 | .882 | -.560 | .384 | .146 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -52.541 | 38.069 | .169 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -56.733 | 34.153 | .098 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | 1.083 | 1.833 | .555 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .828 | .600 | .169 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .456 | .543 | .402 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 32.666 | 53.838 | .545 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.483 | .848 | .570 |

**Words written per minute during the writing process**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 14.327 | .661 | <2e-16 | 25.533 | 5.639 | .000 | 25.581 | 5.636 | .000 | 27.959 | 5.642 | .000 | 3.558 | 5.676 | .000 | 28.718 | 7.325 | .000 | 36.819 | 7.853 | .000 | 18.664 | 11.573 | .108 |
| Age | | | | -.178 | .089 | .050 | -.177 | .089 | .053 | -.177 | .089 | .053 | -.177 | .089 | .053 | -.147 | .116 | .210 | -.275 | .124 | .030 | -.064 | .184 | .729 |
| Sex | | | | | | | -.397 | 1.309 | .763 | -.397 | 1.309 | .763 | -.397 | 1.309 | .763 | 4.134 | 11.499 | .720 | 3.895 | 12.342 | .753 | 21.365 | 18.243 | .243 |
| Task | | | | | | | | | | -4.756 | .525 | <2e-16 | -4.756 | .510 | <2e-16 | -4.756 | .510 | <2e-16 | -2.959 | 5.688 | .000 | .458 | 13.185 | .972 |
| Moment | | | | | | | | | | | | | -1.300 | .312 | .000 | -1.300 | .312 | .000 | -1.300 | .305 | .000 | 7.778 | 4.316 | .073 |
| Age*sex | | | | | | | | | | | | | | | | -.072 | .181 | .693 | -.071 | .194 | .715 | -.347 | .287 | .229 |
| Age*task | | | | | | | | | | | | | | | | | | | .255 | .090 | .005 | .078 | .210 | .709 |
| Sex*task | | | | | | | | | | | | | | | | | | | .479 | 8.967 | .957 | -1.125 | 2.785 | .627 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.001 | .141 | .994 | -.105 | .069 | .126 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -8.735 | 6.804 | .200 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -1.709 | 6.104 | .080 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .146 | .328 | .656 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .138 | .107 | .200 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .088 | .097 | .363 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 5.302 | 9.622 | .582 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.074 | .152 | .628 |

**Number of keystrokes per minute in interval 1 (the start)**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 76.508 | 3.609 | <2e-16 | 154.522 | 3.233 | .000 | 154.499 | 3.256 | .000 | 113.463 | 3.467 | .000 | 106.850 | 3.642 | .001 | 102.246 | 39.602 | .012 | 72.325 | 43.177 | .098 | 101.482 | 71.669 | .158 |
| Age | | | | -1.240 | .477 | .012 | -1.240 | .478 | .012 | -1.186 | .481 | .017 | -1.191 | .481 | .016 | -1.117 | .627 | .080 | -.615 | .684 | .371 | -1.173 | 1.139 | .304 |
| Sex | | | | | | | .135 | 6.976 | .985 | 1.013 | 7.050 | .886 | .944 | 7.047 | .894 | 12.256 | 62.049 | .844 | 68.369 | 67.517 | .314 | 77.267 | 112.380 | .492 |
| Task | | | | | | | | | | 74.866 | 2.993 | <2e-16 | 74.843 | 2.974 | <2e-16 | 74.842 | 2.974 | <2e-16 | 134.352 | 34.018 | .000 | 78.465 | 87.672 | .372 |
| Moment | | | | | | | | | | | | | 3.501 | 1.819 | .055 | 3.499 | 1.819 | .055 | 3.494 | 1.796 | .053 | -11.143 | 29.189 | .703 |
| Age*sex | | | | | | | | | | | | | | | | -.179 | .978 | .855 | -1.130 | 1.064 | .291 | -1.358 | 1.770 | .444 |
| Age*task | | | | | | | | | | | | | | | | | | | -.998 | .539 | .065 | .101 | 1.393 | .942 |
| Sex*task | | | | | | | | | | | | | | | | | | | -113.515 | 53.211 | .034 | -63.610 | 137.782 | .645 |
| Age*moment | | | | | | | | | | | | | | | | | | | 1.924 | .838 | .022 | .280 | .463 | .546 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -4.390 | 45.289 | .923 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 27.949 | 41.111 | .497 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | 1.260 | 2.169 | .562 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .113 | .713 | .874 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.550 | .652 | .400 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -24.736 | 64.024 | .700 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .330 | 1.008 | .744 |

**Number of utterances per minute**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 2.502 | .089 | <2e-16 | .977 | .755 | .201 | 1.017 | .738 | .174 | 1.677 | .744 | .028 | 2.079 | .752 | .008 | 1.664 | .968 | .091 | 1.570 | 1.097 | .156 | 1.463 | 1.929 | .449 |
| Age | | | | .024 | .012 | .047 | .025 | .012 | .033 | .025 | .012 | .036 | .026 | .012 | .034 | .032 | .015 | .040 | .034 | .017 | .054 | .024 | .031 | .437 |
| Sex | | | | | | | -.287 | .171 | .097 | -.290 | .172 | .097 | -.285 | .172 | .104 | .730 | 1.509 | .631 | 1.307 | 1.709 | .446 | 2.782 | 3.023 | .358 |
| Task | | | | | | | | | | -1.293 | .092 | <2e-16 | -1.294 | .089 | <2e-16 | -1.293 | .089 | <2e-16 | -1.106 | 1.033 | .285 | -.949 | 2.449 | .699 |
| Moment | | | | | | | | | | | | | -.213 | .055 | .000 | -.214 | .055 | .000 | -.214 | .055 | .000 | -.127 | .815 | .876 |
| Age*sex | | | | | | | | | | | | | | | | -.016 | .024 | .501 | -.026 | .027 | .340 | -.049 | .048 | .305 |
| Age*task | | | | | | | | | | | | | | | | | | | -.003 | .016 | .832 | .019 | .039 | .631 |
| Sex*task | | | | | | | | | | | | | | | | | | | -1.154 | 1.606 | .473 | -.578 | 3.845 | .881 |
| Age*moment | | | | | | | | | | | | | | | | | | | .019 | .025 | .443 | .005 | .013 | .727 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -.771 | 1.264 | .542 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -.149 | 1.150 | .897 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .008 | .061 | .898 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .012 | .020 | .545 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.010 | .018 | .582 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.200 | 1.789 | .911 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .005 | .028 | .870 |

