Language and cognition in healthy aging and dementia

Ingeborg Sophie Bjønness Ribu

Department of Linguistics and Scandinavian Studies
Faculty of Humanities
University of Oslo

Dissertation submitted for the degree of PhD
October 2019

Series of dissertations submitted to the
Faculty of Humanities, University of Oslo

This document was created with LaTeX 2_ε; the individual papers were created with Microsoft Word (2010). Figures were created with Ti*k*Z, ggplot2 and Microsoft Visio (2010).

Abstract

This dissertation investigates how changes in language performance in Alzheimer's disease (AD) and primary progressive aphasia (PPA) can help shed light on theories of language processing, the mental lexicon, and how language impairments in dementia can help give a more comprehensive understanding of the complex difficulties associated with the different diseases.

The dissertation is the first of its kind in Norway, focusing on language impairment in AD and PPA from a linguistic perspective. It contributes with new knowledge on deficits in lexical production and sentence comprehension in AD, PPA and healthy aging, shedding light on the structure of the mental lexicon in these populations. The results are in line with usage-based theories and interactive models of language processing, where the mental lexicon is seen a structured network of smaller and larger units at different levels of abstraction, sensitive to effects of frequency, age of acquisition and cognate status. Furthermore, they support a notion that theories of language should consider the multilingual mental lexicon as default. The results also indicate that there may be two different variants of the logopenic subtype of PPA.

The dissertation illustrates the advantages of using a range of different methods for assessing language, to get a detailed picture of possible impairments. By means of eye-tracking, subtle differences in processing speed between healthy adults and persons with dementia could be detected that were not seen in a parallel offline task. A free word association task detected differences that were not captured with traditional naming tasks.

Currently, language assessment plays a minor role in diagnosing dementia. However, this dissertation indicates that language data can add to diagnostic criteria for AD and PPA. While language difficulties in healthy aging and dementia can be seen as a continuum, the use of new methods and better assessment tools may contribute to both diagnosis and suggestions for possible treatment.

Sammendrag

Studiene i denne avhandlinga undersøker hvordan endringer i språkbruk hos personer med Alzheimers sykdom og primær progressiv afasi (PPA) kan gi ny kunnskap om språkprosesseringsteorier, det mentale leksikonet og hvordan språkvansker ved demens kan gi en bedre forståelse av de komplekse symptomene som følger av disse sykdommene.

Dette er den første avhandlinga i Norge som fokuserer på språkvansker ved Alzheimer og PPA fra et lingvistisk perspektiv. Den tilfører ny kunnskap om leksikalsk produksjon og setningsforståelse for personer med Alzheimer og PPA, og ved normal aldring, samtidig som den undersøker strukturene til det mentale leksikonet hos disse gruppene. Resultatene gir støtte til bruksbaserte språkteorier og interaktive prosesseringsmodeller, der det mentale leksikonet betraktes som et strukturert nettverk av større og mindre enheter, organisert på ulike abstraksjonsnivåer. Videre støtter resultatene teorier som går ut fra at det flerspråklige mentale leksikonet er grunnleggende. Resultatene peker også i retning av at det finnes to typer logopenisk PPA.

Studiene viser også hvor viktig det er å ta i bruk flere ulike metoder når man utreder språkvansker, for å få et mer detaljert bilde av mulige vansker. Ved bruk av eye-tracker ble det funnet forskjeller i prosesseringshastighet mellom personer med og uten demens, som ikke ble fanget opp av en samtidig, "offline" oppgave. Analyser av frie ordassosiasjoner viser at det er ulikheter mellom gruppene som ikke kommer fram i tradisjonelle benevnelsestester.

Språkkartlegging er ikke en sentral del av demensutredninga i Norge, men resultatene fra denne avhandlinga viser at språkdata kan fungere som tillegg til diagnosekriteriene for Aalzheimer og PPA. Språkvansker ved normal aldring og demens kan sees på som et kontinuum, og bruk av nye metoder og bedre kartleggingsverktøy kan bidra til både diagnose og mulig behandling.

Acknowledgements

I wish to thank those who have contributed in special ways to the completion of this dissertation.

First of all I wish to thank my supervisors for support, encouragement, and for guidance. It has been inspiring working with all of you, and I hope we can do it again. I was very lucky that I got to start out with a great team of three — Valantis Fyndanis, Marianne Lind and Hanne Gram Simonsen; and then to finish with another great team of three — Minna Lehtonen, Monica Norvik and Hanne still. Without all of you, this project would not have happened. Your different backgrounds and expertise made me look at my own work from different angles, which I think it has benefited from.

I also wish to thank Hanne and Marianne, specifically, for recruiting me as a research assistant in the early days of my master's studies, and for introducing me to this wonderful world of research and clinical linguistics. for this I am extremely grateful.

I also wish to acknowledge all my participants and their families. Thank you for putting up with hours of testing and interviews. I also wish to thank my contact persons at the Memory clinic Anne-Brita Knapskog and Ingrid Tøndel Melbøen, as well as NorCog and Aldring og Helse for recruiting participants and granting access to register data. Prof. Roelien Bastiaanse served as my mid-way evaluator in April 2018, and really pushed me and the project in the right direction. Thank you for taking time to read and comment and for a very fruitful discussion.

Furthermore, I have been lucky enough to get the opportunity to discuss my project with some very inspiring researchers, and I really appreciate Loraine K. Obler and Mira Goral for welcoming me in their labs in New York, and Tomas Bak, Kees de Bot and Barbara Mertins for being available to me when you have visited Oslo.

A huge thanks goes to MultiLing center for multilingualism; thank you for letting me use your lab facilities, and for making me as an associated PhD fellow feel very welcome. Furthermore, I want to thank the Department of linguistics and Scandinavian studies for providing an inspiring work environment. I also wish to thank the MultiLing dementia research group, and of course the Research group for clinical linguistics and language acquisition, for including me in very interesting and challenging research projects over the years.

Thank you to all other PhD fellows and colleagues at ILN and MultiLing over the years, for making the tough times seem bright, and for sharing experiences. Many thanks, especially to Yesim Sevinc, Urd Vindenes, the NoSLiP 2017 group and my office mates, Jessica Hansen, Alexander Lykke, Hanna Solberg Andresen, Jorunn Simonsen Thingnes, Malene Bøyum, Hanna E. Menne and Linda Vatne.

My dear friends, you are the best support crew: Yvonne van Baal, Kjersti Bergersen, Mari Bjørkedal, Pia Edvardsen, Cathrine Einarsson, Sidsel Fjelltun, Pernille Hansen, Katharina Heinz, Anne Fougner Helseth, Elisabeth Holm, Tejaswinee Kelkar, Helene Killmer, Silje Mosgren — Thank you for getting me off my chair, into the pool, up on my bike, out around the lake; for facilitating tea breaks, bringing chocolate, sending encouraging texts, and just always being there for me.

Pernille Hansen, once more. Thank you for tech support and life support. For solving issues I didn't know I needed to get solved, and for being my academic big sister. Thanks also for lending out Thomas Nygreen and his inexhaustible knowledge about R. Also, a great thanks to Ekaterina Kuzmina for life-saving help with the eye tracking analysis, and for fun sessions in the lab.

I am also very grateful for Helge Dyvik at the University in Bergen, who taught me how to search in NorGramBank and find construction frequencies for my sentence comprehension experiment. Anders Ribu recorded those sentences, and Kira Del Mar made illustrations for my naming and matching tasks.

I would also like to thank my father and my brothers for keeping me grounded, and to my extended family — by blood and by choice — in Norway, Belgium and in Wimbledon. Thank you for supporting me through the years, and for always believing in me.

And most importantly, thank you Karel, for everything. Always. Life is so much more fun with you in it.

Oslo, October 2019

Ingeborg Sophie Bjønness Ribu

Contents

List of Figures

List of Tables

Part I

Synopsis

1
Introduction

This dissertation investigates how language skills are affected in healthy aging and by dementia, more specifically Alzheimer's disease and primary progressive aphasia. The focus is on both theoretical linguistic and clinical implications for change in language function in the two diseases. The dissertation consists of four research papers preceded by a summarizing text.

This chapter introduces the field of study, beginning with a short background section in 1.1 and an introduction to the motivation and purpose of the study in 1.2, where the research questions are also introduced. Section 1.3 introduces the field of study, dementia, and the two diseases in focus, Alzheimer's disease (AD) and primary progressive aphasia (PPA). Section 1.4 outlines what we can learn about language processing by studying language impairment. Finally, the structure of the remaining chapters is outlined in 1.5.

1.1 Background

As the world's population grows older, more people are at risk of developing dementia. The World Health Organization (WHO) postulates a growth of 10 million new dementia cases per year worldwide (WHO, 2016b). With rising numbers of dementia cases in the world, more research is needed on areas of life that are affected by this disease.

Cognitive and neural decline in dementia can be explained as a more "exaggerated" manifestation of what is found in healthy, non-pathological aging. This decline will affect all areas of cognition, including language production and comprehension. In this dissertation, I investigate the changes that occur in language production and comprehension in non-pathological aging and in Alzheimer's disease (AD) and primary progressive aphasia (PPA).

Many aspects of language do not change as we age. However, there are changes in language use and behavior which are related to aging. The most prominent change in language behavior with increasing age are difficulties with lexical retrieval and sentence processing in comprehension (Obler & Pekkala, 2008). Most of the existing research in the field has been conducted in English, and with this dissertation I hope to bring in perspectives from a lesser studied language — Norwegian. This can help broaden the general clinical picture, and support findings from other studies, as well as inform about the language specific issues that can be of interest to clinicians in Norway.

Language skills are often poorly or inadequately assessed during screening for dementia. This is because dementia assessment needs to incorporate several elements to gain a complete picture of the complex impairments on different levels — cognitive, daily living, depression, personality, language — and tests in each domain by itself only touch on the surface. Another reason is that the language assessment tools are not sensitive enough, nor have they been specifically developed to account for degenerating language impairment.

In this dissertation I will explore issues related to word-finding difficulties and sentence comprehension in two different kinds of dementia — AD and PPA — as well as in healthy aging. The dissertation has two main goals; 1) to study how language manifestations in the different diseases can help contribute to theories about language organization in the brain, and 2) to see how the study of language impairment can help shed light on the clinical picture of AD and PPA.

1.2 Motivation and purpose

As previously mentioned, language function is often poorly assessed when diagnosing dementia, even though most types of dementia will affect language production and/or comprehension. A more thorough understanding of the language changes that follow as a result of different dementia diseases will not only give the persons who receive such a diagnosis (and the people around them) more knowledge about what can be expected as the disease progresses, but can also contribute to differential diagnosis between different dementia types. Furthermore, there are theoretical-linguistic reasons to study language impairment following disease in, or damage to, the brain. Studying impaired language can give us valuable insight into how language is organized in the brain.

This dissertation is based on work from one overarching project, titled *Language and cognition in healthy aging and dementia*, which includes four separate studies, each with a different focus. These studies will be outlined in detail in chapters 4 and 5. Each study resulted in a paper intended for scientific publication, these can be found in part II of this dissertation.

The first article, study I (Ribu, Under revision), is a literature review of how different psycholinguistic properties affect lexical access in both production and comprehension. Study II (Ribu, Norvik, Lehtonen, & Simonsen, submitted), looks at how a test of free word associations can supplement traditional tasks for lexical access in assessment and research of language in dementia. The third study (Ribu & Kuzmina, submitted) is concerned with sentence comprehension in AD and PPA, and employs eye tracking methodology to study how different sentence types are processed in real time. The focus in study IV (Lind et al., 2018) is on longitudinal changes in language production in PPA. This is a single-case study of lexical retrieval skills over time, and in different languages for one person with a diagnosis that is most likely the logopenic variant of PPA.

Language impairment in dementia is often assessed by means of picture-based tests — though these are often not comprehensive enough to identify all aspects of language and communication that can be affected by dementia. In many countries, tests are merely translated from the original to the local language without any adaptations. New tools and methods, as well as good adaptations of already existing tools for studying language processing and retrieval, may help shed light on both clinical and theoretical aspects of language decline in healthy aging and dementia. It is important to note that translation equivalents of words are not always comparable across languages. Not only because contexts differ, and concepts that are expressed with one lexeme in one language might correspond to two or more different lexemes in another, but also because underlying psycholinguistic variables (such as frequency, age of acquisition, imageability, etc.) differ between languages (see 2.1.2).

To date, there is not much research that looks at the longitudinal changes in language behavior during the course of a dementia disease. This makes it difficult for persons with dementia and their next of kin to know what to expect as the disease progresses. Learning more about the longitudinal aspects of language impairment in AD and PPA will also be of value to speech-language therapists, clinical linguists, doctors and other medical personnel who work with these patients.

Most of the research on language in healthy aging and dementia is based on studies of oral language production. However, language comprehension is just as important for successful communication; yet we know far less about comprehension deficits than we know about production deficits in dementia. Single-word comprehension is often not impaired to the same extent as single-word production. Sentence comprehension deficits may be the result of general cognitive impairment, but the exact underlying difficulties are debated.

AD is the most common cause of dementia, and the more we learn about the different deficits that accompany this disease, the easier it will be to distinguish the cases that are in fact AD from the cases that are not. Unfortunately, some cases of rarer dementia diseases, such as PPA, are sometimes misdiagnosed as AD because of physiological similarities and lack of knowledge of the finer details that separate the diseases. Furthermore, there is some uncertainty surrounding the diagnostic criteria for the subtypes of PPA, which makes it important to continue to study the language manifestations in the different PPA subtypes to keep adding to the knowledge about these diseases.

In this dissertation, I will focus on the following three aspects, and answer research questions related to each:

1. **Linguistic aspects:** How do different psycholinguistic variables influence naming and comprehension in AD and PPA? (study I). What can word associations reveal about lexical retrieval difficulties in AD and PPA? (study II). How is sentence comprehension impaired in AD and PPA? (study III). How can data from language impairments in AD and PPA inform about theories of language processing? (studies I, II, III, and IV).

2. **Methodological aspects:** How can the use of different test methodologies to study language production and comprehension give a deeper insight to the language impairments in AD and PPA? (studies II, III, and IV).

3. **Clinical aspects:** How can language data be used to differentiate between dementia diseases? (studies I, II, III, and IV). How do naming impairments in dementia change over time? (study IV).

With this dissertation I hope to bring more knowledge about how language(s) is organized in the brain, and how it is affected in individuals with dementia.

1.3 Dementia

Dementia is a syndrome characterized by biological mechanisms that damage brain cells, resulting in cognitive decline and functional impairment. The first symptoms are often seen in episodic memory, but also in complex mental tasks. Early behavioral decline is gradual, and most basic abilities such as language and motor functions are relatively spared early on in the disease. This may make it difficult to date the real onset of the clinical symptoms (Lezak, Howieson, Bigler, & Tranel, 2012). Executive functioning deficits in early/mild stage dementia includes impairments in planning, reasoning, foresight and impulse resistance. Patients will have more problems with complex tasks involving planning and flexibility of thinking as the disease progresses (Lezak et al., 2012; Bayles & Tomoeda, 2007).

Contrary to common belief, dementia is not *one* disease, but rather a syndrome[1] that can be caused by a number of different diseases that lead to atrophy of the brain cells, and impairment in multiple cognitive domains. Depending on which areas of the brain are most affected by atrophy, the dementia disease will affect cognitive abilities differently.

[1]In medicine, a syndrome is a constellation of signs and symptoms associated with a morbid process, a set of symptoms that occur together (Bayles & Tomoeda, 2007)

Most diseases that cause dementia are progressive and not reversible. This means that a person who is diagnosed with dementia will progressively get worse as time passes. The most common disease that causes dementia is AD (see 1.3.1 below), followed by different kinds of fronto-temporal diseases, including PPA (see 1.3.2). Throughout the remainder of this thesis, dementia will be used as a collective term to refer to both Alzheimer's disease and primary progressive aphasia.

1.3.1 Alzheimer's disease

Alzheimer's disease (AD) is the most common cause for dementia, accounting for approximately 60-70% of all cases (WHO, 2016a). AD is most commonly recognized by impaired episodic memory, difficulty with learning, and difficulty with recalling recently learned information. In some cases, other domains are more affected than memory initially, these are so-called non-amnestic presentations of AD (McKhann et al., 2011). In these cases, the most common dysfunctions are found in language, visuospatial skills and executive functioning.

AD often originates in areas of the brain that are most commonly associated with episodic memory, especially in hippocampus and the basal forebrain (Braak, Braak, & Hohl, 1993). Once the disease progresses, working memory and semantic memory are also affected. The motor cortex is often spared (Farkas et al., 1982), which means that speech is fluent without signs of apraxia of speech or dysarthria (Bayles & Tomoeda, 2007).

In a study where caregivers were asked to specify which changes in language behaviour they noticed in the patient before the diagnosis of AD was made, they reported word-finding problems (anomia), difficulty naming objects, impaired comprehension of instructions, difficulty sustaining a conversation and problems completing sentences among others (Bayles & Tomoeda, 1991). A more detailed description of how language is affected in AD will be provided in 3.2.

1.3.2 Primary progressive aphasia

Primary progressive aphasia (PPA) is a neurodegenerative disease with, in most cases, semantic degeneration as the core symptom. Three different subtypes of PPA have been identified, and these can be distinguished from each other on the basis of language manifestations and underlying neural pathology (Gorno-Tempini et al., 2011). The three subtypes are: A logopenic variant of PPA (lvPPA); a non-fluent, agrammatic variant of PPA (nfavPPA); and a semantic variant of PPA (svPPA). The variant which is most often linked to AD, is lvPPA. svPPA and nfavPPA are more often associated with frontotemporal dementias.

In the remainder of this dissertation, extra emphasis is made on lvPPA as this subtype is more similar to, and often closely linked to, AD than the other two variants, and because all participants with PPA who took part in this project have lvPPA.[2] The terms lvPPA and PPA are therefore used more or less interchangeably, with lvPPA used when extra emphasis of the subtype is needed.

PPA is not to be confused with stroke-induced aphasia, as the underlying causes are different. In PPA, there is no lesion or brain trauma that causes the language impairment, but rather progressive cortical atrophy to a more or less confined region of the brain (Gorno-Tempini et al., 2011). The language decline in PPA stems from these progressive neuroanatomical changes, and not from injury.

The onset of PPA is slow, and manifests itself as a gradual, progressive impairment of language production, object naming, syntax or word comprehension, that is apparent in conversations as well as in speech and language assessment (Gorno-Tempini et al., 2011). A more thorough description of the language impairments in PPA is supplied in 3.3.

1.3.2.1 The types of PPA

In recent years, there has been some discussion regarding the classification of the PPA subtypes. Some researchers report that as many as 40% of all PPA cases are unclassifiable into any of the three types (Sajjadi, Patterson, Arnold, Watson, & Nestor, 2012; Machulda et al., 2013; Utianski et al., 2019). In many cases, the unclassifiable observations will later go on to develop lvPPA or nfavPPA. This indicates that there is some uncertainty related to at least these two PPA subtypes (Machulda et al., 2013).

Some effort has been made to challenge the current sub-classification system of PPA (Vandenberghe, 2016). Vandenberghe (2016) and Leyton, Ballard, Piguet, and Hodges (2014) argue that there is evidence for two types of lvPPA: one that resembles the non-amnesic variant of AD, with initial language manifestations; and one that resembles the originally described version of lvPPA (Vandenberghe, 2016; Leyton, Ballard, et al., 2014). Similar patterns have also been described by Rohrer et al. (2013) and Teichmann et al. (2013). Matias-Guiu et al. (2019) recognized two different types of lvPPA based on both language profiles and imaging data from a cohort of 68 patients with mild PPA (all three subtypes). These issues will be discussed again in more detail in chapter 6 (see 6.3.3).

1.4 What can we learn from studying language impairment?

Research on language deficits and impairment in dementia may contribute to the development of linguistic theories. Many theories of language processing build on evidence from language-impaired speakers, mainly on data from speakers with post-stroke aphasia.

Historically, the study of how language is impaired after a focal brain injury has served as

[2] The participant in study IV has a more uncertain diagnosis, but it is reasonable to assume that lvPPA is the correct diagnosis (more on this in chapter 5 and in the article).

evidence for left-hemisphere dominance for language. It is believed that the deficits observed in speakers with an acquired language disorder reflect the underlying cognitive architecture consisting of sub-components that may be selectively impaired by an injury or disease (Meuter, 2009).

A central question within neurolinguistics is: *if certain aspects of language are damaged and others not, following damage to or disease in the brain, what can this tell us about the organization of language in the brain*? Studying the language impairments that follow from an injury or disease in the brain can say something about this organization, as a common trait for persons who acquire aphasia or dementia is that they had a fully mature language system before they experienced either a sudden (aphasia), or gradual (dementia) deterioration. The assumption is that language impairment following damage to, or disease in the brain, is not random but depends on constraints determined by the structure of the premorbid system (Caramazza, 1986).

Studying the language of speakers with different kinds of dementia allows us to study the relationship between language and cognitive processes. The pattern of dissociation in dementia can provide valuable information of the dependencies between language and cognition (Obler & Gjerlow, 1999). Language impairments rarely occur in isolation, and are usually accompanied by impairments in memory, executive functioning or other cognitive domains. Cognitive and linguistic functioning should therefore be assessed together, to examine the relationship between language and other cognitive functions.

Traditionally, theorists have assumed that the mental lexicon (see 2.1.1) is monolingual by default, with an option for bilingual storage and processing. However, recently the tables have been turned (Goral, 2019; G. Libben & Goral, 2015). In later years, there has been a growth in studies of bilingual[3] dementia. Parallel to the studies of "monolingual" dementia, these studies have been used to contribute to knowledge on the bilingual organization of language.

Furthermore, there is also an ongoing debate about the bilingual advantage: the claim that persons who speak more than one language have larger cognitive reserve, and that this may delay the onset of dementia (Bialystok, Craik, & Freedman, 2007; Bialystok, Craik, & Luk, 2012).This discussion is outside of the scope of this dissertation, but is important to acknowledge it in a dissertation which focuses on language and cognition in aging and dementia.

1.5 Outline of the dissertation

This dissertation is divided into three parts; part I is a summarizing text that introduces the field of study, theoretical background, previous research, methods and materials used in the different studies, analysis and discussion. Part II consists of the four articles that were written to answer the research questions, and reach the goals outlined above. All appendices are collected in part III.

[3] In the remainder of this dissertation, I will use the term ‘bilingual’ rather than ‘multilingual’ to refer to speakers of more than one language, regardless of whether the number of languages the individual speaks is two or more.

The remainder of part I is structured as follows: The next chapter introduces the theoretical framework of the studies, and the third chapter summarizes previous research on language abilities in healthy aging, AD and PPA. In chapter 4, the methods, materials and plans for analysis of the four studies are outlined. chapter 5 introduces the research studies found in part II, highlighting some of the main findings from each. Chapter 6 offers discussions and conclusions related to the clinical and theoretical implications that can be taken away from the studies in relation to the research questions.

2

Usage-based linguistics and language processing

To properly account for language impairment in aging, we need a good theoretical framework that can explain how language is organized, stored and processed. Based on such a framework we can postulate models for language production and comprehension, and changes in language behavior throughout the lifespan.

In this chapter, I first outline the theoretical framework which serves as grounding for this project (section 2.1), then I introduce some hypotheses about language and aging (section 2.2), and follow on with some models of language processing (section 2.3).

2.1 A usage-based theory of language

The theoretical framework adopted in this dissertation is a cognitive, usage-based approach to language. The main features of usage-based theories of language are: that language is understood as domain-general, neurocognitive capacities that are shaped by individual usage-patterns and experiences; that there is no separation between lexicon and grammar; and that language is a dynamic system, in which various aspects of a language user's linguistic knowledge are constantly reorganized and restructured through use (Langacker, 1987; Taylor, 2002; Bybee, 2010; Diessel, 2017). Each of these aspects will be discussed below.

The first feature, that language is domain-general, implies that there are no brain regions that are involved in language processing alone — the brain areas that are involved in language processing are also involved in other cognitive processes, such as memory, attention, learning and motor planning, to name a few (Dick et al., 2001). All aspects of language are integrated parts of cognition, and rely on the same general mechanisms. This means that the processes which underlie language structure are not specific to language, but applicable to several cognitive domains, which makes language domain-general, rather than domain-specific (Bybee, 2010). Language is acquired and processed by means of general cognitive skills, such as the ability to categorize our experiences based on perceived similarities, and through a vast memory capacity (Bybee, 2001; Taylor, 2002; Langacker, 1987).

Furthermore, in usage-based linguistics there is no distinction between the lexicon and grammar. The lexicon is central, and grammatical structures are abstract representations derived from a language user's experiences with particular words or utterances (Bybee, 2010). To understand these abstractions, we first need to understand how the lexicon is built up and shaped through patterns of usage.

Language users store *tokens* of language — these can be single words, *chunks* of words or whole utterances — as *exemplars* in a rich mental network. This network is organized by *form*, *meaning*, *usage patterns* and the connections between these. Connections are formed between exemplars which are perceived as similar in form or meaning. These perceived similarities between exemplars give rise to hierarchical relationships between general *schemas* and their more specific *instances*. These generalizations are based on different levels of abstraction, within phonology as well as semantics (Langacker, 1987; Taylor, 2002; Bybee, 2010).

All tokens are stored in one rich memory, and map onto other exemplars that are similar to it in form and/or meaning. Representations of tokens are further *entrenched*, or strengthened, by other exemplars that map onto it because of these perceived similarities (Bybee, 2010). One important condition for this entrenchment is frequency of occurrence. Frequency strengthens the representations of linguistic units in the memory, and facilitates the activation and processing of words, chunks and *constructions*, which are "learned pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully general phrasal patterns" (Goldberg, 2006, p.3).

Following Bybee (2006), cognitive representations of grammar are organized into constructions which are partially schematic, conventionalized sequences of morphemes with a direct semantic representation. Langacker (1987) claims that there is no distinction between syntactic, morphological or phonological constructions, rather they are all emerging from generalized abstractions. When the formation of grammatical constructions are regular, these regularities are expressed in the grammar by a schematic symbolic unit (Langacker, 1987). Grammatical structures are entrenched patterns of usage, and motivated by frequency just like other exemplars in the lexicon (Martínez-Ferreiro, Bastiaanse, & Boye, 2019).

Since all exemplars are stored in one rich memory system, storage and processing becomes efficient, as we do not rely on the application of rules (i.e., for phonotactics and/or verb morphology) to select the correct output, but we can retrieve words, chunks, and constructions as whole units (Taylor, 2002; Langacker, 1987).

At every level of language, from phonology to syntax, there is evidence for rich memory representations: exemplar strength (mediated by frequency) means that constructions can easily be accessed and used for analogical extensions, or for the creation of new exemplars. (Bybee, 2010).

The third feature relates to how language is shaped by usage patterns and experiences over the lifespan. Throughout life, language will develop and change. For instance, the memory system expands when new exemplars are acquired and mapped on to existing exemplars, leading to a larger vocabulary. Change in the vocabulary is not only related to the number of words, but also to the content of words. Words will not only be more entrenched with experience, but also attract richer semantic representations, with more connotations and stronger network connections to other words (Simonsen, Lind, Hansen, Holm, & Mevik, 2013).

The mental lexicon belongs to individuals, and not to individual languages. Since the lexicon is individual, a range of factors will influence how the lexicon is shaped, for instance the level of education and the number of languages, dialects and social registers a person speaks can affect the architecture of the lexicon (Street & Dąbrowska, 2010; M. Libben, Goral, & Libben, 2017). This means that the mental lexicon is fully capable of handling more than one language at a time (G. Libben & Goral, 2015). Storage and processing in the mental lexicon are further discussed in the next section.

2.1.1 The mental lexicon

The mental lexicon should be conceptualized as a (dynamic) process, rather than a (static) entity, meaning that the mental lexicon is a manifestation of human capacity for lexical action and ability (G. Libben & Goral, 2015; Jarema & Libben, 2007). Lexical ability is fluid and variable, and there are substantial individual differences in the functional architecture of the mental lexicon which will be related to patterns of change across the lifespan, patterns of use and education, and related to the specific languages each individual maintains at all times. Thus, the mental lexicon accommodates different languages, dialects and situational social restrictions (G. Libben & Goral, 2015).

Word comprehension and production (both spoken and written) are lexical activities which take place in the mental lexicon (Jarema & Libben, 2007). Processing relates to how we activate items in the lexicon, and prepare them for production or comprehension. When a concept in the lexicon is activated for production or comprehension (lexical retrieval), this activation spreads to other semantically or phonologically related words. This spreading activation is an important premise for many usage-based theories of language processing (e.g., the *Spreading-Activation*

Theory (Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997), and in interactive theories of language and aging, such as the *transmission deficit hypothesis* (Burke & Shafto, 2004, 2008), see sections 2.2 and 2.3 below).

Grammatical structures, just like lexical ones, vary across registers, languages, contexts, groups of individuals and even within individuals. Usage-based theories of language propose one common storage system for everything from morphemes to complex syntactic structures. Languages are seen as communication instruments, and all grammatical structures are perceived as functional. This means that grammar is not autonomous from semantic and pragmatic function, but that grammar is conceptualization. Syntax is iconically motivated by function. For instance, constituency is a product of what belongs together semantically and what belongs together at the expression level (Martínez-Ferreiro et al., 2019). Furthermore, grammar emerges from generalizations over exemplars, and is maintained and shaped by usage. Grammar, like language in general, draws on domain-general neurocognitive capacities involved in social cognition, conceptualization and memory (Ishkhanyan, Sahraoui, Harder, Mogensen, & Boye, 2017; Boye & Harder, 2017; Martínez-Ferreiro et al., 2019).

Usage-based theories emphasize that heuristic and probabilistic factors account for language structure, rather than (morpho)syntactic rules and operations (Gahl & Menn, 2016). An example of one such probabilistic factor is verb bias. Some verbs are biased to appear in certain structures; verbs that typically appear in *passive* sentence constructions are so-called *passive-biased*. Passive sentences with passive-biased verbs (i.e. 'the candidate elected for government was pleased')[1] will be easier to process than active sentence with passive-biased verbs (i.e. 'The candidate elected to change the topic') (Gahl & Menn, 2016). This means that it is not necessarily the structure (active vs. passive) which makes processing difficult, but rather the context in which the verbs occur. In other words, both word and construction frequency are important for processing of different sentence structures.

Individual factors can also affect the entrenchment of exemplars; for instance, the level of education has been found to influence the capacity for syntactic comprehension of low-frequency structures. People with a higher level of education are better at understanding low-frequency sentence structures compared to people with a lower level of education. That is, low-frequency structures might be more entrenched for highly educated people (Dąbrowska, 2015).

2.1.2 Psycholinguistic variables

Several factors, or underlying psycholinguistic variables, pertaining to the form, meaning and usage patterns of words will influence how they are stored in, and accessed from the mental lexicon. This section introduces a few variables of importance, namely *frequency*, *age of acquisition*, *imageability* and *cognate status*. This is not an exhaustive list of all variables which can influence

[1]Examples borrowed from Gahl and Menn (2016).

lexical retrieval, since a thorough review of all psycholinguistic variables is beyond the scope of this dissertation. The variables which are introduced here are variables that have an important influence on both language production and comprehension in healthy aging and dementia, and thus are important for the discussion of the results of the four studies in this dissertation. The influence of these factors on naming and comprehension is further reviewed in 3.2 and 3.3. This is also the main focus in study I (Ribu, Under revision).

Psycholinguistic variables can affect lexical-semantic processing at different levels of the mental lexicon — i.e., at the conceptual, lemma or lexeme level (Vonk et al., 2019). These variables are to a large extent language-dependent, meaning that for example, frequency for one word might not be similar for a translation-equivalent of the same word in another language (see table 3.2 in chapter 3 for some examples). Likewise, age of acquisition and imageability may differ for the same concept across languages.

As previously mentioned, all language structures are results of entrenched usage patterns and entrenchment is a function of frequency, and thus of experience (Martínez-Ferreiro et al., 2019). Word and construction frequency are therefore critical variables which affect retrieval performance, as high-frequency items have stronger representations than low-frequency items and are therefore easier to retrieve from the lexicon (Bybee, 2001).

It is common to distinguish between *type* and *token* frequency. Type frequency refers to the number of different lexical items a certain construction is applicable to. Token frequency refers to how often specific items occur. Both types of frequency are important for processing; type frequency is important for productivity of patterns in the lexicon (Bybee, 2001), and token frequency is recognized as one of the most critical factors which affect naming performance. Neurologically healthy speakers name words with high token frequency faster and more accurately than words with lower frequency (Oldfield & Wingfield, 1965; Balota & Chumbley, 1984; Balota, Burgess, Cortese, & Adams, 2002).

How early words are learned in childhood, the *age of acquisition* (AoA) of words, is also recognized as an important variable which affects lexical access. Words that are learned early in life tend to be more entrenched, and often more frequent than words learned later in life (Juhasz, 2005). One explanation for the higher entrenchment of early learned words is that all new words attach to already known words, strengthening the connections and further entrenching the known items. Words that are used more will develop richer semantic representations, which again leads to stronger entrenchment (A. W. Ellis & Lambon Ralph, 2000; A. W. Ellis & Young, 1977; Juhasz, 2005).

Two different AoA measures can be distinguished; objective and subjective AoA. *Objective AoA* of words can be obtained by following children's development over time. *Subjective AoA* is obtained by asking adults how old they think they were when they learned a given word. This last method may seem far-fetched, but there is a strong correlation between objective and subjective AoA, and both can be used as valid measures for how early or late words are acquired (Hansen, 2016; Łuniewska et al., 2016; Juhasz, 2005).

Another factor which influences lexical retrieval is a word's imageability. This is a conceptual feature of words, and refers to the ease of which a word gives rise to a sensory mental image (Paivio, Yuille, & Madigan, 1968). Similar to subjective AoA, imageability measures are obtained by asking peopple how easily different words evoke a mental image (Simonsen et al., 2013).

Imageability is often related to concreteness, but there is no one-to-one relationship between the two. Words with high imageability ratings are most of the time, but not always concrete. For instance, 'ghost is a word which has a high imageability rating for Norwegian (Lind, Simonsen, Hansen, Holm, & Mevik, 2013), but it does not denote a concrete entity. The opposite is also true: the word 'armadillo' is a concrete noun which often has a low imageability rating. Subjective measurements, such as for AoA and imageability, are highly dependent on individual variation, and individual experiences. For instance, research shows that imageability ratings for words increase with age, due to older adults' richer semantic networks (Simonsen et al., 2013).

One form-based variable which is of importance for the current project is word similarity across languages, or the cognate status of words. Cognates are words with similar form and meaning between languages. However, the similarity between form and/or meaning can be more or less overlapping; for instance, the Norwegian word 'katt' and the English 'cat overlap in both form and meaning, whereas the words 'sykle' and 'cycle' are more overlapping in form than in meaning, as the Norwegian meaning is more specific than the English one, and can only mean the verb 'to ride a (bi)cycle', and not the noun 'a cycle' in the sense of a set of elements recurring over an interval.

All of these variables, and many others, influence how items are stored in the mental lexicon, how they relate to each other and how they are processed. Variables like these affect the entrenchment of words and the connections between words in the mental lexicon. Frequency was mentioned in 2.1 as an important factor which influences processing. The other variables are also associated with storage and processing in different ways. AoA influences entrenchment, as new words map on to already existing ones and strengthen the connections between known words in the lexicon, and imageability is associated with the richer semantic connections and connotations between words.

2.2 Hypotheses on language and aging

Language processing in older adults, and by extension also dementia, must be seen in relation to age-related cognitive change. Some hypotheses try to explain the language changes observed in aging by relating them to other changes in cognitive abilities. For example, the *Inhibition deficit hypothesis* (IDH) assumes that aging weakens the inhibitory processes associated with task-irrelevant information (Hasher & Zacks, 1988; Zacks & Hasher, 1994). As a result, older adults activate more irrelevant information than younger adults, and suppress less irrelevant information once it is activated. This means that aging impairs inhibition in all cognitive systems — including memory, language and attention — and this disrupts the use of relevant information (Hasher & Zacks, 1988; Ortega, Gómez-Ariza, Román, & Bajo, 2012).