**Ratio keystrokes product/process**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 9.570 | 1.003 | <2e-16 | 96.540 | 8.859 | .000 | 96.568 | 8.864 | .000 | 96.757 | 8.910 | .000 | 97.076 | 9.154 | .000 | 97.547 | 11.773 | .000 | 111.441 | 15.709 | .000 | 95.717 | 38.664 | .014 |
| Age | | | | -.095 | .140 | .500 | -.094 | .140 | .504 | -.094 | .140 | .504 | -.094 | .140 | .505 | -.101 | .183 | .582 | -.315 | .247 | .204 | -.074 | .614 | .904 |
| Sex | | | | | | | -.180 | 2.043 | .930 | -.181 | 2.043 | .930 | -.178 | 2.043 | .931 | -1.319 | 18.035 | .942 | -11.764 | 24.229 | .628 | 2.657 | 6.582 | .965 |
| Task | | | | | | | | | | -.377 | 1.798 | .834 | -.376 | 1.798 | .834 | -.377 | 1.798 | .834 | -27.999 | 2.780 | .179 | 12.933 | 54.169 | .812 |
| Moment | | | | | | | | | | | | | -.167 | 1.101 | .879 | -.166 | 1.101 | .880 | -.169 | 1.096 | .878 | 7.825 | 17.977 | .664 |
| Age*sex | | | | | | | | | | | | | | | | .018 | .284 | .949 | .166 | .381 | .664 | -.084 | .954 | .930 |
| Age*task | | | | | | | | | | | | | | | | | | | .425 | .330 | .198 | -.186 | .860 | .829 |
| Sex*task | | | | | | | | | | | | | | | | | | | 2.775 | 32.345 | .521 | -25.255 | 85.052 | .767 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.294 | .509 | .564 | -.123 | .285 | .668 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | -7.342 | 27.948 | .793 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | -2.848 | 25.404 | .413 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | .422 | 1.339 | .753 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | .127 | .440 | .774 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | .311 | .403 | .440 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | 23.384 | 39.557 | .555 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -.363 | .623 | .560 |

**Mean pause time (in s) within words**

| | Model 0 | | | Model 1 | | | Model 2 | | | Model 3 | | | Model 4 | | | Model 5 | | | Model 6 | | | Model 7 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. | Est. | SE | Sig. |
| (Intercept) | 4.437 | .177 | <2e-16 | 4.526 | 1.567 | .006 | 4.585 | 1.571 | .005 | 4.688 | 1.565 | .004 | 4.917 | 1.614 | .004 | 5.862 | 2.087 | .007 | 9.020 | 2.701 | .001 | 26.391 | 6.734 | .000 |
| Age | | | | -.001 | .025 | .955 | -.001 | .025 | .968 | -.000 | .024 | .986 | -.000 | .024 | .993 | -.015 | .032 | .640 | -.065 | .042 | .125 | -.324 | .105 | .002 |
| Sex | | | | | | | -.219 | .364 | .550 | -.228 | .362 | .532 | -.233 | .363 | .524 | -2.463 | 3.149 | .438 | -5.510 | 4.031 | .174 | -27.119 | 1.188 | .008 |
| Task | | | | | | | | | | -.304 | .332 | .362 | -.296 | .332 | .375 | -.305 | .332 | .360 | -7.202 | 3.840 | .062 | -27.062 | 1.310 | .009 |
| Moment | | | | | | | | | | | | | -.122 | .203 | .550 | -.118 | .203 | .561 | -.120 | .203 | .553 | -9.003 | 3.151 | .005 |
| Age*sex | | | | | | | | | | | | | | | | .035 | .049 | .480 | .083 | .063 | .192 | .401 | .160 | .013 |
| Age*task | | | | | | | | | | | | | | | | | | | .108 | .060 | .073 | .399 | .162 | .015 |
| Sex*task | | | | | | | | | | | | | | | | | | | 6.641 | 5.934 | .264 | 31.986 | 15.406 | .039 |
| Age*moment | | | | | | | | | | | | | | | | | | | -.104 | .093 | .261 | .133 | .049 | .008 |
| Sex*moment | | | | | | | | | | | | | | | | | | | | | | 11.033 | 4.717 | .020 |
| Task*moment | | | | | | | | | | | | | | | | | | | | | | 1.141 | 4.718 | .033 |
| Age*sex*task | | | | | | | | | | | | | | | | | | | | | | -.478 | .241 | .049 |
| Age*sex*moment | | | | | | | | | | | | | | | | | | | | | | -.162 | .074 | .029 |
| Age*task*moment | | | | | | | | | | | | | | | | | | | | | | -.149 | .074 | .046 |
| Sex*task*moment | | | | | | | | | | | | | | | | | | | | | | -12.943 | 7.086 | .069 |
| Age*sex*task*moment | | | | | | | | | | | | | | | | | | | | | | .191 | .111 | .087 |

## APPENDIX F

**Overview of the different stories in Dutch, with mistakes underlined (Chapter 5)**

| **AD** | | |
|---|---|---|
| NWT1 | M1 | Ik zie een zwarte kat die zit en die kijkt rechts kijkt. Ik zie ook een witachtig vliegtuig voor op reis te gaan. Ik zie een bruine stoel voor op te zitten. Ik zie ook een ring die ik graag zie. Ik denk dat er een diamant op de ring bevindt. Ik vind een stoel is handig om te zitten. Ik zal bij enkele weken het vliegtuig nemen om te reisen naar Malta. Ik vind een kat mooi en lief? mmar helaas zitten er in onze buurt te veel zwerfkatten. Ik zie bij mooi weer veel vliegtuigen vliegen in de lucht en soms vliegen laag boven ons. Ik denk dat veel mensen handig vinden om met het vliegtuig op reis te gaan. Ik denk dat een srtoel onmisbaar is voor mij. Ik zet graag eens bij mooi weer de stoelen buiten. |
| NWT2 | M1 | De cowboy is een store man en heeft hij heeft een been die geanbuteerd is. De cob-wboy heeft een oog die bedekt is. De cobwboy heeft een medaille gewonnen. Hij was de eerste . Hij at veel krabben en kreeften. Hij vind dat lekker . De krab was rood en de kreeft was orange. Hijheeft gefiets met een step die grijs had rode handvaten. Het plankje waar hij zijn voet moet opzetten is rood en de wieltjes waren rood. Hij streek zij kledij op de grijze strijkplank. De cowboy stond wat schuin. |



| **AVS** | | |
|---|---|---|
| NWT1 | M1 | Vorig jaar nam ik het vliegtuig naar India, het land van de heilige koe. Het verkeer is er een chaos. Autos rijden er kris kras door elkaar heen, alsof ze geen rekening houden met elkaar. Als voetganger zou ik er niet aan denken de massa te doorkruisen, nochtans zie je heel regelmatig, zelf oudere mensen gewoon lukraak de staart oversteken. Op een dag nam ik toch de moed om een wandeling te maken door de stad. Na een tijdje werd de chaos minder storend en werd ik wat onoplettender. Ik was mijn zonnebril vergeten en door de warmte begonnen mijn ogen te tranen. Ik kocht een flesje water aan een overdekt kraampje. Toen ik buiten kwam verblinde het zonlicht mij en ik liep recht op een koe die midden op de weg stond. Het beest draaide zich en gaf me zijdelings een stoot. Daar vloog mijn flesje water net op het hoofd van een bedelende vrouw. De vrouw was gezegend en ik probeerde mezelf te verontschuldigen met een stevige duit in haar hoed. |
| | M2 | ik ging met de fiets naar de markt om een kip te kopen. Ik nam voor alle zekerheid een doos mee. Op de markt waren verschillende kippen. De verkoper raadde mij de witte kip soort aan om dat die meer eieren leggen. Dus kocht ik de witte kip. De kip was gelukkig in onze tuin en mocht vrij rondlopen. Op een dag werkt ik in de tuin en verloor mijn verlovingsring. Nergens was hij te vinden. Mijn verloofde was kwaad en verbrak de verloving. Toen ik na drie jaar de kip doodde om op te eten vond ik de ring in de maag van de kip. |
| NWT2 | M1 | Er was eens een tovenaar in Zweden die mee deed aan een zeilwedstrijd. Hij hoopte op de gouden medaille. Maar toen hij het mooie landschap zag was hij helemaal betoverd. Hij nam zijn verrekijker en bewonderde de fauna en flora. Hij was echter teleurgesteld over de herten die leken op de herten uit zijn thuisland. Toen nam hij zijn toverstaf en toverde de herten veel groter. Daarna toverde hij ook de geweien groter. Zo werd de eland geboren. De tovenaar heeft de zeilwedstrijd niet gewonnen, maar in het mooie landschap zijn nog steeds elanden te bewonderen. |
| | M2 | Vorig jaar ging mijn tante op reis om de natuur te bewonderen. Ze wilde graag wilde zwanen zien. Ze streek haar wandelkledij en vulde haar koffer. Ze trok door nederland op weg naar het noorden. Toen ze helemaal in het noorden kwam had ze al een massa aanverschillende vogels gezien. Maar plots midden in de nacht zag ze iets raars aan de lucht. Het was een raket en die landde in de buurt van haar tent. Uit de raket kwam een eskimo. De eskimo begroette haar en had een pauw in zijn armen. Mijn tante en de eskimo werden dikke vrienden. |