Evidence for this theory comes partly from older adults' tendency to produce speech which is perceived as off-topic, or irrelevant (Arbuckle & Gold, 1993). IDH explains this by stating that older adults have a reduced ability to inhibit irrelevant information, which in turn makes it almost impossible to suppress thoughts which digress from the current topic. This will result in production of unrelated information or personal observations (Zacks & Hasher, 1994; Arbuckle & Gold, 1993).

Problems with inhibition are also found in tasks which measure executive control, where older adults show more interference from the incongruent color base-word in the Stroop color-naming task, and from distracting words in picture naming and sentence reading tasks, than younger adults do (i.e. Hasher & Zacks, 1988; Lustig, Hasher, & Zacks, 2007; Ortega et al., 2012, for an overview).

Hasher and Zacks state that inhibition is an essential component of both language production and language comprehension, which would indicate that older adults should be impaired on tasks which tap both lexical comprehension and production (Hasher & Zacks, 1988; Zacks & Hasher, 1994). However, as will be discussed in sections 3.1.1 and 3.1.2, there is an asymmetry in lexical comprehension and production abilities associated with aging, which cannot be accounted for by a theory which assumes similar impairment in the two modalities.

Another view of age-related language change is found in the *transmission deficit hypothesis* (TDH), proposed by Burke and colleagues (Burke & Laver, 1990; Burke & Shafto, 2004, 2008). This hypothesis assumes that language production and perception depend on how fast, and how much priming can be transmitted across the connections between different nodes in the language-memory system. A node is selected for activation only if the priming-level for that node reaches a critical difference separating it from other nodes in the same domain. Connections become stronger with use, especially recent use, but will weaken over time if not frequently used. Aging it self can also weaken the strength between connections more generally.

The TDH assumes that there is only one connection between a phonological node and each lexical node, but many connections between different lexical nodes. This makes the phonological nodes more vulnerable to break-down (i.e., transmission deficit) than the lexical nodes. However, for comprehension, this means that upon hearing a word, priming transmits via the phonological nodes to the lexical nodes. Transmission of priming within the lexical system is aided by the many connections which link related concepts to each other (Burke & Shafto, 2008; MacKay & Burke, 1990). This hypothesis predicts small or no age-effects for language comprehension tasks, but a large age-related effect on production tasks (Burke & Shafto, 2008). The retrieval impairments observed in aging are thus explained as a deficit in retrieving phonological, rather than semantic, information.

2.3 Models of language processing

Most models of language processing operate with at least two levels of representation; one level for semantics and one for phonology/orthography. The main difference is in the way they conceptualize the relationships between these levels, either through serial or interactive activation. Serial models postulate that there is one-way activation between meaning and form, or the other way around. Interactive models, on the other hand, assumes that there is interaction in both directions. In the next sections, some common models for language processing which are relevant for this present thesis will be presented. First, some general models for production and comprehension are outlined in 2.3.1. Then, in the section regarding comprehension (2.3.2), I also introduce some theories of sentence comprehension, which is of particular importance for the third study in this dissertation. The final section (2.3.3) introduces a recent alternative language processing model which incorporates both production and comprehension.

2.3.1 Production models

One of the most influential models of language processing — called *A blueprint for the speaker* — was put forward by Levelt (1989). In this model, it is assumed that activation is unidirectional, with no feedback between different levels of processing. Following this model, language processing happens in different steps: In the *Conceptualizer* (the first step), the intended concepts are selected from the mental lexicon. In the second step, the *Formulator*, the phonological and grammatical form is selected, before the mapping of the phonological form to the concept takes place in the third step, the *Articulator*. The model is presented in figure 2.1.

Levelt's model is recognized as one of the most comprehensive models of speech processing, but it has been criticized for the uni-directional, top-down view of language processing (Dell, 1986). As an alternative to this serial model, Dell (ibid.) proposed a model based on the spreading activation principle. In this model, interaction spreads between the different levels of representation (i.e., phonological, semantic etc.), and the different levels are active at all times (Dell, 1986; Dell et al., 1997).

Since there is activation on several levels at the same time, activation at the phonological levels of the production system feeds back to the semantic level, which activates semantic representations which in turn reinforce activation of the phonological level (Dell et al., 1997). An illustration of how this activation spreads across levels is seen in figure 2.2.

Interaction is necessary to account for mixed-error effects (slips of the tongue where similarity in form tends to increase the probability for semantic substitutions in naming), which is taken as evidence for the simultaneous activation of semantic and phonological information. Activa-

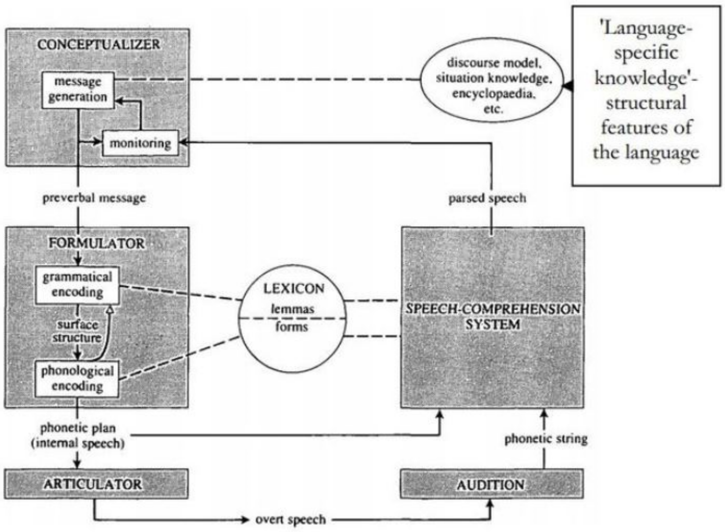


Figure 2.1: The blueprint for the speaker model (Levelt, 1989)

tion thus flows from target word nodes to phonemes, and then to mixed 'neighbors' (Dell et al., 1997).

Interactive models, where the structures are not predetermined, but shaped by feedback, are compatible with usage-based theories of language (Bybee, 2001).

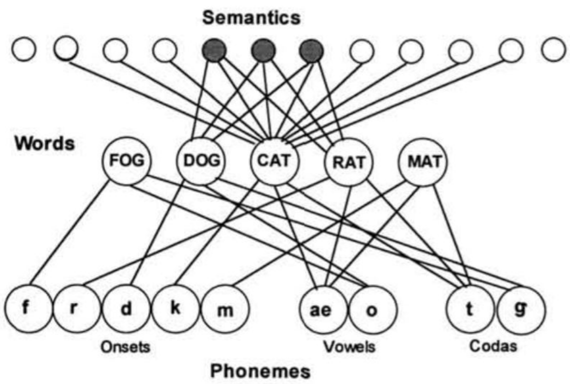


Figure 2.2: The spreading activation theory (Dell, 1997)

2.3.2 Comprehension models

Similar to models of speech production, models of speech comprehension can either be serial or interactive. For instance, the *Cohort model* (i.e., Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1980) assumes three stages of language comprehension that follow each other temporally: The *Access* level, *Selection* level and *Integration* level (see figure 2.3).

Phonemes are received on the access level, and all items which start with the same initial phoneme structure are activated, creating a cohort of candidate words. As more and more of the phonemes are received by the listener, the target word is selected by a process of elimination once competitor words that do not share the sound structure of the target word have been excluded.

At the second level (Selection), the listener uses context, semantics, recency of use and frequency to narrow down the candidates before choosing the word with the best fit on the third level (Integration). This model is uni-directional, and the higher levels do not interact with the lower levels, but rather rely on them for further processing (Marslen-Wilson, 1987).

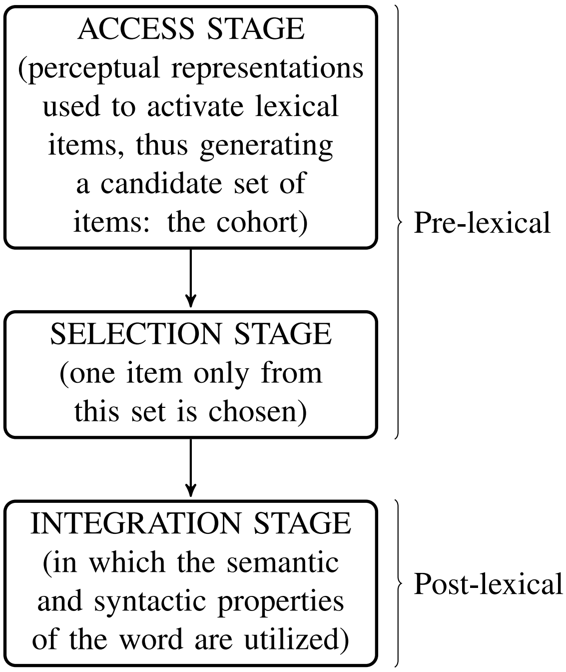


Figure 2.3: Schematic overview of the Cohort model

This feed-forward view of the Cohort model means that the model cannot account for how listeners recognize words which mismatch acoustically or contextually (Tanenhaus, Magnuson, Dahan, & Chambers, 2000), since later levels are dependent on the previous ones and there is no option for retracing and reactivation.

In contrast, the *TRACE model* (McClelland & Elman, 1986) is a dynamic processing structure made up of a network of units, which performs as the system's working memory as well as the perceptual processing mechanism. TRACE is also divided into three levels; the *Feature* level, *Phoneme* level and the *Word* level. Each of these levels relate to a particular perceptual object occurring at a particular point in time, relative to the beginning of the utterance (McClelland & Elman, 1986).

In this model, word elements are organized in a network. The likelihood of successful word recognition is influenced by excitatory connections at both lower levels (features and phonemes) and higher levels (sentential aspects) of representation. Selection of a target word is defined by competition between activated nodes. The node which receives most excitation will win out and be selected (McClelland & Elman, 1986). According to this model, the mind uses physical acoustic features, phonemic information and semantic information to match what has just been heard to a word in the mental lexicon (McClelland & Elman, 1986).

Speech recognition is complicated by the complex nature of speech signals. The model tries to account for the following "problems" with speech perception: 1) the temporal aspects of the speech signal, 2) overlapping phonemes and words, 3) context-sensitivity of cues, or the fact that articulation of phonemes is affected by the sounds that come before and after it, and 4) noise and indeterminacy in the speech signal (McClelland & Elman, 1986). These issues can be accounted for by interactive models which assume that the different processing levels affect each other, and that activation flows both upwards and downwards within the network.

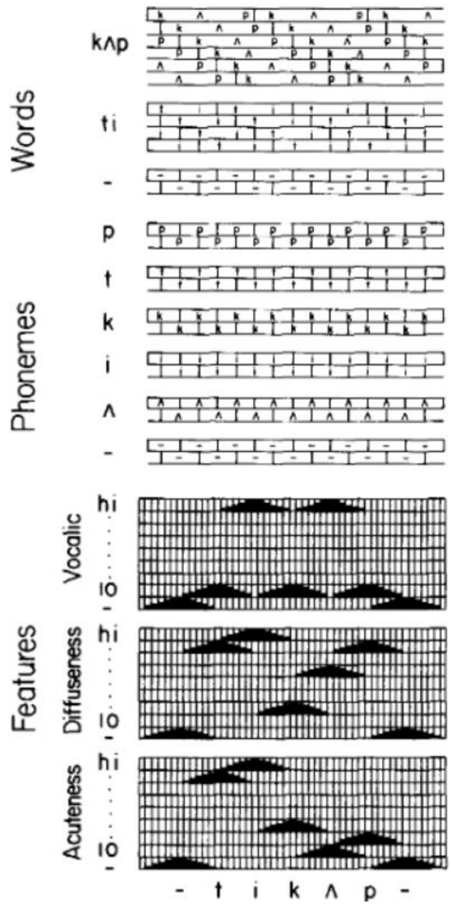


Figure 2.4: The TRACE model (McClelland and Elman, 1986)

Figure 2.4 shows a subset of the Units in the TRACE model, where each rectangle represents a different unit. The labels indicate what each unit represents. The horizontal edges of the rectangle indicate the portion of the TRACE which is covered by each unit. This specific image shows the feature specifications for the phrase "tea cup", preceded and followed by silence.

2.3.2.1 Sentence comprehension

Sentence comprehension is a complex task which requires comprehension of the individual words in the sentence, mapping of thematic (i.e., syntactic) roles onto grammatical structures, establishing the time-frame of the sentence, etc. In the following paragraphs, I introduce a few theories which try to explain how complex grammatical structures are comprehended and processed.

Some theories of sentence processing, for instance the *constraint-based model* (Trueswell & Tanenhaus, 1994) take statistical and probabilistic aspects of language, such as frequency of structures, into account when they explain how sentence processing is affected. Parallel to the frequency effects on the lexical level, the relative frequency of constructions affects sentence parsing (MacDonald, Pearlmutter, & Seidenberg, 1994; Jap, Martinez-Ferreiro, & Bastiaanse, 2016). Despite the fact that lexical frequency is recognized to be an important factor in language processing, frequency of grammatical constructions has not been incorporated into theories which explain comprehension of different syntactic structures in speakers with acquired language disorders, such as aphasia and dementia (Jap et al., 2016; Gahl & Menn, 2016).

Other models, such as the *Derived Order Problem Hypothesis* (DOP-H), assumes that sentences with derived word order require more processing capacities than sentences which follow base word order (Bastiaanse & van Zonneveld, 2006). The derived sentence order is, following Bastiaanse and van Zonneveld (ibid.), the result of syntactic movement operations. Since older adults and persons with dementia have reduced processing capacities compared to younger adults (see Park & Reuter-Lorenz, 2009), sentences with derived word order are expected to be even more difficult to process for them. An important premise for this hypothesis is that the derived sentences are harder, but not impossible, to process (Bastiaanse & van Zonneveld, 2006; Jap et al., 2016).

In a recently proposed statistical modeling account for language processing, Frank and Yang (2018) suggest that hierarchical syntactic operations (i.e., movement) are not necessary to explain sentence comprehension; relying on lexical properties of the stimulus will suffice (Frank & Yang, 2018). The authors argue that sentence comprehension requires at least some knowledge of word meaning, and constructing a sentence's hierarchical structure requires information about the word's possible syntactic categories (e.g., if a word can be a noun, a verb or an adjective in a given situation) (Frank & Yang, 2018). The only linguistic information which was available in Frank and Young's model resided on the lexical level, there was no phrase- or sentence level processing, only representation of lexical information (Frank & Yang, 2018; Frank & Christiansen, 2018).

This view is compatible with a usage-based view of language, where there is no distinction between the lexicon and grammar, and lexical items are seen as central. Models of sentence processing do not need to be based on a notion of movement and syntactic operations if statistical and probabilistic aspects of language, such as frequency of structures, are taken into account.

Statistical preferences for certain structures are used by the cognitive system for learning how to comprehend and produce utterances (Frank & Christiansen, 2018).

2.3.3 The Multilink model for production and comprehension

Not many models try to incorporate both production and comprehension, and even fewer do so while at the same time considering a bilingual mental lexicon as default. Most models of language processing see monolingualism as the norm, and define bi- and multilingual processing, and the bilingual mental lexicon, in relation to this monolingual default. But, as discussed above, the mental lexicon is driven by individual experiences, and is fully capable of handling more than one language at the same time (G. Libben & Goral, 2015).[2]

Studies of the bilingual mental lexicon have traditionally been based on the assumption that a multilingual speaker has two or more mental lexica which partially overlap. However, this hinges on the assumption of the monolingual norm and the bilingual exception. Recent advances in the study of multilingual processing claim that there is one mental lexicon which is shared for both or all languages, and that the bilingual mental lexicon thus includes the monolingual lexicon (G. Libben & Goral, 2015; M. Libben et al., 2017).

The *Multilink* model for language processing incorporates a unified account for (bilingual) word comprehension, lexical-semantic processing and word production (Dijkstra et al., 2019b). The model is the first to consider processing aspects of word production, comprehension and translation, and addresses how cognates can be processed (Goral, 2019). The model takes psycholinguistic variables into account, such as frequency, also when frequency is dependent on language proficiency and language exposure. Furthermore, it is also the first comprehensive model of language processing which explicitly states that the multilingual lexicon should be considered as the norm (Dijkstra et al., 2019b; Goral, 2019). This model is interactive, and postulates that different levels can be activated at the same time, and that activation spreads both forward and backward between levels (Dijkstra et al., 2019b).

Multilink is modelled as a lexical network where an input word activates various representations. These representations in turn activate their semantic and phonological counterparts, as well as associated language membership representations. All activation in Multilink is bi-directional (Dijkstra et al., 2019b). Figure 2.5 shows how this network is structured. One important aspect to note in this figure, is that the languages are represented as two separate nodes, even though the model argues for an integrated lexicon. Goral (2019) suggests that we do not yet have the proper experimental tools and/or the theoretical terminology to distinguish clearly between an "integrated one" and "interconnected two" lexical systems.

[2] In this dissertation, I assume that the multilingual lexicon is the default, and argue that the "monolingual" participants in the studies reported in part II are not, in fact monolingual, but rather use mainly one language for most purposes of their daily linguistic activity (see 4.3.1).