| JP | | |
|---|---|---|
| NWT1 | M1 | Ik had te diep in de kles gekeken en moet nog met de wagen naar huis rijden.Maar tegenwoordig met de controles op de weg, is het zeker niet aangeraden om met dewagen naar huis te rijden. Dus maar naar huis gebeld om mijn vrouw naar raad te vragen. Na een paar kwade opmerkingen, want ze zat juist onder de lampadair een boe te lezen, over dramatische ongevallen.Ze stelde dan ook maar voor van mijn zoon op te bellen en samen, mij te komen ophalen.Zo gezegd zo gedaan, een half uurtje later waren zij ter plaatse, nog steeds zeer mis tevreden, en met reden.NA EEN PAAR DAGEN STILTE PRAAT ZE NU TOCH TERUG MET MIJ. vOOR MIJ IS HET EENLES: wil je met de wagen rijden blijf dan van de fles. |
| | M2 | Toen ik verleden week mijn broer ging bezoeken , een reisje dat ik met de elecrische trein deed, stelde ik ,vast dat ik thuis mijn bril vergeten was. Dat veroorzaakte wel een paar probleempjes in plaats van koeien in de wei zag ik olifanten, Toen ik bij mijn broer was ging ik naast mijn stoel zitten en viel op de grond. In het vervolg zal ik wat beter opletten en mijn bril niet meer vergeten. |
| | M3 | toen ik een fiettochtje ging maken kreeg ik telefoon van mijn vriendin of ik een pakje wilde halen bij haar thuis, DIT PAKJE MOEST OP EEN VAST UUR BIJ EEN KLANT ZIJN ik keek op mijn uurwerk en zg dat er niet veel tijd meer overbleef. Dan maar een sprintje ingezet om toch op tijd te komen. Een overstekende hond deed mij nog bijna vallen. Ik kon mijn eenr nog redden door toch nog op tijd te komen.Eind goed , alles goed. |



| JP | | |
|---|---|---|
| NWT2 | M1 | Een familielid van mij trok op reis naat Australie. Hij ontmoete er veel vreemde mensen en deed er rare belevenissen op.Hij onmoette er maories in de woestijn die zich bezig hielden met fetis beeldjes. In australie is het geen zeldzaamheid kangoeroes tegen te komen, maar eekhoorntes heeft hij er niet gezien. Hij heeft er ook een zeilboot tocht gemaakt in een echte antieke zeilbbot. Hij moest er zelf zijn was doen , hij had geen klasiek wasspelden om zijn was op te hangen kon er ook niet beschikken over een strijkplank en moest dus met alles zijn plan trekken. Het was wel een enige reis waar hij nog vaak over verteld. |
| | M2 | Marlijn de tovenaar is een figuur uit lang verleden tijden, hij had onntelbere mogelijkheden. Hij kon van zwaan een pauw maken, Maar bij mijn weten kon hij geen gitaar spelen. Op rolschaatsen heb ik hem ook nooit gezien, maar die bestonden toen misschien nog niet, en gezien zijn ouderdom zou hij gevallen zijn. Tatoeages bestonden vermoedelijk ook niet in zijn tijd anders had hij ook ,wel een anker op zijn borst laten tatoeen. |
| | M3 | Mijn nonkel Albert hield van avontuurlijk verhalen, niet alleen ven verhalen maar zelfs van avonturen. Zo was hij ooit gaan zwemmen in de stille Oceaan, en was er aangevallen door een haai, dat hem zijn been gekost had. Toen hij aan land gerraakt was hadden ze hem in een kruiwagen naar een dokter gebraccht. Van daar was hij in een gasthuis geraakt waar hij na lange tijd buiten kwam met een houten kunstbeen. Voor hem was dit een rijke bron om allerhande verhalen te verzinnen.Zo vertelde hij dat hij geredwerd door een dolfijn. De Kruiwagen werd soms een step en de persoon die hem per kruiwagen naar de dokter bracht kreeeg later nog een decoratie. |

| | | BDB |
|---|---|---|
| NWT1 | M1 | Ik ga volgende week op reis en hoop dat de aangengeven stoel in het vliegtuig op een goede plaats zal zijn [enter] Ik neem dan zeker geen juwelen of ringn mee. Deze verliezen op reis zou me verdriet doen [enter] Wat me het meest gaan mankeren zijn onze honde, we hebben er 3 grote ( duits herders) en onze klein Happu een King Charles Cavalier,maar die gaat op zijn beurt dan ook ,op vakatie [enter] Wij kijken er al enorm nar uit, omdat we deze trip gedurende de laatste 12 maande voorbereid hebben, speciaal is dat we op expeditie gaan met Nat Geo en we ijn enrom benieuwd awt deze bestemming voor ons in petto heeft |
| | M2 | Met de vakantie in aantocht, zullen de meeste mensen een reisje of uitstap plannen [enter] Gan fietsen met familie en vrienden is dan wel een uitstkende gelegenheid om de natuur in te trekken [enter] Iedereen heeft dan wel een voorkeur voor het logement, want een goed ben is voor iedereen een noodzaak [enter] Een pretpark of speeltuin waar de kinderen met de bal kunnen spelen of zwemmen of wandelen zal dan bij iedereen in goede aarde vallen en zeker indien een bezoekje aan de kinderboerdrij mogelijk is, waar de kinderen kennis maken met koeien, schapen, kippen enz. |
| | M3 | ik ben onlangs op reis geweest. Mijn reisinformatie haalde ik uit een boek over onze bestemming [enter] Dit boek bestelde ik via en verzendpost en dat werd geleverd in een doos [enter] Maar ter plaatse kwam de kip uit het ei. Een verrassing voor mij, want elke verplaatsing die ik ging maken om bezoeken aan steden en musea te doen, zou per bus plaats hebben |



| | | BDB |
|---|---|---|
| NWT2 | M1 | Wat zag de tovenaar door zijn verrekijker, een pauw?? neen het was een zwaan [enter] de witte kleuren van de zwaan hadden hem verblind [enter] Zou jij gaarne met een raket naar de maan vliegen?? neen was het antwoord want dan zou ik met een wasknijper alles bijeen moeten houden wat ik anders zou verliezen |
| | M2 | Hij zou het me wel allemaal vertellen, wat hij meemakte in zijn droom [enter] Eerst wilde hij de verste zeeen bevaren, en daar zijn anker uitslaan om het land van de elanden te gaan verkennen, maar zijn verbazing was groot toen hij eenlandbouwer op zijn traktor tegen kwam die hem vertelde dat het damherten waren en geen elanden [enter] Dit zou hij wel kunnen vertelen aan zijn "mechanisch vriendje" waar hij al zijn droomverhalen aan kwijt kon [enter] Wat zou die vragen hebben voor hem e, wat zou hij kunnen vertellen over zijn belevenissen |
| | M3 | Hij zielde de ganse wereld rond [enter] Overal nam hij zijn gitaar meen, want hij was een vrolijke kerel di er ook opstond als hij bij zijn bemanning kwam er voor zorgde er pikfijn uit te zien [enter] Daarom gebruikte hij ook zijn strijkplank regelmatig om een plooi strak in zijn broek te strijken [enter] Hij was verzot op het eten van kreeft en krab, de schaaldieren die hij op zijn vele tochten vaak tegen kwam |

# APPENDIX G

**Overview of participant scores in the case studies**

| | AD | | AVS | | | | JP | | | JP | | | BDB | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NWT | 1 | 2 | 1 | | 2 | | 1 | | | 2 | | | 1 | | | 2 | | |
| Moment | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| **Lexical** | | | | | | | | | | | | | | | | | | |
| Total words | 141 | 91 | 106 | 173 | 96 | 102 | 136 | 74 | 82 | 112 | 75 | 117 | 63 | 94 | 113 | 53 | 73 | 92 |
| Word length | 3.89 | 4.23 | 3.88 | 4.39 | 4.75 | 4.19 | 3.96 | 4.26 | 4.07 | 4.54 | 4.65 | 4.49 | 4.38 | 4.81 | 4.35 | 4.43 | 4.33 | 4.62 |
| Word frequency | 5.07 | 4.31 | 4.67 | 4.55 | 4.33 | 4.6 | 5.29 | 4.88 | 5.23 | 4.57 | 4.9 | 4.65 | 4.61 | 4.96 | 4.96 | 4.15 | 4.39 | 4.19 |
| Type token ratio | 0.48 | 0.59 | 0.62 | 0.66 | 0.61 | 0.65 | 0.63 | 0.66 | 0.7 | 0.68 | 0.76 | 0.67 | 0.86 | 0.67 | 0.73 | 0.74 | 0.74 | 0.68 |
| Adjectives* | 85.11 | 109.89 | 56.6 | 63.58 | 62.5 | 58.82 | 58.82 | 54.05 | 36.59 | 62.5 | 53.33 | 51.28 | 0 | 53.19 | 88.5 | 18.87 | 95.89 | 43.48 |
| Nouns* | 148.94 | 208.79 | 226.42 | 219.65 | 187.5 | 215.69 | 183.82 | 216.22 | 182.93 | 160.71 | 200 | 188.03 | 253.97 | 255.32 | 150.44 | 226.42 | 150.69 | 141.3 |
| Verbs (=v)* | 205.67 | 263.74 | 179.25 | 184.97 | 197.92 | 166.67 | 176.47 | 202.7 | 219.51 | 214.29 | 253.33 | 196.58 | 206.35 | 180.85 | 203.54 | 226.42 | 178.08 | 260.87 |
| Concrete v* | 7.09 | 43.96 | 9.43 | 23.12 | 0 | 9.8 | 22.06 | 0 | 12.2 | 8.93 | 13.33 | 8.55 | 0 | 31.91 | 0 | 0 | 27.4 | 0 |
| Abstract v* | 85.11 | 197.8 | 113.21 | 92.49 | 145.83 | 68.63 | 117.65 | 148.65 | 109.76 | 142.86 | 173.33 | 145.3 | 142.86 | 127.66 | 141.59 | 188.68 | 54.79 | 206.52 |
| **Grammatical** | | | | | | | | | | | | | | | | | | |
| D-level | 1.92 | 2.36 | 1.1 | 2.08 | 2.22 | 1.6 | 2.38 | 5.33 | 1.83 | 1.71 | 3 | 3 | 2 | 3.75 | 3 | 1.6 | 2.5 | 5.5 |
| Distance 1 | 1.45 | 0.86 | 2.13 | 1.42 | 1.31 | 1.82 | 3.08 | 2.73 | 2.21 | 4.17 | 0.73 | 2.37 | 3.73 | 6.25 | 2.88 | 2.67 | 1.29 | 1.89 |
| Distance 2 | 0.17 | 0.19 | 0.09 | 0.13 | 0.29 | 0 | 0.12 | 0.08 | 0.18 | 0.36 | 0 | 0.21 | 0 | 0.74 | 0.36 | 0.1 | 0.55 | 0.15 |
| Density pers | 120.57 | 153.85 | 122.64 | 132.95 | 125 | 166.67 | 95.59 | 202.7 | 146.34 | 116.07 | 160 | 153.85 | 111.11 | 53.19 | 159.29 | 113.21 | 178.08 | 184.78 |
| Density conj | 49.65 | 43.96 | 47.17 | 40.46 | 62.5 | 58.82 | 44.12 | 81.08 | 48.78 | 17.86 | 40 | 42.74 | 79.37 | 85.11 | 70.8 | 18.87 | 41.1 | 54.35 |

# Chapter eight

Summary

Samenvatting

# SUMMARY

Given that changes in language production already occur at the onset of several progressive disorders such as Alzheimer's disease, insight into the initial language changes compared to language characteristics of cognitively healthy ageing adults, could aid the diagnostic screening process (Kavé & Dassa, 2018; Pekkala et al., 2013; Tsantali et al., 2013). Various studies have already contributed to mapping these linguistic changes, ranging from picture naming studies (e.g., Hardy et al., 2020) to picture description studies (e.g., Forbes-McKay et al., 2014) and studies into free writing (e.t.: Le et al., 2011), thereby expanding our knowledge on of linguistic changes on a word, sentence and even text level. Nevertheless, little is known about the linguistic changes in spontaneous written language upon healthy ageing, and few of the existing tools/methods allow for tracking those language changes, especially within a test-retest setting.

To tackle this challenge, this dissertation sought to find out how spontaneous written language changes upon ageing, by adopting a linguistic approach. To that extent, four main research lines were set out: (1) to establish how age, image characteristics and repeated testing affect picture naming, (2) to find out how ageing and sex affect the cohesion of written narratives, (3) to examine the test-retest reliability of spontaneous written language in healthy ageing and (4) to explore how spontaneous written language can be used to make a differential diagnosis. Additionally, we set out to create a toolkit that was non-intrusive, low-cost and time effective. Therefore, we also wanted to (5) create a toolkit that bridges the gap within differential diagnosis with regard to spontaneous written language generation, (6) create tasks within that toolkit that are congeneric in nature and that are interchangeable on certain key variables and to (7) create and standardise that toolkit within a healthy ageing population and pilot its validity within a clinical setting. These various research lines are addressed in the various chapters within this dissertation.