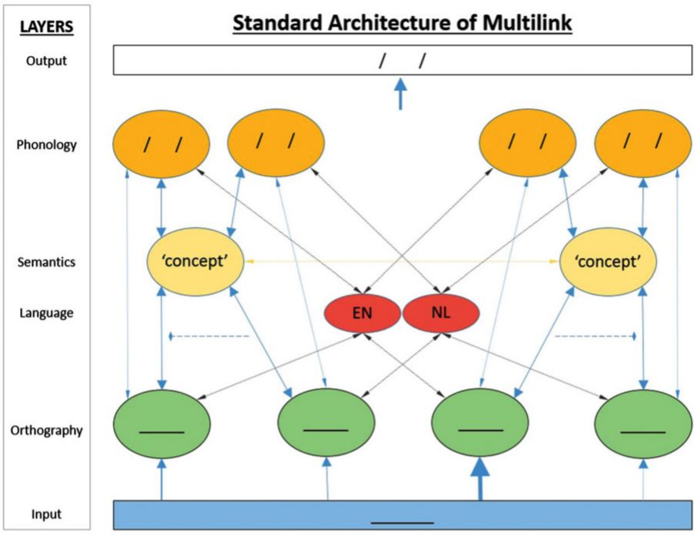


Figure 2.5: The Multilink model (Dijkstra, 2019)

Multilink simulates interaction between several codes (orthographic, phonological and semantic) at an interval measurement level. The model assumes that word retrieval involves language non-specific processing. The bilingual lexicon is integrated in the model, which means that there is only one storage for words from different languages. Furthermore, there is a link between translation equivalents only on the semantic level (Dijkstra et al., 2019b).

The model includes a task/decision system, which allows it to simulate word processing in psycholinguistic tasks such as lexical decision, orthographic and semantic priming, word naming and word translation production. The task/decision system checks the language membership of the input and output, and the degree of activation on phonological/orthographic and semantic levels as a requirement for the release of a response. Word naming and word translation happen when the phonological representation of the target language reaches a certain critical threshold. For word naming, the target and input languages are the same, whereas for word translation, target and input languages are opposite (Dijkstra et al., 2019b).

The input level is indicated by the blue line at the bottom, and orthographic representation by the green circles. Phonological representations are noted as slashes (/ /). Output is task dependent; in this illustration, the slashes indicate a phonological output in either language (L1 or L2). (see Dijkstra et al., 2019b, for more detailed information). The figure does not depict the production component in a satisfactory manner, but simulations of bilingual naming data performed on Multilink correlated well with empirical data (Dijkstra et al., 2019b).

Currently, Multilink is based on data from English and Dutch, and simulations have been run mainly on comprehension data rather than on production data, but the authors claim that other languages and language data can be implemented to further develop the model (Dijkstra et al., 2019b).

This model fits the theories outlined in sections 2.1 and 2.1.2, since it takes different underlying variables into account, especially frequency — even when this is influenced by language proficiency — and assumes that the multilingual lexicon should be seen as central. Furthermore, excitation between nodes spreads forwards and backwards in the system, giving activation at different levels of representation.

The model's merit lies in the fact that it assumes an integrated lexicon, and accounts for both production and comprehension. However, certain extensions and improvements can still be made: for instance, the model does not account for inhibition and suppression of non-target language representations, it is based on languages with very similar scripts and it lacks integration beyond the word-level. Furthermore, the primary infrastructure of the model is based on orthographic input, and there can be differences in written and spoken input which are not clearly accounted for in the model's current form (Goral, 2019; Ivanova & Kleinman, 2019; Van Hell, 2019; Dijkstra et al., 2019a).

The models and theories outlined in this chapter are all considered to be relevant for the studies in the current project. They will be discussed and evaluated in the discussion in 6.6.

3
Language and aging

In this chapter, previous research on language decline in healthy aging, AD and PPA is outlined in sections 3.1, 3.2 and 3.3 respectively. In section 3.4, I outline how language and cognition are traditionally assessed in dementia, both for diagnostic purposes and research.

It is important to note that when discussing language change in AD and PPA, these refer to generalized observations. Depending on the individual neural pathology, the impact of the language impairment will vary across patients, and two people with the same diagnosis may have different language profiles.

3.1 Language in healthy aging

Language abilities are said to not change much across the lifespan (Wingfield & Stine-Morrow, 2000), but some aspects of language processing are influenced by aging (Obler & Pekkala, 2008; Obler et al., 2010). Most notably, there is a change in naming abilities, and most adults will be familiar with the feeling of not being able to come up with the correct names for objects and persons. Furthermore, older adults may have more difficulty interpreting complex sentence structures compared to younger adults. This is normal in aging, and does not necessarily imply a cognitive or language impairment. However, as both naming impairments and sentence comprehension difficulties are common in AD and PPA, it is important to identify the changes that are considered "normal" in aging and those that can be indicative of an impairment.

Furthermore, general cognitive abilities, such as episodic and short-term memory, attention and inhibition will start to decline with advanced age (Park & Reuter-Lorenz, 2009; Goral, 2004). Changes in language as a result of aging may therefore be due to dependence on other cognitive factors that decline with age.

3.1.1 Language production in healthy aging

Though our vocabularies continue to grow throughout our lifespan, our lexical retrieval abilities reach a peak at around age 30 before they start to decline (Connor, Spiro, & Obler, 2004; Goral, Spiro, Martin, Obler, & Connor, 2007; Obler & Pekkala, 2008). Longitudinal studies have found the most drastic decline to happen after age 70 (Au, Albert, & Obler, 1989), but initial signs of decline have been found for people in their fifties (Au et al., 1995), and even as early as in the late thirties (Connor et al., 2004). This might be explained by a general slowing in processing with increasing age (Salthouse, 2010), or because it takes longer to search through a larger vocabulary (Vitevitch & Luce, 1998). When naming becomes difficult, one compensatory strategy can be to name a synonymous word, or overuse semantically "empty" words like ‘thing’ or ‘do’ to not interrupt speech flow (Adlam, Patterson, Bozeat, & Hodges, 2010).

Older adults are more prone to so-called "tip of the tongue states" (TOT) than younger adults (Astell & Harley, 1996; Burke, MacKay, Worthley, & Wade, 1991). This is the term used to explain the feeling of knowing a word without being able to articulate it. Results from naturalistic diary-keeping studies show that older adults have far more TOTs than younger adults (Burke & Shafto, 2008, 2004). TOTs can also be experimentally induced, often by asking trivia-like questions such as "What is the name of the river that runs through Rome?"

In a study by Burke and colleagues applying this latter paradigm (Burke & Laver, 1990; Burke et al., 1991), participants were asked questions about the real world, and to rate how certain they were that they knew the answer to the questions. They were also asked if they were able to recall the answer. The results showed that older adults had a significantly larger proportion of TOT states than younger adults, especially for proper nouns (Burke et al., 1991).

Decline in lexical retrieval skills is most often associated with proper noun retrieval, but difficulties have also been reported for other types of words and linguistic units (Obler & Pekkala, 2008). The main difficulty with lexical retrieval seems to be access to the phonological form, rather than impaired semantic knowledge. This can be seen in studies where cues are given to aid retrieval. Semantic cues are usually not very helpful for older adults, whereas phonological cues can be very beneficial and aide lexical retrieval (Obler & Pekkala, 2008).

Language production is more than just naming words in experimental settings, and it is important to study language in use to learn more about how the naming difficulties affect speech and communication. Research finds that sentence production is less affected by increasing age than naming in isolation, and syntax seems to be well preserved. If people experience difficulties with sentence generation, this is usually at the word level. (Kavé & Goral, 2016b; Peelle, 2018;

Kemper, 2012).

However, some changes have been observed in narrative production when comparing younger and older speakers. In tasks where people are asked to retell a recently heard story, older adults will often provide shorter answers than younger adults. In contrast, if they are asked to produce free narratives, older adults' narratives are longer than those of younger adults, but they contain fewer cohesive ties, more irrelevant information and are perceived by raters as less "dense" (Juncos-Rabadán, Pereiro, & Rodríguez, 2005).

The difference in story length can be related to speech fluency, and to more frequent and longer pauses. Researchers have found that older adults have an increased number of pauses in their speech, and that these are typically longer than for younger adults (Burke & Shafto, 2004; Tyler, Russell, Fadili, & Moss, 2001). This adds to the perception of older adults' slower and more hesitant speech. More frequent pausing can be seen as a compensatory strategy for older adults, to allow them self to search for appropriate alternatives when the initial word search is unsuccessful (Mack et al., 2015).

For speakers of more than one language, the same general picture is found as for monolingual speakers. However, the decreasing language abilities can develop later than for monolinguals, or they can affect one language more than the other(s) (Goral, 2013). Some people may experience *language attrition*, which is defined as the loss of one language, often the speaker's L1 (Schmid & De Bot, 2004).

In sum, lexical retrieval difficulties in production are apparent in older adults, especially in experimental settings where single words are assessed. Single word retrieval is also one of the issues that can be observed in naturalistic settings, and one that is reported on by speakers in diary studies. It is also clear that older adults do perform differently from younger adults on tasks that are based on more naturalistic speech output.

3.1.2 Language comprehension in healthy aging

Difficulties with language comprehension associated with aging are not observed on the single-word level, like we see for production difficulties. However, comprehension deficits in healthy aging have been reported on sentence and text level (Obler & Pekkala, 2008). These problems are often not exclusively language-related problems, but problems that originate from an interaction between language and other cognitive changes that are associated with aging — such as processing speed, attention and memory (Goral et al., 2011).

In studies that look at sentence structures that require a lot of inference, researchers find that such sentences are comprehended more slowly than sentences with less complex sentence structures. For instance, Kempler, Almor, and MacDonald (1998) and Ferreira (2003) found that older adults have more difficulty understanding complex sentences, for example, sentences that do not follow basic word-order, or longer sentences that include more than one finite verb. In some cases, the difficulty lies in processing pronouns, and deciding who is doing what to whom (Almor, Kempler, MacDonald, Andersen, & Tyler, 1999). Marked sentence types, such as passives, relatives or sentences with double negation, are especially difficult to comprehend (Caplan, 1996; Waters, Caplan, & Rochon, 1995; Rochon, Waters, & Caplan, 1994). In a study that compared older and younger participants who were matched on verbal and general intelligence and working memory, Zhu and colleagues found that older adults had significantly lower accuracy scores on a semantic and syntactic acceptability judgment task compared to younger adults (Zhu, Hou, & Yang, 2018).

When sentence processing becomes difficult, more executive functions are called upon (Goral et al., 2011), which begs the question of why sentence comprehension deficits are not more common in healthy aging, given the natural decline in memory functions during aging (Tyler et al., 2010).

Other issues that often affect language comprehension in older adults relative to younger adults, are hearing loss and reduced auditory acuity (i.e., difficulties with hearing certain sound frequencies). These issues often affect language comprehension in sub-optimal settings, such as noisy environments, or conversations including several participants (De Bot & Makoni, 2005).

To sum up, comprehension of syntactically complex sentences, and dense text material can be impaired in older adults, but single-word comprehension is usually not. However, it is difficult to disentangle the linguistic deficits from other cognitive deficits, such as reduced processing speed and changes in attention and memory. Physiological changes, such as hearing loss, may also contribute to deficits in spoken language comprehension.

3.2 Language in Alzheimer's disease

Changes in language behavior can be an early symptom in AD, specifically changes related to lexical retrieval. Anomia is often prominent early on in the disease, and other language abilities will be gradually affected as the disease progresses (Damico, Müller, & Ball, 2010). In the most severe cases of AD, patients' abilities to partake in meaningful communication are severely impaired, and in some cases almost impossible. The verbal output is often reduced to repetitions of their own or other's speech, production of meaningless sounds, and in extreme cases mutism (Sloot & Jonkers, 2011).

The first linguistic abilities that are affected are verbal fluency, auditory comprehension, and reading and writing comprehension (Tsantali, Economidis, & Tsolaki, 2013). Syntax seems to be relatively well spared, although this has mainly been investigated in oral production studies (Rochon et al., 1994; Lee, Yoshida, & Thompson, 2015).

3.2.1 Language production in Alzheimer's disease

Anomia is often apparent in both spontaneous speech and confrontation naming early on in the disease, and it often affects proper names before common names. The word finding difficulties will worsen throughout the course of the disease, making communication more and more difficult with time. In extreme cases, the anomia can even lead to mutism (Sloot & Jonkers, 2011).

Anomia in AD has been explained as either an underlying semantic impairment, or as an impairment in access to semantic information (Kavé & Goral, 2016a). An underlying semantic impairment could imply that some words are completely inaccessible from a speaker's mental lexicon, causing problems with both production and comprehension. However, if the impairment lies in access during production, speakers will still be able to recognize and comprehend words that they might not be able to articulate, and will still be able to provide descriptions of the word or produce synonyms and other semantically related words.

Some researchers find that persons with AD have more problems with naming nouns than verbs (White-Devine et al., 1996; Druks et al., 2006), while some find the opposite pattern (Williamson, Adair, Raymer, & Heilman, 1998; Hernández, Costa, Sebastián-Gallés, Juncadella, & Reñé, 2007). Almor and colleagues found that persons with AD have difficulties with pronouns (Almor et al., 1999). Others find a dissociation between animate (for instance, animals, fruits and vegetables) and inanimate objects (tools, furniture, vehicles etc.), where one category can be better preserved than the other (Zannino, Perri, Pasqualetti, Caltagirone, & Carlesimo, 2006; Cuetos, Arce, Martínez, & Ellis, 2017). All of these, and other underlying word properties (i.e. frequency, AoA, imageability and word structure, see 2.1.2) can influence how easy or hard words are to retrieve from the mental lexicon, both in terms of production and comprehension (Cuetos, Rodríguez-Ferreiro, Sage, & Ellis, 2012; A. W. Ellis & Lambon Ralph, 2000; N. C. Ellis, 2002).

Persons with AD tend to retrieve high-frequency words more easily than low-frequency words, and these are often more general, semantically empty words, such as 'thing' or 'do, or the use of superordinate category-names rather than more specific target words (i.e. 'bird' for 'penguin)[1] (Domoto-Reilly, Sapolsky, Brickhouse, & Dickerson, 2012). This can be attributed to the fact that their semantic network is impoverished, and more general lexemes are more easily retrieved due to their higher frequency and early AoA (Adlam et al., 2010).

On verbal fluency tasks, where one is instructed to name as many words they know belonging to a phonological or semantic category within a given time frame, persons with AD produce far fewer items than healthy participants, and the difference is greater for semantic than for phonemic fluency. Persons with AD also produce more repetitions of the same words, which may be related to a decline in working memory (Marczinski & Kertesz, 2006; Vita et al., 2014).

Analyses of speech samples show that persons with AD differ significantly from healthy

[1]Example from the picture naming task used in studies II and III.

controls on several measures: the respective amount of content words, nouns, and pronouns, as percentages of the total number of words (Kavé & Goral, 2016b); articulation rate; speech tempo and hesitation ratio (Hoffmann et al., 2010). Persons with AD produced fewer content words than controls, and a smaller percentage of the total words were nouns. Persons with AD also produced a higher number of pronouns than healthy controls. A finer analysis of the words produced by the two groups showed that persons with AD produced shorter words, and words of higher frequency than the control group did (Kavé & Goral, 2016b). The temporal measures can help distinguish both between persons with AD and healthy controls, but also between persons with mild, moderate and severe AD (Hoffmann et al., 2010; Martínez-Sánchez, Meilán, García-Sevilla, Carro, & Arana, 2013). More frequent, and longer, pauses in connected speech may be seen as compensatory mechanisms to overcome lexical retrieval difficulties (Pistono et al., 2019). Sentence production in AD is often short, slow and characterized by an overuse of empty words and a decline in relevant information content (Taler & Phillips, 2008; Fraser, Meltzer, & Rudzicz, 2015; Szatloczki, Hoffmann, Vincze, Kalman, & Pakaski, 2015).

In early stages of the disease, apraxia is not a common observation, but may develop in later stages. Dysfluencies, paraphasias, bizarre word combinations and intrusions are common mid-stage speech defects (Bayles & Tomoeda, 2007). In very late stages speech becomes non-fluent, repetitive, and largely non-communicative, with many patients displaying partial or pure mutism (Bayles, Kaszniak, & Tomoeda, 1987).

There is a growing body of research on bilingual dementia, but the results from this field show mixed results. In general, the majority of the studies on bilingual dementia show that persons with AD perform better in their dominant language on a range of tasks and measures. However, the pattern of performance across languages was similar for bilingual speakers with AD, suggesting that both languages are affected (Stilwell, Dow, Lamers, & Woods, 2016). It is important to note that the persons' first language is not always the dominant language. Both Gollan, Salmon, Montoya, and da Pena (2010) and Gómez-Ruiz, Aguilar-Alonso, and Espasa (2012) reported that their participants performed better in their L2 on measures where they had been more proficient in that language (i.e., literacy in the study by Gómez-Ruiz et al. (2012), and on the Boston Naming Test in Gollan et al. (2010)).

In sum, persons with AD show an early and gradual deterioration in lexical access abilities, which influences naming in isolation as well as in connected speech. Reduced lexical richness in context might be an indication of a cognitive impairment or early indication of AD, and these changes should be recognized as an important diagnostic feature for these patients. However, to date, there is no linguistic test or test battery that can successfully diagnose Alzheimer's disease on its own, and evaluation of linguistic parameters should be used in combination with other diagnostic measures (Szatloczki et al., 2015).

3.2.2 Language comprehension in Alzheimer's disease

Language production has been studied more extensively than language comprehension in AD. The study of language comprehension in AD has in recent years mainly been focused around syntactic comprehension, since single-word comprehension is less impaired than single-word production.

In most cases, semantic knowledge and word comprehension are spared early on in the disease. However, as the disease progresses, language comprehension may also deteriorate in line with other language and cognitive functions, such as auditory comprehension and reading (Bayles & Tomoeda, 2007).

Although most research concludes that single-word comprehension is spared in AD, some researchers found that persons with AD may have impairments in single-word recognition and/or comprehension relative to neurologically healthy persons (Hodges & Patterson, 1995; Hodges, Patterson, Graham, & Dawson, 1996; Cuetos, Herrera, & Ellis, 2010; Cuetos et al., 2017). These studies are based mainly on findings from lexical decision or word recognition studies.

Cuetos and colleagues too find some impairment in single-word comprehension (see Cuetos, Gonzalez-Nosti, & Martínez, 2005; Cuetos et al., 2017; Cuetos, Rodríguez-Ferreiro, & Menéndez, 2009), however, these impairments are less prevalent than sentence comprehension difficulties. Sentence comprehension can be affected by both a general processing decline (working memory impairment) or by a failure to comprehend the individual parts of sentences when these are presented in marked order (Croot, Hodges, & Patterson, 1999; Waters & Caplan, 1997; Caplan, 1996).

Some studies have shown that patients with AD have more difficulty comprehending more complex sentences (see Lee et al. (2015) for an overview), although the notion of what constitutes a complex sentence is not consistent across the studies. Some authors measure complexity by the number of words in the sentence (Tomoeda, Bayles, Boone, Kaszniak, & Slauson, 1990); some consider complex sentences to deviate from the base syntactic structure in the language (Rochon et al., 1994); while others discuss complexity by the number of propositions (finite verbs) (Croot et al., 1999; Waters et al., 1995). All of these issues do contribute to sentence complexity, often interacting with each other; shorter sentences with only one proposition in the active voice are easier to parse than longer sentences with more propositions and maybe even deviant syntactic structures, which will require more working memory capacity to parse correctly.

Meyer and colleagues (2012) found that it was the number of propositions (two or more) which made the sentences harder to comprehend, rather than what is often referred to as complex sentences, such as passives and other less frequent sentence structures (Meyer, Mack, & Thompson, 2012; Kempler, Almor, & MacDonald, 1998; Kempler, Almor, Tyler, Andersen, & MacDonald, 1998).

All of these studies show that language comprehension in AD is differently impaired relative to production. Production of sentences and connected language output is relatively spared in AD, but comprehension is impaired. The opposite is found for single-words, where the production seems to be impaired but the comprehension spared. This means that while persons with AD can still comprehend the individual parts of sentences, the syntactic parsing is difficult. Sentences that require a lot of inference rely more on cognitive capacities than when the sentences require less inference.

3.3 Language in primary progressive aphasia

The main diagnostic criterion for primary progressive aphasia (PPA) is an isolated and progressive language impairment for at least two years, with relative sparing of other cognitive functions (Gorno-Tempini et al., 2011; Mesulam, 2001; Gorno-Tempini & Miller, 2013). As previously mentioned, three variants of PPA have been identified: a semantic variant (svPPA), a non-fluent/agrammatic variant (nfavPPA) and a logopenic variant (lvPPA). There are no concrete numbers showing the distribution between the three subtypes, and some researchers claim that in as many as 40% of the cases it is not possible to make a correct classification (Sajjadi et al., 2012). The three variants are distinguished based on language manifestations and underlying neuro-pathological differences. Table 3.1 summarizes the main language features of the three PPA variants. The main focus of this and future sections on PPA will be on the logopenic variant of PPA (lvPPA), unless stated otherwise, because it is the variant that is most similar to AD in both underlying neural pathology and linguistic symptomatology.

In many cases, it is difficult to set a certain diagnosis until after a few years post-onset, as the diseases can be very similar (Gorno-Tempini et al., 2011; Henry & Gorno-Tempini, 2010). There is also some debate regarding the diagnostic unity for the lvPPA and nfavPPA subtypes (see 1.3.2.1 and 6.3.3), and some researchers argue that each of these two types can be further split into two different variants (Matias-Guiu et al., 2019; Vandenberghe, 2016).

3.3.1 Language production in lvPPA

The main language impairment found in lvPPA is one of lexical retrieval. This is especially apparent in experimental settings, such as picture naming. Word search is also present in spontaneous speech, resulting in slow speech, with frequent pauses allowing for word search. Some patients experience speech-sound errors both in confrontation naming and spontaneous speech. These are mainly phonemic paraphasias with speech-sound substitutions that are well articulated, without signs of agrammatism or impaired motor speech. This last point is one of the main differences between lvPPA and nfavPPA (Gorno-Tempini et al., 2011; Wilson et al., 2010; Henry & Gorno-Tempini, 2010; Mesulam, 2007).

Table 3.1: Main language manifestations in the three subtypes of PPA

	svPPA	**nfavPPA**	**lvPPA**
Production	**Impaired** confrontation naming **Spared** repetition and speech production (grammar and motor speech)	Agrammatism **and/or** effortful, halting speech with inconsistent speech-sound errors and distortions (apraxia of speech)	**Impaired** single-word retrieval in spontaneous speech and naming, repetition of sentences and phrases, (phonological) speech errors in naming and spontaneous speech **Spared** motor speech and absence of agrammatism
Comprehension	**Impaired** single-word comprehension and object knowledge for low-frequent or low-familiarity items	**Impaired** comprehension of syntactically complex sentences **and/or spared** single-word comprehension **and/or spared** object knowledge	**Spared** single-word comprehension and object knowledge

The operational criteria for lvPPA include: 1) presence of aphasia; 2) impaired sentence repetition and comprehension; 3) presence of anomia with evidence of relatively spared single-word comprehension; 4) evidence of phonemic paraphasias; 5) slowed rate of verbal expression due to pauses for word retrieval or verbal formulation; and 6) absence of agrammatic or telegraphic verbal output (Gorno-Tempini et al., 2011).

Noun naming is usually better spared than verb naming for lvPPA (Mack et al., 2015; Lind et al., 2018) and nfavPPA, but opposite for svPPA (Hillis, Tuffiash, & Caramazza, 2002; Hillis, Oh, & Ken, 2004; Hillis et al., 2006). High frequency, familiarity and low AoA have all been recognized as factors that aid lexical retrieval for persons with PPA (Hirsh & Funnell, 1995; Ukita, Abe, & Yamada, 1999; Vonk et al., 2019).

Persons with lvPPA often have frequent repetitions in their spontaneous speech, and difficulties with sentence repetition (Grossman et al., 1996; Hoffman, Sajjadi, Patterson, & Nestor, 2017; Louwersheimer et al., 2016), which can be an indication of impaired working memory (Meyer, Snider, Campbell, & Friedman, 2015; Silveri et al., 2014).

Bilingual PPA is still a young field of research, and there are not may studies that investigate the patterns of deterioration in both or all languages of bilinguals with PPA. In a review by Malcolm, Lerman, Korytkowska, Vonk, and Obler (2019) on all three PPA subtypes, the authors find two main patterns: either a parallel deterioration of all languages, or better sparing of L1.

In sum, the language impairments in lvPPA are similar to the ones observed in AD (i.e., impaired lexical retrieval and relatively spared sentence production), but with a relative sparing of general cognitive functioning. Two important clinical markers for lvPPA are impaired sentence repetition, and absence of agrammatism. However, there are some uncertainties about the lvPPA diagnosis, and the divergent results reported in the literature may in fact reflect the ongoing discussion about the possible division of lvPPA into two different subtypes.

3.3.2 Language comprehension in lvPPA

Since lvPPA is a relatively young diagnosis, it was the last one to be implemented in the PPA classification system (Gorno-Tempini et al., 2011; Henry & Gorno-Tempini, 2010), there is also less research on this sub-type of PPA compared to the two other types.

On a general note, single word comprehension and knowledge and grammatical abilities are usually well preserved throughout the course of the disease for lvPPA (Gorno-Tempini et al., 2011). Comprehension impairments have not been observed to the same degree as impairments in language production.

Two studies of sentence comprehension in PPA found that persons with lvPPA are more impaired on comprehension of longer sentences (Wilson, Galantucci, Tartaglia, & Gorno-Tempini, 2012; Charles et al., 2014). In the study by Wilson et al. (2012), persons with lvPPA performed well on short sentences, even passives, but showed impairment as the sentence length increased. Charles et al. (2014) could only identify a sentence comprehension deficit on sentences which included center-embedded clauses (‘the fox that followed the domesticated cat was fierce), and relate this to atrophy in fronto-temporal brain areas which are associated with sentence comprehension deficits, rather than an impairment in working memory or mapping of semantic and thematic roles.

However, since persons with lvPPA have impaired working memory, as evidenced by their impaired sentence and word repetition skills (Leyton, Savage, et al., 2014) and impaired phonological short term memory (Meyer, Snider, et al., 2015), it is reasonable to assume that sentence comprehension deficits similar to those observed in AD can be expected. Recent research also suggest that auditory acuity is impaired in PPA, which will influence language comprehension (Utianski et al., 2019; Hardy, Johnson, & Warren, 2019).

3.4 Dementia diagnostics

Since the main symptoms for most diseases that lead to dementia are primarily memory-related, language screening during the diagnostic process receives less focus. The diagnostic criteria developed by MacKhann and colleagues (McKhann et al., 1984; McKhann et al., 2011), which are recognized as a standard for dementia screening, suggests that language abilities should be assessed at the very least by naming, semantic similarity judgment and sentence repetition tasks. In cases of so-called non-amnesic debut, especially in cases where the patient's subjective claims are language-related, a more thorough language screening is necessary (McKhann et al., 2011).

The *Consortium to Establish a Register for Alzheimer's Disease* (CERAD) was founded in 1984 to develop a screening battery that could be implemented as a standard globally (Morris et al., 1989). This screening tool includes picture naming — 15 items from the *Boston Naming Test* (BNT), (Kaplan, Goodglass, & Weintraub, 1983) — verbal fluency and a word list to be learned and later recalled for assessment of verbal working memory. Changes in language behavior can also be an early indicator of cognitive decline, as found in non-amnesic mild cognitive impairment (MCI).[2] For differential diagnosis, especially between AD and lvPPA a more thorough language screening might be needed to offer the correct diagnosis.

Tests for general cognitive functioning, such as the *Mini-Mental State Examination* (MMSE) (Folstein, Folstein, & McHugh, 1975) usually include a section on language functioning. Such tests are meant for quick screening, and do not include more than a few items of picture naming (language production), reading (sentence comprehension), and sentence repetition.

Many tests for cognitive functioning rely on verbal information: for instance, tests for verbal episodic memory, such as the word list learning and recall test found in the protocol developed by CERAD (Morris et al., 1989). Despite the fact that the majority of cognitive and neuropsychological tests are language-based, there seems to be little recognition of the language impairments in AD and PPA during the diagnostic process, at least in a Norwegian context.

3.4.1 The Norwegian dementia screening protocol

The Norwegian register of persons assessed for cognitive symptoms (NorCog)[3] works to develop assessment tools for dementia in Norwegian, and to run quality control on these, and to ensure that patients are assessed according to the same protocol across the country.

The main screening protocol developed by NorCog follows the lines of the CERAD test battery, and includes the 15 item version from the BNT, verbal fluency for phonemes and categories, a subjective judgment of language fluency made by a neurologist during the screening process,[4] and a translated version of the word list for learning, recall and recognition from the original CERAD test battery, consisting of translation equivalents of the English word list items (Kirsebom et al., 2019).

[2]Mild cognitive impairment is a disorder characterized by memory impairment, difficulties with learning and reduced ability to concentrate. None of the symptoms are so severe that a diagnosis of dementia can be made (WHO, 2016a).

[3]*Norsk register for personer som utredes for kognitive symptomer i spesialisthelsetjenesten (NorKog)* is the Norwegian quality register for dementia which was opened in 2007 to increase knowledge about diagnostics, assessment and treatment of cognitive symptoms and dementia (https://www.aldringoghelse.no/norkog/).

[4]The judgment includes six yes/no-questions: *Does the patient have poor fluency? Does the patient produce repetitions in speech? Does the patient have dysarthria? Is the patient's speech coherent? Is the patient's comprehension poor? Does the patient have word finding difficulties?*

This protocol is not intended for language assessment, and the language tests included are therefore not very comprehensive. Furthermore, these tests are translated rather than adapted to Norwegian, which can potentially be problematic for language assessment. The BNT is usually not employed by speech-language therapists in Norway because there are no Norwegian norms available for this test. The version included in the NorCog battery is the same as the 15-item version extracted from the original test. However, for the (American) English version the items are arranged from the lowest to the highest frequency ("easier" words before "difficult" words). But, as mentioned earlier in section 2.1.2, underlying psycholinguistic variables are not always comparable across languages. A Norwegian translation equivalent of a "difficult" word in the original English version might have higher or lower frequency or familiarity, and therefore be easier or more difficult to name in Norwegian than in English. Table 3.2 shows how the frequencies differ between Norwegian and English for the 15 items in the short version of the BNT, based on recent frequency data from the Norwegian Web as Corpus (NoWaC) (Guevara, 2010) and the Corpus of Contemporary American English (COCA) (Davies, 2008).[5] This overview shows that there are certain items in the Norwegian version that should have been ordered differently compared to the original English version. Furthermore, among the Norwegian words there are more words with low frequency (under 1000 counts per million) than among the English words. The English word list contains more words of medium frequency (1000 - 3000 counts per million), even among the so-called ‘easy words’ (the five first items in the list).

Table 3.2: Frequency counts (per million) for items in BNT

	Item	Norwegian Frequency (NoWaC corpus)	English frequency (COCA corpus)
Easy	House	111407	316918
Easy	Comb	3525	2337
Easy	Toothbrush	971	1165
Easy	Squid	574	1765
Easy	Bench	12457	14026
Medium	Volcano	631	3042
Medium	Canoe	1204	3041
Medium	Beaver	580	2655
Medium	Cactus	323	2007
Medium	Hammock	206	1062
Difficult	Stethoscope	102	572
Difficult	Unicorn	386	676
Difficult	Trellis	48	565
Difficult	Sphinx	10	596
Difficult	Pallette	314	11

From a linguistic perspective, there are certain issues to be noted concerning the translated

[5]Both corpora give lemma frequencies in counts per million.

word list for learning and recall from the CERAD battery. These translations do not take underlying variables into account, and the words in the Norwegian version differ from the ones in the original English version in number of syllables and phonological structure. A greater number of longer words in the Norwegian version can potentially make the learning and recall even more problematic for persons with phonological short-term memory impairments.

Furthermore, translation equivalents may not always translate the full meaning of the source word. For instance, in the English version, the words 'letter', 'pole' and 'arm' have homonyms which are not covered in their Norwegian translations 'brev', 'stokk' and 'arm'.[6]

Even if words are translation equivalents of each other, they can be potentially different in cultural interpretation, familiarity or frequency of occurrence. These potential differences can result in different patterns of responses when used in different languages. If the purpose of the test is to measure linguistic and/or cognitive ability, a translation without consideration of functional, cultural, and frequency equivalence may introduce bias (Peña, 2007).

These issues can be used to argue for adaptation, or even creation of language-specific tests rather than an uncritical translation of tests between languages.

[6]'Brev' can only mean a letter one sends in the post, and not the letters of the alphabet. A better translation for 'stokk' would be 'stick' or 'cane', whereas 'pole' could also denote magnetic/geographic poles, or wooden constructions which would better be translated to 'pol' and 'påle' respectively. The Norwegian word 'arm' can only refer to body parts, and never to weaponry.

4

Methods and materials

One way of learning as much as possible about language impairment in dementia is to use different methods to study the same phenomenon. In this dissertation, I have combined several tests that involve the same domains, and used both behavioral and experimental methods in the same experiments. This chapter is structured as follows: First the studies are introduced in section 4.1, followed by a more detailed description of the methods, participants and analysis of each of the studies (sections 4.2 to 4.5). Since there is some overlap between participants, materials and procedures in studies II and III, these issues are only reported on once, for study II, with some further specifications for study III. In section 4.6, some considerations regarding validity and reliability are discussed. Finally, some ethical considerations are discussed in 4.7.

4.1 The studies

Table 4.1: Overview of the studies

Study	**Type of study**	**Participants**
Study I	Literature review	N/A
Study II	Multi-case study to investigate lexical retrieval in healthy aging, AD and PPA	7 AD, 2 PPA 29 controls
Study III	Multi-case study of sentence comprehension abilities in AD, PPA and healthy aging	5 AD, 2 PPA 9 controls
Study IV	Single-case follow-up study of bilingual PPA	1 lvPPA?

Table 4.1 gives an overview of the four studies. More detail about each one can be found in the next sections of this chapter, in chapter 5 and in each of the individual papers.

4.2 Study I - Literature review

The main objective of this study was to identify which (psycho-)linguistic variables of words are most central for naming and comprehension success in AD and PPA, and whether these variables can be used to distinguish between different dementia types (AD and PPA), and between healthy and pathological aging. If that is the case, these variables and properties could and should be taken into account when constructing assessment tools for language impairment in AD and PPA. The variables under investigation were: word class, frequency, age of acquisition (AoA) and imageability (see 2.1.2).

4.2.1 Search criteria and selection

Three literature searches were conducted between November 2016 and March 2018 to ensure that the latest and most updated articles were included in the review. The searches were carried out searching for specific terms in the PubMed and Science Direct electronic databases.The two first searches were carried out using the following search terms: ‘Alzheimer’s disease’, ‘primary progressive aphasia’, combined with ‘lexical access’, ‘imageability’, ‘age of acquisition’, ‘word frequency’, and ‘word class’ In the third search, the terms ‘bilingualism’ and ‘multilingualism’ were added to specifically identify studies involving lexical retrieval in bilingual speakers — studies which might not have shown up in the first two searches. In total, the three searches returned 433 titles. After a thorough screening of the titles, abstracts, methodology and discussion of the papers, as well as removing duplicates and articles written in other languages than English, only 48 articles were included in the review. A schematic overview of the selection process is found in study I (Ribu, Under revision).

Papers that were included in the review were original, peer-reviewed research papers written in English. There was no restriction in regards to publication year. Articles that reported on other types of dementia than AD and/or PPA were excluded, unless they also reported on either one or both of the diseases in question.

Only papers reporting on behavioral data were included, since methodologies and results from structural and functional imaging studies, ERP and EEG studies, as well as eye tracking and other online methods, are difficult to compare to strictly behavioural data.

4.2.2 Analysis

The 48 articles that fulfilled the inclusion criteria were sorted based on which diagnosis they reported on — AD, PPA or both — and whether the main focus was on production or comprehension. Next, the methodologies that were employed in each study was considered, and the effect of the variables were studied in relation to the different tasks. This means that I could study the individual effects of, for instance, AoA on picture naming tasks or verbal fluency tasks.

Many of the studies included were written before the consensus on diagnostic criteria for PPA was made and implemented (Gorno-Tempini et al., 2011). Thus, a number of studies operated with fuzzy categorizations of the subtypes, or used names for the subtypes that are now outdated. No effort was made to try to classify the participants in each study following the 2011 criteria, since in most cases, there was not enough information provided in the original papers to properly do so. This of course means that there is some uncertainty about how the lexical variables affect the different subtypes of PPA.