In order to capture the linguistic changes of spontaneous written language, both the writing product and writing process was considered. The writing product was analysed in light of cohesion, lexical and grammatical characteristics with T-Scan (Pander Maat et al., 2014) and Comproved (Lesterhuis et al., 2016). The writing process was logged and analysed with respectively ScriptLog (Frid et al., 2014) and Inputlog (Leijten & Van Waes, 2013).



**Chapter two** presents our image validity study ('*Name agreement and naming latencies for typed picture naming in ageing adults*'), with which we aimed (1) to create new images that could be used to elicit a certain target word and when combined could be used to trigger a narrative. These images would serve as a baseline for the continuation of this dissertation, and allowed us to reach our first aim, namely 'to establish how age, image characteristics and repeated testing affect picture naming'. Single-object images were carefully selected based on certain characteristics (e.g., word frequency, age-of-acquisition, and reaction times) found previous studies and created based on strict visual criteria (colour, position). The images were divided into two categories: 'highly relatable' images (e.g., 'dog') and 'less relatable' images (e.g., 'wheelbarrow'). All images can be found in the newly founded Open Linguistic Picture Database (OLPD, www.olpd.eu, Paesen & Meulemans, 2020) A representative sample of 60 healthy controls aged 50 and over was given a typed picture naming task, in which they were asked to name the prompted images as fast as they could. With the use of the keystroke logging tools ScriptLog and Inputlog, we were able to log and analyse the naming process. The naming product was manually checked, filtered, and divided into categories based on mistakes (the use of synonyms, spelling mistakes, incorrect naming and blanks). Results indicate that object recognition in older adults, i.e.: the naming correctness of images when taking synonyms into account, was influenced by the word/image characteristics. Additionally, we found that older adults also rely on synonyms and alternative naming more often. Image characteristics affected naming latencies, in the sense that the less relatable images triggered longer response times. Age overall also affected the naming speed, whereas the test moment had no effect, thereby leading to a great test-retest reliability. The 'interkey latencies' (pauses within the typing process) showed a learning effect in a test-retest situation and could not be used to predict naming correctness. Given these findings, we concluded that both healthy ageing and image characteristics influence picture naming, and that certain variables are stable within a test-retest setting. Additionally, we also expressed the need to not treat the participant group as a homogeneous group, but rather as a set of individuals.

These findings gave rise to **chapter three**, '*Ageing and sex differences in the cohesion of written narratives*', in which the creation and standardisation of narrative writing tasks is discussed. The images created in chapter two allowed us to construct these narrative writing tasks; tasks that can be used to elicit spontaneous written language while still controlling for certain key variables (target nouns elicited by the images). This chapter, therefore, contributed to our aims of (2) exploring the potential effects ageing and sex have on the cohesion of written narratives, (6) creating congeneric tasks and (7) standardising the toolkit in a healthy population. Additionally, this study also contributes to the current set of existing test batteries by laying the basis for a linguistic toolkit and potential test. 257

healthy adults were given two narrative writing tasks (respectively NWT1 and NWT2); they were asked to write a story based on the given images and their output was analysed in light of the writing product, and more specifically, text cohesion with the use of both lexical and grammatical variables. Based on the recommendations made in chapter two, the participants were treated as a set of individuals through multilevel analysis; in the presentation of the results and the normative dataset, however, they were presented per decade. Results indicate that age did not affect sentence and word complexity and word variety. Age had an effect on word length and number of words, a result that could be attributed to the tip-of-the-tongue encountered upon ageing. Sex had an impact on various lexical and grammatical variables. For instance, we found that word length increase upon ageing for men, whereas it decreases for women. These results stress the importance of differentiating between the different sexes in research and indicate that age had no effect on the complexity of discourse. Moreover, we stress the importance of including the writing process upon analysis, as it could provide additional insight into the difficulties encountered upon constructing a narrative. Given the size of our cohort, a normative dataset could be constructed that can be used as reference in future research.

In **chapter four** ('*Clinical tool for the evaluation of written spontaneous speech in healthy adults*') we build on that normative dataset to (3) explore the test-retest reliability of spontaneous written language in healthy ageing and (7) standardise the toolkit within a healthy ageing population (in a longitudinal setting). In total, 58 healthy ageing adults participated in our study thrice, with an interpose of three months every time. They were given the same design as participants in chapter three, with two narrative writing tasks that needed to be completed. The output did not only consider the product measures used in chapter three (with the use of Comproved and T-Scan); the writing process was also studied (through ScriptLog and Inputlog). Results for the writing product showed great test-retest reliability, with no effects of the different test moments to be found, and no significant age effect was found. Other effects could be contributed to the design of the templates of NWT1 and NWT2. The writing process results show greater variability, even though the active writing time and typing speed remained stable over the three test moments. Therefore, we concluded that we were able to create congeneric tasks that allow for retesting in a longitudinal setting and that more research into the spontaneous writing process is needed.



The final chapter**, chapter five** ('*A preliminary study into the use of narrative writing tasks for an ageing population in a clinical setting*') aims to (5) create a toolkit that bridges the gap within differential diagnosis regarding spontaneous written language generation by exploring whether or not spontaneous written language could be used to make a differential diagnosis and (7) pilot its validity within a clinical setting. Within this pilot study, we focus on four case studies who received the same design and procedure as in chapter four. These four cases were found within our healthy cohort, however, due to their MoCA and / or GDS scores and after consultation with a language pathologist, were removed from the healthy cohort and placed within this study. Their narrative writing tasks were compared to the normative dataset found in chapter three. Additionally, we also considered their picture naming tasks in line with the design of chapter two. The results lead to two important considerations: (1) we believed more insight into potential pathologies could be gained with spontaneous written language rather than merely a picture naming task; our results confirm this finding. (2) Our results indicate that the variable 'Density of abstract verbs' should be further studied as a potential language biomarker.

In sum, we were able to construct a new clinical toolkit for the evaluation of spontaneous written speech, that is non-intrusive, low-cost and time effective. The different chapters within this dissertation help shed more light on spontaneous written language changes and laying the basis of a test that could potentially be used for differential diagnosis. We found that picture naming is affected by both age and image characteristics, but not repeated testing. In terms of spontaneous written language, cohesion is affected by both age and sex. Spontaneous written language remains stable in a test-retest setting, especially with regards to the writing product. More research into the writing process is needed to fully grasp possible issues encountered during spontaneous written language generation. Our pilot study indicates that further research is needed for the 'Density of abstract verbs' to explore its potential as a linguistic biomarker. In a more technical sense, we contribute to current research by providing an image database that can be used in various types of other studies (OLPD, www.olpd.eu)(Paesen & Meulemans, 2020), by creating narrative writing tasks that allow for a spontaneous written language flow while still controlling for key variables (target nouns triggered by the images), by establishing a normative dataset of cohesion measures that can be used for reference in other studies, by designing a toolkit that can be used in retest settings, and by laying the basis for a test that can be used in clinical settings. We hope that our toolkit and findings help to shape future studies on spontaneous written language, differential diagnosis, and ageing. We believe this dissertation promises to be a strong addition to the current battery of screening tools in ageing of language.