The number of studies that investigated the same issues was often small, and it would be too difficult to perform robust regression analyses of these studies for a meta-analytic review, so the results from all studies were considered qualitatively.

4.3 Study II - Free word associations

In study II, 38 participants were recruited to investigate how free word associations (FWA) can be used as a supplement to more traditional tests of lexical access. For this purpose, a new picture naming and a word-to-picture matching test were also created, taking into account the psycholinguistic variables discussed in study I (see 4.3.5). The next sections focus on recruitment of participants, and selection and creation of different tests for cognitive and linguistic screening. The same tests were also used in study 3 (see 4.4).

4.3.1 Participants and recruitment

Participants with AD and PPA were recruited from the memory clinic at Oslo University Hospital (OUS), at Ullevål, Oslo, and control participants were recruited from personal networks and online information posts, or were spouses and/or family members of the participants with AD/PPA. The young adults came mainly from the student body at the University of Oslo (UiO) and personal networks. Posters with information about the project containing a QR-code that led to a sign-up website were hung in study halls, at the main libraries and around hallways at the University of Oslo's Faculty of Humanities. Three information posts with a link to the sign-up website were posted to Facebook, which were shared a total of 67 times. Despite the heavy activity in the comment sections of these posts, only a handful of people signed up after reading the information on the sign-up page.

Table 4.2 gives an overview of how many potential participants were contacted or initiated contact through the sign-up website, and how many of those that were, in turn, tested and finally included for the studies.

Table 4.2: Overview of participant inclusion

	AD	PPA	Older controls	Younger controls
Contacted	16	4	17	26
Declined / withdrew	9	2	1	9
Tested	7	2	16	17
Excluded	0	0	2	2
Total included	7	2	14	15

Participants with a pre-morbid language deficit (e.g., stroke-induced aphasia), or other neurological illnesses and disorders were not included in the study. Neither were left-handed participants; persons with a history of drugs or alcohol abuse; participants that the neuropsychologists at the memory clinic at OUS deemed to be unfit for long testing sessions.[1] Control participants with an MMSE score below 26/30 were also excluded.

Four participants were excluded from the control population as they did not fulfill the inclusion criteria; two participants from the older control group were had too low MMSE scores, one of the younger control participants was left-handed, and one reported having a known neurological condition.

A Wilcoxon rank-sum test showed that there was no significant difference between the neurologically healthy older control participants (HCE) and the AD and PPA groups with respect to age (AD vs. HCE: $W= 63.5, p= 0.291$, PPA vs. HCE: $W= 21.5, p= 0.258$) or education (AD vs. HCE: $W= 22.5, p= 0.0508$, PPA vs. HCE: $W = 5.5, p= 0.199$). However, the education levels between AD and HCE is close to significant; HCE participants had slightly higher education than AD participants. Table 4.3 gives information about demographic details at a group level for the four participant groups. The participants with PPA are reported individually, since calculating averages for two persons only is futile.

Education is reported in number of completed years, starting from first grade in elementary school. A majority of the younger control participants were still students, and most were tested mid-semester. In those cases, years were counted up until the last completed semester. Among the older control participants, approximately 2/3 had higher education corresponding to Master's level by today's standards. One participant with AD had no education after high school, one had four years of vocational training after (academic focused) high school, and the five others

[1]This last point only pertain to the participants with AD and PPA.

[2]In some cases biological 'sex' and social 'gender' do not coincide. In the current study, the participants' self-reported gender is recorded.

Table 4.3: Demographics of participants in study II

	AD (N=7)	PPA (N= 2)	Control (old) (N=14)	Control (young) (N=15)
Gender (M:F)	2:5	2:0	5:9	3:12 [2]
Age Age span	m: 74,85 (SD: 4,3) 66-79	75;81	m: 73,14 (SD: 6,3) 66-89	m: 24,8 (SD: 2,9) 19-30
Education Education span	m: 14,28 (SD: 2,49) 10-18	15;14	m: 16,60 (SD: 2,37) 15-20	m: 16,3 (SD: 1,8) 13-20

all had at least three years of university or college education. Both participants with PPA had an education level comparable to a Master's degree by today's standards.

All participants reported that Norwegian was their main language in daily life, and all considered themselves to be monolingual. However, all participants reported having learned at least one foreign language in school, and a majority of the participants had at some point in their lives used more than one language in their daily lives, either by living in a country where another language was spoken, or by studying one or more languages at a higher level. Only three participants reported not actively using any other language than Norwegian in their daily life (speaking, writing or reading), but all participants were regularly exposed to foreign languages — mainly English — through television shows, films and from listening to music/radio.[3]

Dementia diagnosis (AD or PPA) was made at the memory clinic at OUS by trained neurologists. The diagnosis follows the international guidelines for Alzheimer's diagnosis (McKhann et al., 2011), and is based on results from a range of standardized tests and supported by physiological findings, such as biomarkers in cerebrospinal fluid and neuro-imaging (MRI, SPECT or CT). One of the participants with PPA had initially been diagnosed with AD before this was later changed to PPA.

Persons who are diagnosed with a cognitive impairment in Norway are asked if they want to be registered in a national register for treatment and research — NorCog, see 3.4.1). All participants with AD or PPA in this study are also in this register. Ethical considerations regarding research with clinical populations are discussed in 4.7.

One of the main recruitment challenge was to find willing participants. This can be seen in table 4.2, as 16 persons with AD and 4 with PPA received an information package by mail and were contacted personally via telephone, but only 7 with AD and 2 with PPA consented to participate. Furthermore, 1 participant with AD initially consented to participate but withdrew the day before testing was scheduled. The same can be seen for the younger control participants, where 26 initially showed interest, but only 17 participated.

[3] Some participants provided responses on the verbal fluency task, the picture naming tests and the word association tests in other languages than Norwegian, although the test session was kept in a constant monolingual environment. This means that even "monolingual" persons show activation for several languages simultaneously, and can therefore not be called "pure" monolinguals.

It was surprisingly difficult to recruit participants to the older control group. Potential relevant participants were very active in the comment sections below most of the shared Facebook postings, but very few followed the link to the sign-up website. The project got a lot of online exposure in relevant environments, but the exposure did not translate into actual participants. Two possible reasons for this might be due to the long testing sessions (see 4.3.6), and geography; some potential participants from areas outside of South-Central Norway showed interest in the project, but were not interested in traveling to Oslo to participate. Furthermore, participation was voluntary, and no compensation beyond travel reimbursement was offered.

4.3.2 Cognitive test battery

To assess the functional cognitive level of all participants, everyone included in this project went through a comprehensive screening of cognitive and linguistic functioning. There is a substantial overlap between the test battery used at the memory clinics in Norwegian hospitals and the protocol used in this project. Patients who were referred from the memory clinic were therefore given a shortened version of the full battery that was used with control participants (see Appendix C). However, as the test battery from OUS is used for screening of cognitive function, and not language, some additional tasks were included (see 4.3.3).

Tests for screening of cognitive functioning rely heavily on linguistic as well as cognitive processing, and some cognitive tests have been proven to correlate more with language processing skills than others (MacDonald & Christiansen, 2002). For instance, set-shifting and sequencing problems are common for people with AD, which is often expressed through impairment in impulse resistance. This often results in slowness on the inhibition condition of the Stroop task (Lezak et al., 2012). Results on the inhibition sub-test of the Stroop task have also been found to correlate with the ability to suppress a contextually irrelevant language for speakers of more than one language (Bialystok et al., 2012), and to suppress irrelevant and personal responses on restrictive language activation tasks. This has implications for what kinds of responses one can expect on confrontation naming tasks (see Ribu et al., submitted).

The cognitive test battery included the *Mini-Mental State Examination* (MMSE) (Folstein et al., 1975), *Trail-Making Test A* and *B* (TMT-A+B) (Tombaugh, 2004), and a word-list learning, recall and recognition test, and a figure copying and recall test from *The Consortium to Establish a Registry for Alzheimer's Disease* (CERAD) (Morris et al., 1989); three different *Digit span* tasks (digits forward, digits backward and digit ordering); and the *Stroop* task.

The MMSE assesses orientation to time and place, memory recall and working memory, concentration, language and visuo-spatial drawing.

In the word-list task in CERAD, participants are presented with ten words, one at a time, printed in large letters on a piece of A4 paper, and asked to read each one out loud. The test administrator flips the pages to reveal the next word as soon as the participant has read the word on the current page. Immediately after reading the ten words, the participant is asked to recall as many words as they can remember. This procedure is repeated another two times with the same words, but in a different order of presentation. After 20-30 minutes, during which other, non-verbal tasks are completed, the participant is asked to recall the word list. Finally there is a word recognition list, including ten new words in addition to the ten words from the learned list. In the figure test, participants are first asked to copy four geometrical shapes, and after a series of non-visual tasks they are asked to recall these shapes.

The *Trail Making Test* (TMT) is a measure of visual conceptual and visuo-motor tracking, mental flexibility, visual attention and executive function (Lezak et al., 2012; Crowe, 1998). *TMT-A* consists of a page on which the numbers 1 through 25 are encircled and scattered around the page. The participant is asked to draw a line connecting the numbers in order as fast as they can. Errors are pointed out and can be corrected, but are not scored. In *TMT-B*, the page is filled with both numbers (1-13) and letters (A-L). In this task the participant is asked to draw a line alternating between numbers and letters (1-A-2-B-3-C etc.). The time it takes to complete TMT-A is subtracted from the time to complete TMT-B, which gives a total switching cost: a measure of mental flexibility, attention and switching.

Digit span tasks are also common tests for verbal working memory, and are often said to correlate with different measures of language proficiency (MacDonald & Christiansen, 2002). In the current study, participants were tested on digits forward, digits backward and digit ordering. In the first task, participants hear a series of numbers and are asked to repeat them back to the test administrator in the same order of presentation. In digits backward, the participants have to repeat the digits back in reversed order of presentation. The digit ordering task, where participants are asked to order digits from the lowest to the highest number, is supposed to be particularly difficult for persons with AD (MacDonald & Christiansen, 2002; MacDonald, Almor, Henderson, Kempler, & Andersen, 2001).

A two-panel version of the Stroop task (Delis, Kaplan, & Kramer, 2001) was used to measure executive functions. The two panels consisted of word reading (all words in black ink) and a color-word-interference panel (names of colors written in another colored ink than the color they represent). Time to complete the word reading panel was subtracted from the time it took to read the color-word-interference page, resulting in an inhibition cost score.

Table 4.4 gives an overview of the tests included in studies II and III and the cognitive and/or linguistic domains they measure. A superscript ^{0}, indicates that the test is also used in the test battery from the memory clinic.

4.3.3 Linguistic test battery

At the intersection between cognitive and linguistic tests are a sub-set of tests which assess both linguistic and cognitive abilities, such as tests for semantic knowledge and association — e.g.,

Table 4.4: Tests included in studies II and III

Test	Measure
MMSE [0]	Orientation to time and space, memory recall, working memory, concentration, language and visuospatial drawing
CERAD [0] word list learning and recall	Verbal learning and verbal working memory
CERAD [0] figure copying and recall	Constructional praxis, visuospatial drawing and -memory
Trail Making Test A + B [0]	Visual conceptual and visuo-motor tracking, mental flexibility, visual attention and executive function — especially shifting
Digit span Forward Backward Ordering	Working memory, attention, phonological storage and processing
Stroop	Executive functions — especially inhibition and speed of processing
Verbal fluency [0] (phonetic and semantic)	Verbal retrieval and recall, auditory attention, short-term memory, cognitive flexibility, and long-term vocabulary storage
Pyramids and Palm Trees and Semantische associatie test	Semantic knowledge and memory of objects and actions
Free word association	Semantic processing, lexical retrieval and semantic memory
Picture naming	Confrontation naming/lexical access
Word-to-picture matching	Lexical comprehension
Sentence-to-picture matching (Eye tracking)	Sentence comprehension
Cartoon description	Semi-spontaneous speech and narrative production

the *Pyramids and Palm Trees test* (PPT) (Howard & Patterson, 1992, 2005).

In the picture version of PPT knowledge about how objects belong together in the world is assessed through a procedure where participants match pictures to show their semantic knowledge. A similar test has recently been developed in the Netherlands for assessment of actions. This test, *De semantische associatie test* (SAT) (Visch-Brink, Stronks, & Denes, 2005), is used in this study together with the PPT to assess semantic memory and knowledge of actions and objects respectively. In both tests, the participant is asked to choose one picture that is related to another picture. In PPT the participants choose between two competing items, whereas they choose between three items in SAT.

A *verbal fluency* test, where participants are asked to name as many words as they know within a given time frame, is often used in neuropsychology as a measure of updating and verbal working memory (Lezak et al., 2012; Marczinski & Kertesz, 2006; Damico et al., 2010), but can also be used in linguistic screening as a measure of word generation/lexical retrieval. Both phonemic and semantic fluency was tested, and participants were instructed to name as many words they knew beginning with the sounds /f/, /a/ and /s/, as well as name as many animals they knew, in separate sessions of 60 seconds each.

Free word associations (FWA) can be used to study both lexical activation and lexical organization. The test included here is an oral test of 30 cue words to which participants have to freely associate. The test administrator reads a word from a list, and the participant is instructed to respond with the first word they can think of as a response to that word. A short introduction to how the FWA test was created is found below in 4.3.4.

Two different picture naming tests were included to assess lexical retrieval of both nouns and verbs. The Boston Naming Test (Kaplan et al., 1983) is the only test of language assessment that is included in the neuropsychological test battery from the memory clinic at the University hospital. BNT is not adapted to Norwegian, and the items are ranked in difficulty following frequency of the original test items in English (see 3.4.1). A second, locally constructed picture naming task was also included to investigate how test results may be influenced by taking underlying linguistic variables into account when constructing naming tasks. In the end, only results from the locally constructed naming task (henceforth the *Picture Naming Task*) were analyzed. This decision was made as there are no norms for the BNT, and it is neither translated nor adapted to Norwegian, and therefore not validated for use in a Norwegian context. In the Picture Naming Task, both verbs and nouns are assessed, and the total number of tested items is 60, compared to 15 nouns in the BNT. This makes the Picture Naming Task more comprehensive than the BNT.

The same items were also included in a word-to-picture matching task, for assessment of lexical comprehension. Inclusion of the same items in both a production and comprehension task was considered as one possible way to assess if knowledge of the items are spared even if access to the phonological form is impaired. How the items for these two tests were selected is detailed in 4.3.5.

A cartoon description task was included with the intention of analyzing word finding difficulties in narrative production (see below for study IV, 4.5.2), but was left out as it would make more sense to analyze these data in a separate study designated for that issue, this will be discussed more in 6.4.

Finally, sentence comprehension was tested with a sentence-to-picture matching task with an eye tracker (more on this in section 4.4).

4.3.4 Creation of the free word association test

A word list of 30 items (18 nouns and 12 verbs) was created based an already existing test for free word association (FWA) for Norwegian (Bøyum, 2016). This original test has been tried out on two cohorts of neurologically healthy participants: older adults (over 60 years) and younger adults (18-30 years). The FWA test in Bøyum (2016) differs in two main respects from the one included in the current study: First, the original test consists of 100 items. Secondly, the long list is developed as a written version, whereas this shortened version is presented to participants orally.

The 100-words list was initially reduced to 50 items by removing all adjectives, homophones, words that had low inter-rater agreement, and words that elicited more than 20% multi-word responses by participants older than 60 years in the 100-word study. Some words that triggered emotional reactions (“to miss”, to fear”, etc.) were also removed after trying out a pilot version of 50 words on a small set of participants. The FWA test should be fast and easy to administer, so the 50 words were then further reduced to the final version of 30 words. During the pilot testing it was also clear that the instructions had to be updated, as persons needed to be repeatedly reminded to reply with single words only.

The final selection included 20 words that had predominantly received meaning-based responses and ten words that received both meaning- and position-based responses in the 100-words version. Order of presentation was randomized using Microsoft Excel. The 30 items included in the final version with English translations can be found in Appendix E. Bøyum, Hansen, and Ribu (forthcoming) explains in more detail how both the 100 and the 30-word lists were created and how they both can be used for research purposes. The second paper in this dissertation investigates how the FWA test can contribute to knowledge on theoretical questions in linguistics, and can also be used in clinical assessment of AD and PPA.

4.3.5 Selection of items for Picture Naming and word-to-picture matching

To create the picture naming and matching tasks, the online database "Norwegian Words" (Lind et al., 2013) was consulted to select all nouns and verbs that were coded as having either high or low frequency, imageability and age of acquisition (AoA), excluding median values. This search returned a list of 1188 words: 747 nouns and 441 verbs.

First, a manual selection was made by removing homonyms (e.g., ‘huske’ *a swing* or *to swing* and even *to remember*, or ‘et bein’ *a leg* or *a bone*). Words with competing synonyms (with either similar AoA, imageability or frequency) were excluded (e.g., ‘lege’ and ‘doktor’ both translate to *(medical) doctor*, and have quite similar AoA and imageability ratings). Objects that were either too specific (‘en ponni’ *a pony*) or too general (‘et dyr’ *an animal*) were left out. Depictability was considered in cooperation with an artist, and words she considered to be too complex or difficult to draw in a simplistic style were left out. After this selection 431 words remained; 307 nouns and 124 verbs. Distribution across the three variables for the 431 words was explored in R (“R: A Language and Environment for Statistical Computing,” 2019) with the Shapiro-Wilk test (Happ, Bathke, & Brunner, 2019).

AoA and frequency for the selected words were slightly skewed towards words with a high positive predictive value for lexical access — i.e., high frequency, and low AoA (see 2.1.2). This skewness was reduced by applying a logarithmic transformation (log(Xi)) to remove outliers and select words that had a normal distribution of these variables.

For imageability, the data had a slight negative skewness, with more words with a negative value (i.e., low imageability). This skewness was corrected by applying an exponential transformation (exp(Xi)). When the skewness was reduced, a total of 120 words were extracted from the two subsets (60 verbs and 60 nouns), trying to keep the selection normally distributed with regards to all three variables.

A Shapiro-Wilk test was applied for each word class to calculate the deviance from normal distribution. For the noun subset, all three variables showed normal distribution; however, in the verb subset there was a significant skewness towards words with low imageability, even after the necessary reductions and transformations (w = 0.93, p = 0.002). Verbs are in general less imageable than nouns (Bird, Howard, & Franklin, 2003), and this can therefore be expected.

Black-and-white line drawings for the 60 nouns and 60 verbs were prepared by the same artist who assisted in the selection process. A group of ten adult, native speakers of Norwegian (age range 29 – 78, mean: 40,9 SD: 16,26; education range 13 - 20 years, mean: 17,9 SD: 2,34) rated all of the 120 pictures in two separate sessions (60 images in each session). In both sessions they first named the picture (name agreement), and then rated on a five-point scale how well they felt the picture represented their mental image of that word (image agreement). Only the words with a 90% or higher name agreement and high image agreement (4-5 on the 5-point scale) were included in the final selection.

Figures 4.1 and 4.2 show the distribution of the final 60 words that were included in the naming test. A list of items that were included in the picture naming and word-to-picture matching task can be found in Appendix D, together with some examples of the illustrations.

A post-hoc Shapiro-Wilk analysis of the normal distribution of the final 30 nouns and 30 verbs show that the verb selection is still not normally distributed for imageability (w = 0.90, p = 0.01), meaning that the verbs in the naming and picture-matching tasks are generally of lower imageability than the nouns. This is seen in figures 4.1 and 4.2, where the majority of the words are clumped together on the left side of the x-axis. Note also that the scales of the x-axes are different; the imageability range is greater for nouns than for verbs.[4]

The figures show the words' frequency in color shading: a more purple color denotes higher frequency, and a more red color denotes lower frequency. On the x-axis, the words are scattered from lowest to highest imageability, and on the y-axis, the words are scattered from lower to higher age of acquisition. Verbs are in general less imageable than nouns (Bird et al., 2003, 2000), which is also visible in these figures.

There is a plan for further trialling and norming of this test, so it can hopefully be used more

[4]The scales run from 500-1100 for the nouns and 300-700 for the verbs, which is a result of the different transformations that were employed to ensure (near-)normal distribution of the values.

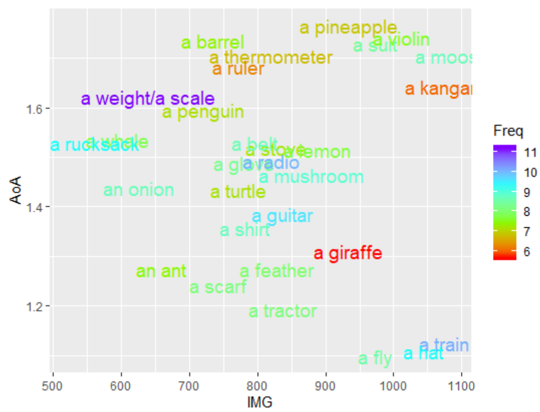


Figure 4.1: Final noun selection

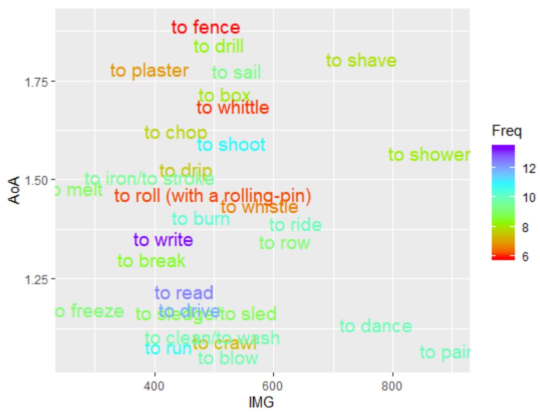


Figure 4.2: Final verb selection

widely.

4.3.6 Procedure

All participants with AD and PPA were assessed and diagnosed by a neurologist at the Oslo University Hospital, and asked if they wanted to be contacted about participation for this project. All participants with AD or PPA, and a majority of the (young and old) control participants were tested by myself, while 13 participants were tested by research assistants employed for this project. The research assistants received training in how to conduct the testing, and were instructed to always provide the same instructions for each testing session.

MMSE, CERAD, TMT A + B, verbal fluency and BNT were administered at the hospital as part of the diagnostic process, whereas the Pyramids and Palm Trees (PPT), De semantische associatie test (SAT), digit span tasks, Stroop and the Picture Naming, word-to-picture matching, free word association and cartoon description tasks were administered by the researchers involved in this project.

Participants were tested in one session of approximately two to three hours. Testing mainly took place in the Socio-Cognitive laboratory at MultiLing, Center for Multilingualism in the Society Across the Lifespan at the University of Oslo (UiO). In four cases, participants with AD received a home visit instead of coming to the University. One participant with AD was tested in two separate sessions of 90 minutes each.

The order of testing differed slightly between participant groups, because participants with AD and PPA were assessed with some of the tests at the memory clinic. A schematic overview of the order of testing can be found in 4.3.

To reduce test anxiety for the participants with AD and PPA, it was important to start the test session with an easy task. Furthermore, it was important to leave enough time between the picture naming and the word-to-picture matching tasks, for all participant groups. For control participants it was also important to leave enough time between the learning and recall tests from

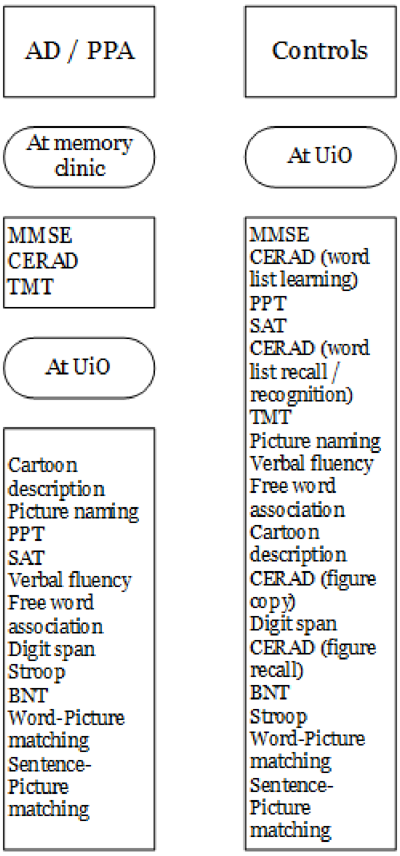


Figure 4.3: Order of tests per participant group

CERAD, and to have non-verbal tasks between the word list learning and recall sessions, and non-visual tasks between the figure copying and recall sessions. [5]

4.3.7 Analysis

In the following sections, I account for how the different tasks have been scored and analyzed.

To compare results between the participant groups, the Wilcoxon rank sum test has been used. This is computed in R ("R: A Language and Environment for Statistical Computing," 2019). The Wilcoxon sum rank test is a non-parametric test that is well-suited for operational statistics when the sample sizes are small. In this test the sums of the averages are ranked, rather than compared directly, which means it is less likely that the results are shifted because of outliers (Happ et al., 2019).

4.3.7.1 Accuracy

Responses in most tests were scored as either correct or incorrect, but for the Picture Naming and the FWA tasks, types of answers were recorded in order to investigate the relationship between

[5] This last point is of course also true for participants with AD and PPA, but the issue was resolved at the memory clinic.

the actual response and the target (for Picture Naming) or cue (for FWA).

Table 4.5 show the normative values for each of the tests included in this study. TMT and Stroop are timed, and each part is expected to be performed in less than 5 minutes. It is not possible to set a maximum performance score on these tests, and the column Max. (score) is therefore left blank. The same column is empty also for verbal fluency, as it is an individual measure for each participant. For these three tests, the average time and/or score for persons over 65 is given in the column Normative values. The FWA test is excluded from this table, since it does not operate with a numeric score, and results are not scored as correct or incorrect.[6]

Table 4.5: Max. scores and norms on the cognitive and linguistic tests

Test	**Max.**	**Normative values**
MMSE	30	26-30
CERAD WL-L	30	21,5/30
CERAD WL-R	10	7,2/10
CERAD F-C	11	10-11
CERAD F-R	11	8,51/11
TMT - A		34-40 sec.
TMT - B		80-104 sec.
Digits forward	16	
Digits backward	14	
Digit ordering	15	
Stroop 1		40,05 sec.
Stroop 2		135,50 sec.
Verbal fluency phon.		F: 15,8; A: 12,7; S: 17,5
Verbal fluency sem.		23,5
Pyramids and Palm Trees	52	98,7%
Semantische associatie test	20	93,43%
Word-to-picture matching	60	99-100%
Picture Naming, nouns	30	98,6%
Picture Naming, verbs	30	98,1%
Sentence-to picture naming	45	99-100%

The sentence-to-picture matching task have not been normed, but all 29 control subjects that participated in the study scored over 99% correct on this test, indicating that the test was not very difficult for most participants. The norms from the SAT are based on the percentile scores reported in Visch-Brink et al. (2005) for Dutch, and the norms from PPT in Howard and Patterson (2005). The word-to-picture matching task and the Picture Naming task were normed on an additional group of 20 neurologically healthy adults, aged 25-67 (see section 4.6).

Norms for verbal fluency among Norwegian-speaking adults are found in Egeland, Landrø, Tjemsland, and Walbækken (2006), and digit forward and backward in the Norwegian version

[6]CERAD WL-L = Word List Learning; CERAD WL-R = Word List Recall; CERAD F-C = Figure Copying; CERAD F-R = Figure Recall: Stroop 1 = control condition; Stroop 2 = inhibition condition

of WAIS -IV (Wechsler, 2008). CERAD and TMT norms come from Ivnik, Malec, Smith, Tangalos, and Petersen (1996), Kirsebom et al. (2019),and Tombaugh (2004). Norms for the Stroop test are based on data from Boone, Miller, Lesser, Hill, and D'Elia (1990). Considerations regarding the validity and reliability of these tests are discussed in section 4.6.

4.3.7.2 Scoring verbal fluency

Three rounds of phonemic and one round of semantic fluency was recorded for all participants. In the phonemic fluency task, words were scored if the initial sound was /f/, /a/ or /s/ — depending on which one was prompted. Common names and numbers, inflections of previously named words, series of compounds including the same first words, and repetitions were not counted. Dialectal variation was taken into account for words beginning with /s/. Participants who spoke a dialect where the initial sounds in the consonant cluster <sj> is pronounced /sj/ would get credit for words such as 'sjø' (*sea*), whereas the majority of the participants would not be credited for this, as the same word in their dialect would be pronounced with an initial /ʃ/. The total number of words generated for the phonemic fluency is the sum of items generated across all three trials.

Semantic fluency was only measured for one category: animals. All types of animals, living or extinct, were counted, but names of pets were excluded. Both supracategories (e.g., 'fish') and subcategories (e.g., 'salmon', 'cod' etc.) were accepted, so were different genders and parent/child categories of the same animal (e.g., 'cow', 'bull' and 'calf' would all be counted). Fantasy animals such as 'dragon' or 'unicorn' were also point-giving answers.

4.3.7.3 Scoring Free word association

Traditionally, free word association tasks have been scored using a binary system where responses are rated as either syntagmatic or paradigmatic in relation to the cue (Fitzpatrick, Playfoot, Wray, & Wright, 2015; Santo Pietro & Goldfarb, 1985; Eustache, Cox, Brandt, Lechevalier, & Pons, 1990). This dichotomous system is not always sensitive enough to capture the finer relationships between cues and responses, as many instances can be both syntagmatic and paradigmatic. For the FWA test used in this study, 11 scoring categories were used. Examples can be found in table 4.6. The high number of categories were included to better account for the detailed relationship(s) between cue and response words.

The word association data was scored individually by two raters, and disagreement was resolved through discussion between the two raters and a third consultant. For each cue word, all recorded responses were noted, and scoring was done based on the individual cue word, rather than per participant. This means that, for this study, it has not been possible to look at perseverations or priming. However, the individual responses are available for such studies at a later point: this topic is revisited in 6.4.

Table 4.6: Response categories for the word association task

Main category	**Subcategory**	**Cue**	**Response**	**English translation**
Meaning-based	Synonym	"sykehus"	"hospital"	*hospital → infirmary*
	Other semantic	"tolke"	"språk"	*interpret → language*
Collocation		"konsert"	"hus"	*concert → venue*
Form-based		"nyte"	"nype"	*enjoy → rosehip*
Mixed associations	Meaning + collocation	"avis"	"papir"	*newspaper → paper*
	Meaning + form	"håndtere"	"håndtak"	*to handle → a handle*
	Two-step	"tolke"	"Ringenes herre"	*interpret → Lord of the rings* (via J.R.R. Tolkien)
No association	Description	"hverdag"	"alle utenom lørdag og søndag"	*weekday → all but Saturday and Sunday*
	No answer	"Nyte"	"nei, det har jeg ikke noe ord til"	*enjoy → no, I don't have a word for that*
	Unrelated	"informere"	"hanske"	*inform → glove*
	Personal	"religion	"bra"	*religion → good*

4.3.7.4 Error analysis for picture naming

For the picture naming task, nine scoring categories were identified based on responses provided by participants in the norming sample. These categories were used to judge the participants' understanding of the tested items, and to evaluate if the persons with dementia have an impairment that is mainly due to accessing the correct item or if there are underlying semantic impairments which result in lexical retrieval difficulties. Table 4.7 gives an overview of the scoring categories and an example of each.

Results from the naming tests were scored by three different raters: one rater scored all 38 participants, and two raters scored 15 participants each — of which three overlapped between the two. Inter-rater reliability was calculated with Fleiss' exact kappa using the package *irr* in R. A kappa score is a number between 0 and 1, where numbers close to 0 indicates poor agreement between the raters, and a score closer to 1 indicates near perfect agreement between the raters. The reliability was good between the three raters: for Nouns, the kappa = 1, and for Verbs, the kappa = 0.926. Inter-rater reliability is considered acceptable for every number above 0.8 (McHugh, 2012).

Table 4.7: Response categories for the naming tasks

	Target	**Response**	**English translation**
Synonym	"brekke"	"knekke"	both *to snap, to break*
Hyponym	"flue"	"insekt"	*fly → insect*
Hypernym	"sekk"	"ryggsekk"	*bag → rucksack*
Description	"pingvin"	"sort og hvit fugl"	*penguin → black and white bird*
Code switch	"ananas"	"pineapple"	*English translation of target*
Different word class	"danse (v.)	"dans (n.)	*to dance → a dance*
Visual confusion	"palett"	"pensel"	*palette → paintbrush* (both in picture)
Semantic confusion	"hval"	"sel"	*whale → seal*
No answer	"løk"	"jeg vet ikke"	*onion →* "*I don't know*"

4.4 Study III - Sentence comprehension

The third study was designed to investigate sentence comprehension deficits in dementia as a factor of change in cognitive abilities, such as working memory and executive functioning. In this study, sentence comprehension was measured by means of an eye tracker, as well as behaviorally through judgment. This and the previous study rely on the same materials, therefore that information will not be repeated in this section.

The same participants as in the previous study were also included in this study. However, the number of participants was reduced. Several of the participants were excluded, as it was very difficult to obtain good eye tracking data from them.[7] The total number of participants reported on in this paper is thus 16. Due to the low number of young control participants with satisfactory eye tracking data, or enough recordings, this whole group had to be excluded from the study. Demographics of the included participants are found in table 4.8.

Table 4.8: Demographics of participants in study III

	AD (N=5)	PPA (N= 2)	Control (N=9)
Gender (M:F)	2:3	2:0	4:5
Age Age span	m: 74 (SD: 4,6) 66-79	75;81	m: 70 (SD: 4,34) 66-79
Education Education span	m: 13,8 (SD: 2,71) 10-18	15;14	m: 17,61 (SD: 0,73) 15-20

[7] It is not possible for me to determine the exact reason why the eye movements of these persons were unable to be tracked, as there can be many possible explanations for this. Reasons for poor tracking data can be droopy eyelids, glasses, eye makeup and other external factors (e.g., poor lighting conditions) which makes infra-red recording of the cornea difficult.

4.4.1 Visual world paradigm

Sentence comprehension was assessed by employing a sentence-to-picture matching paradigm combined with eye tracking methodology. Eye tracking allows us to study language processing "online", as it happens. By studying eye movements during a language comprehension task, one can pinpoint an individual's exact attention shifts as they happen, since gaze-shifts are considered to take place approximately 250 milliseconds (ms) after processing of the audio stimuli takes place. Eye tracking data can give some information about what is happening as language is processed, but it cannot tell us exactly what goes on in our brains. When (auditory) linguistic information is coupled with visual information, one can measure how the attention is driven by the sentence information in real time (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

The test paradigm used for studying sentence comprehension in study III is a so-called *visual world paradigm* (VWP) which includes both images and sound. In VWP, one usually measures how long it takes from when the participant hears a critical word until they direct their gaze towards the target picture (Tanenhaus et al., 1995). This paradigm is based on the so-called *eye-mind hypothesis*, which assumes that people automatically look towards what the brain is processing.

Three sentence types were tested in this experiment: active sentences, subject cleft sentences, and object cleft sentences. Cleft sentences have previously been identified as problematic for participants with AD and PPA (Croot et al., 1999; Thompson, Ballard, Tait, Weintraub, & Mesulam, 1997, see chapter 3). Each sentence was split into three *regions of interest* (ROI) for the eye tracking analysis (more on this in 4.4.4).

An example of each sentence type with its English translation is provided in table 4.9, while a full list of test sentences can be found in Appendix F.

Table 4.9: Regions of interest in the sentence comprehension experiment

ACTIVE (A)	*Norwegian* *English*	Gutten The boy	maler paints	jenta the girl
Region of interest		Region 1	Critical region	Region 2
SUBJECT CLEFT (SC)	*Norwegian* *English*	Det er gutten som It is the boy who	maler paints	jenta the girl
Region of interest		Region 1	Critical region	Region 2
OBJECT CLEFT (OC)	*Norwegian* *English*	Det er jenta It is the girl	gutten the boy	maler paints
Region of interest		Region 1	Region 2	Critical region

4.4.2 Selection of test items

The experiment consist of 45 reversible sentences - sentences that are still meaningful even when the subject and the object reverse roles (i.e. 'The boy paints the girl' in table 4.9 or in reversed order: 'The girl paints the boy'[8]). The sentences were constructed to fit 15 different verbs, divided into three experimental blocks, with 15 sentences in each block. In all three blocks, each sentence type appeared five times, but each verb only once.