# SAMENVATTING

Gegeven dat veranderingen in taalproductie al optreden bij het begin van verschillende progressieve aandoeningen zoals de ziekte van Alzheimer, zou inzicht in die taalveranderingen en de vergelijking met het taalpatroon van cognitief gezonde volwassenen kunnen helpen bij het diagnostische screeningsproces (Kavé & Dassa, 2018; Pekkala et al., 2013; Tsantali et al., 2013). Verschillende studies hebben al bijgedragen aan het in kaart brengen van deze linguïstische veranderingen, variërend van plaatjesbenoemingsstudies (bijv.: Hardy et al., 2020) tot plaatjesbeschrijvingsstudies (bijv.: Forbes-McKay et al., 2014) en vrije schrijfstudies (bijv.: Le et al., 2011); deze brachten inzicht in linguïstische veranderingen op woord-, zin- en zelfs tekstniveau. Toch is er weinig bekend over de linguïstische veranderingen in spontaan geschreven taal bij gezonde ouderen, en weinig van de bestaande instrumenten/methoden stellen ons in staat om die taalveranderingen te volgen binnen een test-hertest setting.

In dit proefschrift werd getracht te achterhalen hoe spontaan geschreven taal verandert bij veroudering, gebruik makend van een linguïstische benadering. Daartoe werden vier hoofdlijnen voor het onderzoek uitgeschreven: (1) onderzoeken hoe leeftijd, woord/afbeeldingskenmerken en herhaald testen van invloed zijn op het benoemen van afbeeldingen, (2) onderzoeken hoe veroudering en geslacht de samenhang van geschreven verhalen beïnvloeden, (3) de test-hertest betrouwbaarheid van spontane geschreven taal bij gezonde volwassenen onderzoeken en (4) onderzoeken hoe spontane geschreven taal gebruikt kan worden om een differentiële diagnose te stellen. Bovendien wilden we een toolkit creëren die niet invasief, goedkoop en tijdsefficiënt was. Daarom wilden we ook (5) een toolkit creëren die de kloof overbrugt binnen de differentiaaldiagnostiek met betrekking tot het spontaan genereren van geschreven taal, (6) taken creëren binnen die toolkit die congenerisch zijn van aard en uitwisselbaar op bepaalde sleutelvariabelen, en (7) die toolkit creëren en standaardiseren binnen een gezond ouder wordende populatie en de validiteit ervan pretesten binnen een klinische setting.

Om de linguïstische veranderingen in spontaan geschreven taal vast te leggen, werden zowel het schrijfproduct als het schrijfproces onderzocht. Het schrijfproduct werd geanalyseerd op vlak van cohesie, lexicale en grammaticale kenmerken met behulp van T-Scan (Pander Maat et al., 2014) en Comproved (Lesterhuis et al., 2016). Het schrijfproces werd gelogd en geanalyseerd met respectievelijk ScriptLog (Frid et al., 2014) en Inputlog (Leijten & Van Waes, 2013).



**Hoofdstuk twee** presenteert onze validiteitsstudie van afbeeldingen (‘*Name agreement and naming latencies for typed picture naming in ageing adults*’), waarmee we beoogden (1) nieuwe afbeeldingen te creëren die gebruikt konden worden om een bepaald doelwoord uit te lokken en wanneer gecombineerd gebruikt konden worden om een narratief te initiëren. Deze afbeeldingen zouden als basis dienen voor de rest van dit proefschrift, en stelden ons in staat ons eerste doel te bereiken, namelijk ‘bepalen hoe leeftijd, kenmerken van plaatjes, en herhaald testen het benoemen van plaatjes beïnvloeden’. Afbeeldingen van individuele objecten werden zorgvuldig geselecteerd op basis van kenmerken uit eerdere studies (bijv.: woordfrequentie, leeftijd van verwerving, en reactietijden), en gemaakt op basis van strikte visuele criteria (bijv.: kleur, positie). De afbeeldingen werden verdeeld in twee categorieën: ‘zeer relateerbare’ afbeeldingen (bijv.: ‘hond’) en ‘minder relateerbare’ afbeeldingen (bijv.: ‘kruiwagen’). Alle afbeeldingen zijn te vinden in de nieuw opgezette Open Linguistic Picture Database (OLPD, www.olpd.eu, Paesen & Meulemans, 2020) Een representatieve steekproef van 60 gezonde volwassenen van 50 jaar en ouder kreeg een plaatjesbenoemingstaak, waarbij hen gevraagd werd de gevraagde afbeeldingen zo snel mogelijk te benoemen (door te typen). Met behulp van de toetsregistratiesoftware ScriptLog en Inputlog, waren we in staat om het benoemingsproces te loggen en te analyseren. De output werd handmatig gecontroleerd, gefilterd en gecategoriseerd op basis van fouten (het gebruik van synoniemen, spelfouten, onjuiste naamgeving en spaties). De resultaten tonen aan dat objectherkenning bij oudere volwassenen, d.w.z.: de nauwkeurigheid van het benoemen van afbeeldingen wanneer rekening wordt gehouden met synoniemen, werd beïnvloed door de woord/beeldkenmerken. Bovendien vonden we dat oudere volwassenen ook vaker gebruik maakten van synoniemen en alternatieve naamgeving. Afbeeldingskenmerken beïnvloedden de snelheid voor het benoemen: minder relateerbare beelden veroorzaakten langere reactietijden. Leeftijd had ook invloed op de benoemsnelheid, terwijl het toetsmoment geen effect had wat leidt tot een grote test-hertest betrouwbaarheid. De pauzes binnen het typeproces lieten een leereffect zien in een test-hertest situatie en konden niet gebruikt worden om de juistheid van het benoemen te voorspellen. Gezien deze bevindingen concludeerden we dat zowel leeftijd als afbeeldingskenmerken het benoemen van plaatjes beïnvloeden, en dat bepaalde variabelen stabiel zijn binnen een test-hertest setting. Daarnaast gaven we ook aan dat we de deelnemersgroep niet als een homogene groep moesten behandelen, maar eerder als een geheel van individuen.

Deze bevindingen gaven aanleiding voor **hoofdstuk drie**, '*Ageing and sex differences in the cohesion of written narratives*', waarin de creatie en standaardisatie van narratieve schrijftaken wordt besproken. De beelden die in hoofdstuk twee werden gecreëerd, stelden ons in staat om deze narratieve schrijftaken te construeren; taken die spontane schrijftaal uitlokken terwijl er nog steeds kan worden gecontroleerd voor bepaalde veraf bepaalde variabelen (meer bepaald: de namen die door de afbeeldingen worden uitgelokt). Dit hoofdstuk heeft daarom bijgedragen aan ons doel om (2) de mogelijke effecten van veroudering en sekse op de samenhang van geschreven verhalen te onderzoeken, (6) congenerische taken te creëren en (7) de toolkit te standaardiseren in een gezonde populatie. Daarnaast levert deze studie ook een bijdrage aan de huidige set van bestaande testbatterijen door de basis te leggen voor een linguïstische toolkit en potentiële test. 257 gezonde volwassenen kregen twee narratieve schrijftaken (respectievelijk NWT1 en NWT2); hen werd gevraagd een verhaal te schrijven op basis van de gegeven afbeeldingen en hun output werd geanalyseerd met een focus op het schrijfproduct, en meer specifiek tekstcohesie, met behulp van zowel lexicale als grammaticale variabelen. Op basis van de aanbevelingen in hoofdstuk twee werden de deelnemers behandeld als een set van individuen door middel van multilevel analyse; in de presentatie van de resultaten en de normatieve dataset werden ze voor de duidelijkheid echter per decade voorgesteld. De resultaten gaven aan dat leeftijd geen effect had op zins- en woordcomplexiteit en woordvariëteit. Leeftijd had een effect op de woordlengte en het aantal woorden. Geslacht had een effect op verschillende lexicale en grammaticale variabelen. Zo vonden we bijvoorbeeld dat de woordlengte bij mannen toeneemt bij het ouder worden, terwijl die bij vrouwen afneemt. Deze resultaten benadrukken het belang van differentiatie tussen de verschillende seksen in onderzoek en geven aan dat leeftijd geen effect had op de complexiteit van het discours. Bovendien benadrukken we dat het belangrijk is om het schrijfproces mee te nemen in de analyse, omdat dit extra inzicht kan geven in de moeilijkheden die men ondervindt bij het construeren van een narratief. Gezien de omvang van onze cohort, kon een normatieve dataset geconstrueerd worden die als referentie gebruikt kan dienen in toekomstig onderzoek.

In **hoofdstuk vier** ('*Clinical tool for the evaluation of written spontaneous speech in healthy adults*') bouwen we voort op die normatieve dataset om (3) de test-hertest betrouwbaarheid van spontane geschreven taal bij gezonde volwassenen te onderzoeken en (7) de toolkit te standaardiseren binnen een gezond ouder wordende populatie (in een longitudinale setting). In totaal namen 58 gezonde ouder wordende volwassenen driemaal deel aan onze studie, met een tussenpose van telkens drie maanden. Ze kregen dezelfde onderzoeksopzet als de deelnemers in hoofdstuk drie, met twee narratieve schrijftaken die voltooid moesten worden. Bij de analyse werd niet alleen gekeken naar de productmaten die in hoofdstuk drie werden gebruikt (met behulp van Comproved en T-Scan); ook het schrijfproces werd bestudeerd (met behulp van ScriptLog en Inputlog). De resultaten voor het schrijfproduct lieten een grote test-hertest betrouwbaarheid zien, waarbij geen effecten van de verschillende testmomenten te vinden waren, en er werd geen significant leeftijdseffect gevonden. Andere effecten zouden kunnen worden toegeschreven aan het design van twee verschillende schrijftaken. De resultaten van het schrijfproces vertonen meer variabiliteit, hoewel de actieve schrijftijd en de typesnelheid stabiel bleven over de drie testmomenten. Daarom concludeerden we dat we in staat waren om congenerische taken te creëren die hertesten in een longitudinale setting mogelijk maken en dat meer onderzoek naar het spontane schrijfproces nodig is.

Het laatste hoofdstuk, **hoofdstuk vijf** ('*A preliminary study into the use of narrative writing tasks for an ageing population in a clinical setting*') heeft als doel (5) een toolkit te creëren die de kloof overbrugt binnen de differentiaal diagnostiek met betrekking tot het spontaan genereren van geschreven taal door te onderzoeken of spontaan geschreven taal wel of niet gebruikt kan worden om een differentiaal diagnose te stellen en (7) de validiteit ervan te testen binnen een klinische setting. Binnen deze pilotstudie richten we ons op vier casussen die dezelfde opzet en procedure kregen als in hoofdstuk vier. Deze vier casussen werden gevonden binnen ons gezonde cohort. Echter, vanwege hun MoCA en/of GDS scores en na overleg met een taalpatholoog, werden ze uit het gezonde cohort gehaald en binnen deze studie geplaatst. Hun narratieve schrijftaken werden vergeleken met de normatieve dataset gevonden in hoofdstuk drie. Daarnaast werd ook gekeken naar hun benoemingstaak, in overeenstemming met de opzet van hoofdstuk twee. De resultaten leidden tot twee belangrijke bevindingen: (1) we dachten dat meer inzicht in potentiële ziektebeelden kon worden verkregen met spontane geschreven taal in plaats van alleen een plaatjesbenoemingstaak; onze resultaten bevestigen deze bevinding. (2) Onze resultaten geven aan dat de variabele 'Dichtheid van abstracte werkwoorden' verder bestudeerd moet worden als een potentiële talige biomarker.

Kortom, dit doctoraatsonderzoek construeerde een nieuwe klinische toolkit voor de evaluatie van spontaan geschreven spraak, die niet-invasief, goedkoop en tijdseffectief is. De verschillende hoofdstukken in dit proefschrift helpen meer licht te werpen op spontane veranderingen in geschreven taal en leggen de basis voor een test die mogelijk gebruikt kan worden voor differentiële diagnose. Onze resultaten geven aan dat het benoemen van afbeeldingen beïnvloed wordt door zowel leeftijd als afbeeldingskenmerken, maar niet door herhaald testen. Spontane schrijftaal wordt beïnvloed door zowel leeftijd als geslacht, maar het schrijfproduct blijft stabiel in een test-hertest setting. Meer onderzoek naar het schrijfproces is nodig om de mogelijke onderliggende problemen bij spontane taalgeneratie volledig te doorgronden. Onze pilotstudie geeft aan dat verder onderzoek nodig is naar de 'Dichtheid van abstracte werkwoorden'; deze variabele heeft namelijk het potentieel om als linguïstische biomarker te dienen. In meer technische zin dragen we bij aan het onderzoeksveld door een afbeeldingsdatabase te creëren die gebruikt kan worden in verschillende soorten onderzoek (OLPD, www.olpd.eu), door narratieve schrijftaken te maken die spontane taalgeneratie uitlokken terwijl er nog steeds gecontroleerd kan worden voor bepaalde variabelen (meer bepaald, de zelfstandige naamwoorden die getriggerd worden door de afbeeldingen), door een normatieve dataset op te stellen die kan dienen als referentie in andere studies, door een toolkit te ontwerpen die gebruikt kan worden in hertest setting, en door de basis te leggen voor een test die mogelijks kan worden ingezet in klinische settings. We hopen dat onze toolkit en bevindingen helpen om toekomstige studies over spontane schrijftaal, differentiële diagnose, en veroudering vorm te geven. Wij geloven dat dit proefschrift een sterke aanvulling is op het huidige scala aan screeningsinstrumenten op het gebied van veroudering en taal.



## Bibliography / bronvermelding


Forbes-McKay, K., Shanks, M., & Venneri, A. (2014). Charting the decline in spontaneous writing in Alzheimer's disease: a longitudinal study. *Acta Neuropsychiatrica*, 26(04), 246–252. https://doi.org/10.1017/neu.2014.2

Frid, J., Johansson, V., Johansson, R., Wengelin, Å., & Johansson, M. (2014). Developing a keystroke logging program into a writing experiment environment. *Presentation at Writing Across Borders,* WRAB, Paris February 2014.

Hardy, S. M., Segaert, K., & Wheeldon, L. (2020). Healthy Aging and Sentence Production: Disrupted Lexical Access in the Context of Intact Syntactic Planning. *Frontiers in Psychology*, 11(February). https://doi.org/10.3389/fpsyg.2020.00257

Kavé, G., & Dassa, A. (2018). Severity of Alzheimer's disease and language features in picture descriptions. *Aphasiology*, 32(1), 27–40. https://doi.org/10.1080/02687038.2017.1303441

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing,* 26(4), 435–461. https://doi.org/10.1093/llc/fqr013

Leijten, M., & Van Waes, L. (2013). Keystroke Logging in Writing Research. *Written Communication*, 30(3), 358–392. https://doi.org/10.1177/0741088313491692

Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & De Maeyer, S. (2016). Comparative Judgment as a promising alternative to score competences. In E. Cano & G. Ion (Eds.), *Innovative Practices for Higher Education Assessment and Measurement* (pp. 119–138). IGI Global. https://doi.org/10.4018/978-1-5225-0531-0

Paesen, L., & Meulemans, C. (2020). *Open Linguistic Picture Database [Data set]*. Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.3738213

Pander Maat, H., Kraf, R., Bosch, A. van den, Dekker, N., Gompel, M. van, Kleijn, S., Sanders, T. J. M., & Sloot, K. van der. (2014). T-Scan : a new tool for analyzing Dutch text. *Computational Linguistics in the Netherlands Journal*, 4, 53–74.

Pekkala, S., Wiener, D., Himali, J. J., Alexa, S., Obler, L. K., Liu, Y., & Kee, A. N. N. M. C. (2013). Lexical retrieval in discourse: An early indicator of Alzheimer's dementia. *Clinical Linguistics & Phonetics*, 27(December), 905–921. https://doi.org/https://doi.org/10.3109/02699206.2013.815278

Tsantali, E., Economidis, D., & Tsolaki, M. (2013). Could language deficits really differentiate Mild Cognitive Impairment (MCI) from mild Alzheimer's disease? *Archives of Gerontology and Geriatrics*, 57(3), 263–270. https://doi.org/10.1016/j.archger.2013.03.011

# Chapter nine

## Acknowledgments

Dat niet enkel het schrijfproduct belangrijk is maar ook het schrijfproces, heb ik de voorbije jaren aan den lijve ondervonden. Het waren vijf enorm leerrijke jaren waar ik enorm dankbaar voor ben. Omdat je zo'n proces niet alleen aangaat, wil ik verschillende mensen ontzettend hard bedanken:

**Het promotorenteam**

Miet en Mariëlle, of zoals ik jullie stiekem wel eens noemde: M&M. En net zoals de snoepjes, gaven jullie mee kleur aan dit doctoraat. Zonder jullie zou dit werk er niet liggen. Mariëlle, dank je wel om in me te geloven en me de kans te geven mijn schouders onder dit project te zetten. Merci voor de vele koffiebesprekingen en eerlijke feedback. Jouw inzicht in tekststructuur en taakdesign is er eentje waar ik nog steeds veel uit leer. Miet, ook al ben je later ingestapt, de manier waarop je deze uitdaging bent aangegaan, was voor mij hartverwarmend. Jouw enthousiasme en geloof in het onderzoek was een enorme motivator. Ik blijf verbaasd over je enorme kennis en ik kan er gefascineerd naar blijven luisteren. Dank je wel om steeds een luisterend oor en goede raad te bieden. Peter, het voelt een beetje raar deze woorden te schrijven. Ook al was onze samenwerking te kort, toch wil ik je enorm bedanken voor de vele feedback en tips in het eerste jaar van dit doctoraat. Het was een hele eer om met jou en Dorien te mogen samenwerken en van je te leren.

**Het UA-team**

Luuk, ik ben enorm vereerd samen met jou te mogen hebben gewerkt. Ik had het gevoel altijd bij je terecht te kunnen, met welke vraag dan ook. Je bent een inspiratiebron! Eric & Tom, een enorme merci voor al jullie Inputlog hulp, eender wanneer, eender hoe complex, en de vele fijne badminton momentjes! Sven, de statisticus-engel. Zonder jou bestond dit doctoraat enkel uit gemiddeldes en medianen. Merci voor vele hulp bij statistiek en voor het experiment dat ik tijdens jouw les mocht uitvoeren. Ook de mensen van D-Pac/Comproved: dank jullie wel om mee jullie schouders onder mijn onderzoek te steken! Tanya: merci dat ik met letterlijk elke vraag bij jou terecht kon; als ik het even niet meer wist, was jij de rots in de administratieve branding. Datzelfde geldt ook voor Rosina, dank je wel voor de vele geruststellende telefoontjes die me de laatste maanden doorheen dit doctoraat hebben geloodst. Paul, merci voor de vele leuke babbels en de sta bureau die je speciaal maakte! I loved it. Suzy, dank je wel voor de fijne babbels en hulp bij onze database! Ten slotte wil ik ook algemeen alle (ex-)inwoners van de C4 bedanken om mee deze doctoraatstijd zo fijn te maken. Merci!

**Het schrijfteam**

De schrijfonderzoekers van Nederland en België. Wat was het fijn om in deze groep te worden ontvangen. Merci om voor inspirerende sessies en leuke samenwerking!



**Het schrijvende team**

De deelnemers. Ontzettend bedankt om allen zo enthousiast mee te willen werken aan dit onderzoek! Zonder jullie had dit boekje er nooit gelegen en hadden we deze leerrijke data nooit verkregen. Dank jullie wel voor de vele koffies en fijne babbels die daarmee gepaard gingen. Merci ook aan allen die mee op zoek gingen naar deelnemers, die experimenten hebben ondersteund.

**The Swedish team**

Victoria, Johan, Sven thank you so much for having us, discussing various studies, allowing us to teach a class, creating new ScriptLog modules together. It was an honour to learn from you and work with you.

**Het koffieteam**

Gilles, Didi en Lies: merci voor jullie bakjes troost. Samen met de fijne babbels deden ze enorm deugd!

**Het vriendenteam**

Liefste vrienden, merci om me al die jaren te ondersteunen. Onze vele dates, de lieve woorden van aanmoediging, de unagi-momentjes, spelletjes, kattenfilmpjes en zakken chips. Love you guys.

**Het paranimfenteam**

Officieel gebruiken we die term hier niet, maar Catherine en Nina, woorden schieten me te kort. Dank jullie wel voor de vele jaren die we samen doorbrachten. De vele koffies, lunchkes, uitstapjes, schone deurversiering, dak momentjes en zoveel meer. De honderden keren dat ik jullie mocht lastigvallen met een last-minute vraag, gezaag of stressje. Zonder jullie had ik hier niet meer gezeten. Eeuwige dank.

**The dreamteam**

In alfabetische volgorde (dat is eenvoudiger): Joris, Lotje, Lucas, Mama en Papa – oftewel man van mijn leven en papa van ons Joske, metie en beste schoonzus dat je je kan wensen, peter koekie en liefste broer, allerbeste ouders en soon-to-be moeke en vake. Omdat de krop me al in keel schiet bij het schrijven van deze tekst, hou ik het kort: dank jullie wel voor al die jaren van onvoorwaardelijke steun en liefde, al de kansen die ik kreeg, fijne momenten, troostende woorden, en nog veel meer. Zonder jullie geen doctoraat. *Ik z'n blij da gij in mijn team zit*. Merci voor alles.