Inspiration for the sentences included in the experiment and corresponding drawings was taken from the sentence comprehension subtest from the *Verbs and sentence test* (VAST) (Bastiaanse, Lind, Moen, & Simonsen, 2006). However, there were certain modifications made to make the test suitable for an eye tracking experiment. These modifications include: reduction from four competing pictures (one target and three distractors) to two (one target and one distractor); addition of an introductory 'context sentence' to familiarize the participants with the actors in the sentences and the pictures; and finally, creation of 15 new sentences based on the discarded distractor pictures to have enough sentences for an experimental setting.

A search in the NorGramBank, a syntactically annotated tree bank (Dyvik et al., 2016; Rosén, De Smedt, Meurer, & Dyvik, 2012), [9] was carried out to check how common each of the three sentence types are in Norwegian. This search returned a total of 781 356 hits divided over the three different sentence types. The distribution between the sentence types was as follows:

Table 4.10: Hits in the NorGramBank

Sentence type	No. sentences	percentage
Active (A)	744 358	95,3%
Subject cleft (SC)	32 342	4,1%
Object cleft (OC)	4 656	0,6%

After the stimuli items were selected, they were programmed into an eye tracking experiment using *Senso-Motoric Instruments* (SMI)'s software *Experiment centre* version 3.6. The sentences were audio recorded by a male professional voice actor. Afterwards, these recordings were manipulated to ensure that the verb in each sentence type came at a specific time slot depending on the sentence type. This time slot is referred to as the *critical ROI*, as it is upon hearing the verb that one can properly parse the sentences in question.

The pictures were quadratic black-and-white line drawings against a white background (550 x 550 pixles). The target alternated between the left and right side of the screen. A grey frame lay around both pictures to physically separate them from each other. All parts of the screen that

[8]Compare to 'The boy paints the apple' and 'The apple paints the boy', where the second sentence does not make sense when the roles are reversed.

[9]The tree bank is part of the Norwegian Infrastructure for the Exploration of Syntax and Semantics, INESS: http://clarino.uib.no/iness/page

are not either the target or the distractor image —that is, the grey frame, and anything outside it — is refered to as *white space*. An example of the test drawings can be seen in figure 4.4.

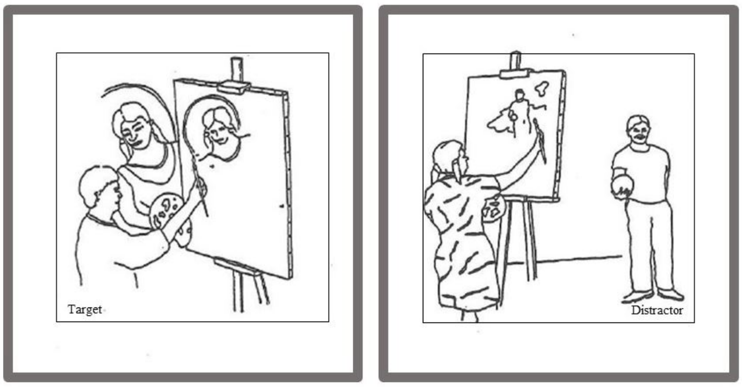


Figure 4.4: Example of a test pane from the Sentence-to-Picture matching task

An *area of interest* (AOI) was assigned to the target picture in each trial, covering a large proportion of the target image. A large AOI is good to capture gazes to target even if calibration is momentarily lost (in cases where participants blink or look away from the screen during testing), or if there is a lot of gaze drift.

Recordings were performed using the SMI RED 250 eye tracking device attached to an external screen linked to a laptop computer. The test administrator ran the experiment from the laptop, and participants saw the pictures on the external screen.

There are some eye tracking studies that have investigated sentence comprehension in healthy aging and in stroke-induced aphasia (Tanenhaus et al., 1995; Bickel, Pantel, Eysenbach, & Schröder, 2000) and single-word comprehension in AD and PPA (Seckin et al., 2016). Furthermore, there are some studies that have used eye tracking methodology to investigate general cognitive processes in AD. To my knowledge, there have not been any eye tracking studies of sentence processing in AD and PPA to date.

In general, studies show that cognitive impairments underlie the many changes in eye movements for patients with Alzheimer's disease. As the disease progresses, AD patients will develop impairments in visual cognition, in addition to the more well-known deficits in episodic and working memory (Ko & Ally, 2011). Researchers have recently directed their attention towards the slowed visual processing as an early indicator of AD, using eye trackers to study cognitive processes in general, and not specifically language processing (Pereira, Camargo, Aprahamian, & Forlenza, 2014; Ko & Ally, 2011; Molitor, Ko, & Ally, 2015).

4.4.3 Procedure

Participants were seated in front of a laptop computer with an SMI RED 250 Hz eye tracker attached to it. Recording frequency was set to 250 Hz. Three participants were mistakenly

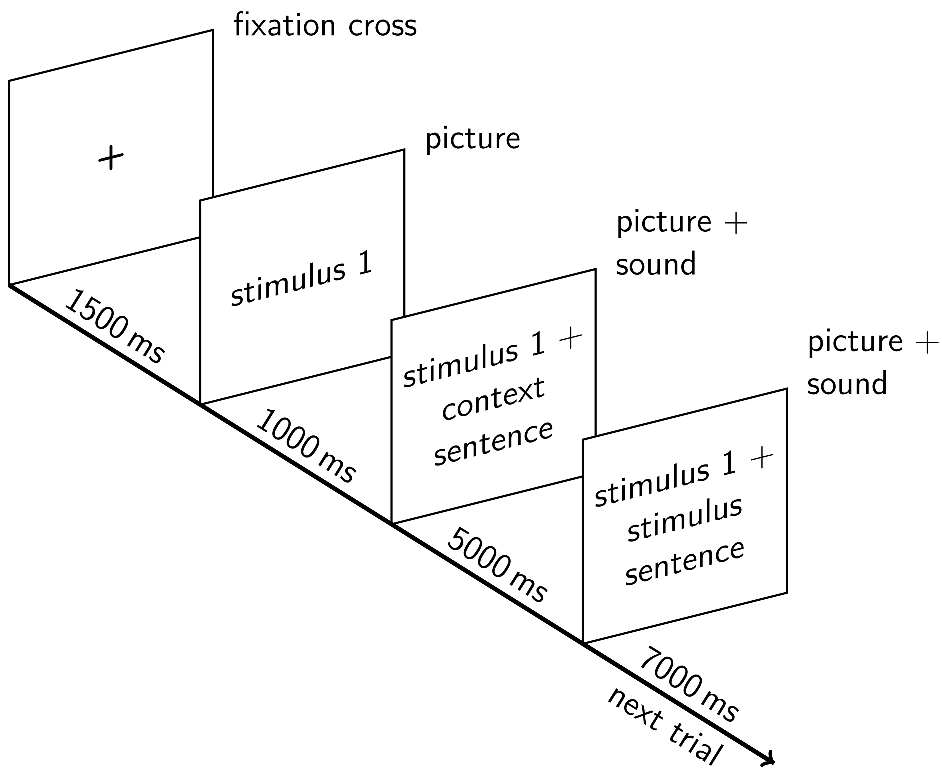


Figure 4.5: Test procedure eye tracking experiment

tracked at a sampling rate of 60 Hz, but this did not influence the data analysis. Between each experimental block there was a short break. Total testing time per participant was approximately 15 minutes for all three experimental blocks, including pauses.

Each trial began with a fixation cross in the middle of the screen, to direct the participants' gaze towards the center. After 1500 ms, two images (the target and the distractor) appeared on the screen. One second (1000 ms) later, the context sentence played (e.g., 'On these two pictures you can see a girl and a boy'). A pause of 2500 ms followed the offset of the context sentence, after which participants heard the stimulus sentence (i.e. 'The boy paints the girl'). After the stimulus sentence had been played, and the participants had indicated whether the left or the right picture corresponded to the stimulus sentence, there was a silence before the next trial automatically started. The full time from stimulus sentence onset until the next fixation cross was 7000 milliseconds. Stimuli were randomized across participants. The test procedure is schematically represented in figure 4.5.

4.4.4 Analysis

The data from the sentence-to-picture matching experiment were analyzed both online and offline. The online analysis includes the gaze data captured with the eye tracker. The offline measure was participants' accuracy in judging which picture corresponds to the spoken stimulus sentence.

The gaze data analysis was performed on the number of fixations from the critical ROI (from verb onset until 250 milliseconds after verb offset). It takes adults about 200 ms to launch a saccade, and 250 ms is a commonly used measure in visual world paradigm studies (Tanenhaus et al., 1995). Fixations after 300ms following the verb offset were considered to be the ones that were triggered by comprehension.

The independent variable for analysis was sentence type (active, subject cleft or object cleft), and within each critical ROI, the proportion of time the participants spent fixating on the target picture was calculated, and treated as the dependent variable. This was determined by applying the following formula:

$$\frac{Target}{(Target + Distractor + WhiteSpace)}$$

The total number of looks to target in a given ROI (sound) was divided by the number of looks to all AOIs in the images — target, distractor and white space —within that same ROI.

Each participants' scores on the MMSE, verbal fluency, Stroop, TMT A + B, and digit span were compared to their online and offline data, to investigate if comprehension is influenced by general cognitive functioning, executive functioning, short-term and/or working memory. All calculations were performed in R ("R: A Language and Environment for Statistical Computing," 2019).

4.5 Study IV - Longitudinal changes

The last study was conducted as part of a pilot project on multilingual dementia in Norway, and aimed to investigate lexical retrieval and naming strategies in the participant's two languages in a longitudinal perspective.

In this study, psycholinguistic elicitation techniques were combined with conversation analysis (CA) to study naming of nouns and verbs in different contexts. The focus here will be on the psycholinguistic part of the study.

4.5.1 Participant and recruitment

The focus in this study is on one participant, JJ, with a probable PPA diagnosis, and his performance on several lexical measures in different languages over the course of 18 months. The neuropsychological tests used in this study were performed at the participant's local hospital in Oslo during four different follow-up appointments. Additional tasks for general cognitive functioning, executive functioning and a linguistic assessment were conducted by members of the *MultiLing Dementia* research group on two occasions.

JJ was recruited through the conversation group for persons with dementia that he attended. The conversation group was organized by the local chapter from the *Nasjonalforeningen for folkehelsen*[10] dementia support group in Oslo. He received written information about the project, and both he and his wife gave written informed consent for him to participate.

JJ was an American national who had lived and worked in Norway for most of his adult life. He was 68 years old at the first test point (TP1), and nearly 70 at the second test point (TP2). JJ had higher academic education, attained in the US. At work, he used both Norwegian and English on a daily basis. He had learned most of his Norwegian through immersion, but also had some formal training. In his family, they spoke mainly English, but with some code-mixing between Norwegian, Swedish (his wife's first language) and English. His wife rated him as very proficient in Norwegian before the onset of his disease, and he spoke Norwegian with friends and neighbors, watched Norwegian television and read Norwegian newspapers. JJ reported that since the onset of his dementia, he had been using Norwegian less frequently.

JJ had been referred to a specialist in neurology at age 67 because of language problems. A *single-photon emission computerized tomography* (SPECT) scan showed bilateral reduced perfusion of large areas of the brain. According to his medical journal, his diagnosis was "A variant of Alzheimer's disease, frontal lobe dementia, primary progressive aphasia". I do not have access to any other information about the underlying pathology, and from the unclear description in his medical journal, it is impossible to say anything certain about the possible subtype of PPA. However, the results from the language testing, and the initial diagnosis as AD could suggest lvPPA.

His first two assessments on cognitive tasks performed at the hospital, show only a mild impairment. This indicates that his cognitive functioning is still intact early in the course of his disease, which is in line with the operational definition of PPA (Gorno-Tempini et al., 2011).

4.5.2 Materials

Cognitive tests in this study included the MMSE, TMT-A+B, the Erikson flanker task (Eriksen & Eriksen, 1974) and the *Rowland Universal Dementia Assessment Scale* (RUDAS) (Rowland, 2009; Wong, Martin-Khan, Rowland, Varghese, & Gray, 2012).

RUDAS and MMSE both measure general cognitive functioning, and assess mainly the same domains. The main difference is that RUDAS was developed as an alternative to MMSE, to be more sensitive in detecting cognitive decline in people with diverse cultural backgrounds. Similar to MMSE the total score is 30, and a score below 26 indicates mild cognitive impairment. RUDAS includes tests for the following domains: memory, spatial orientation, praxis, visuoconstructional drawing, judgment and language.

[10] *The National Association for Public Health* is a non-profit organization in Norway that works both with dementia and heart disease

The Flanker test measures the participants' abilities to suppress irrelevant information, and is generally seen as a measure of inhibition capacities (Eriksen & Eriksen, 1974). This is a computerized test where the participant sees five arrows on a screen. The arrows can either point in the same direction (congruent condition), or the central arrow can point in the opposite direction from the arrows flanking it (incongruent condition). The participant has to judge as quickly as possible, indicating by a key press, if the central arrow points to the left or to the right.

The linguistic battery included a subset of items for confrontation naming from the *Psycholinguistic assessments of language processing in aphasia* (PALPA) (Kay, Lesser, & Coltheart, 1996, 2009), and the *Verb and sentence test* (VAST) (Bastiaanse et al., 2006), and a cartoon description from the *Bilingual Aphasia Test* (BAT) (Paradis, 1998; Knoph, 2010). Recordings of spontaneous speech during a background interview and small talk with the participant are also available. Furthermore, JJ's wife filled in the *Communicative Effectiveness Index* (CETI) (Lomas et al., 1989, 2006), evaluating how she experienced his language impairment since the onset of his dementia.

To test word production through confrontation naming a sub-set of the test items from PALPA and VAST was used. Both tests originally consisted of 40 items each, however, only 30 items were included from each test in the current study to reduce testing load and time. An overview of the selected items can be found in Appendix G. The same set of pictures was used to elicit single verbs and nouns in both English and Norwegian.

The picture naming test was computerized to be able to measure reaction time for each item, programmed in E-prime version 2.0; (Schneider, Eschman, & Zuccolotto, 2002). However, this was not taken into account in the analysis because of poor timing quality. Nouns were elicited before verbs in both languages and at both test points.

The items were not matched for any underlying properties, but a post hoc analysis showed that 16/30 of the nouns and 15/30 of the verbs are cognates in English and Norwegian. A Wilcoxon rank-sum test showed that there was no significant difference in the frequencies of the words across languages. However, the selection of items is slightly skewed, with a few very high-frequency items in both the object and action naming subtests.

The cartoon description task was borrowed from the BAT (Paradis, 1998; Knoph, 2010). In this task, the participant is presented with a six-picture cartoon without any text, and is asked to use the pictures to tell a story. For scoring purposes, ten key content components of the cartoon's plot were identified; these are rendered in 4.5.4.

4.5.3 Procedure

JJ was tested on the language battery at two *test points* (TP), approximately one year apart. Each language was tested separately at the two TPs; first English, and a week later, Norwegian. RUDAS, flanker, picture naming and cartoon description were assessed by members of the MultiLing Dementia team. Before and between the two TPs JJ was followed up at his local hospital where he was also tested with the clock drawing task, TMT A + B and MMSE.

At each test session, two people were present with JJ in his home. To keep each test session as monolingual as possible, different research assistants were present at the different sessions, but the same main examiner (the second author, HGS) was present at all four test sessions.

Each session began with approximately 30 minutes of free talk between JJ and the examiners. In the first session, a biographical interview was conducted to learn as much as possible about JJ and his background, both socially and linguistically.

RUDAS and flanker tasks were only conducted during the English testing session, since this was his strongest language; CETI was filled out by his wife in Norwegian; and picture naming, cartoon description and small talk were recorded in both languages. Table 4.11 gives an overview of which tests and modalities were assessed at both TPs.

Table 4.11: Different recordings in each language at both TPs

Language	Test/Recording
ENGLISH	Interview/Small talk RUDAS Flanker Picture naming Cartoon description
NORWEGIAN	Small talk Picture naming Cartoon description CETI

4.5.4 Analysis

In this study, both qualitative and quantitative measures were employed to investigate JJ's naming skills and strategies in different task environments, though, the conversation analysis data will not be discussed in this dissertation.

For the picture naming task, items were scored as correct or incorrect using *composite scoring*. Composite scoring means that the responses were coded regardless of the language he used to respond (e.g., 'fish' was coded as correct for the Norwegian target 'fisk'). This was a strategy JJ employed frequently, especially for cognate words. The number of light verbs ('to do' or 'to make' etc) in the verb naming task were also counted.

For the cartoon description task, ten key content components that should be included to form a coherent story were identified. These components can be found in table 4.12. The cartoon description was further analyzed in terms of number of produced words and lexical density. The narratives were transcribed using conversation analysis conventions. Examples of transcribed excerpts can be found in paper IV (Lind et al., 2018).

Table 4.12: Content components for the cartoon description task

1	Introduction of the man (main character)
2	Introduction of the woman (main character)
3	Introduction of the birds (main characters)
4	The man climbs the tree
5	The branch snaps
6	The man falls down
7	The man's leg is broken
8	The ambulance arrives
9	The man is taken to the hospital
10	The mother bird cries as the chicks have died

4.6 Reliability and validity

Reliability and validity are considered to be the main measurement properties to assure researchers that a specific tool they use will give an accurate representation of what they wish to investigate. If a tool or an instrument is considered to be accurate, valid, interpretable, and provides robust results, it is suitable for use in clinical and research practice (Souza, Alexandre, & Guirardello, 2017). In this section, I present how the tasks that were specifically developed for the current project (FWA, picture naming, word-to-picture matching and sentence-to-picture matching) are considered to be reliable and valid for the purpose of this project.The validity of the project as such, is discussed towards the end of this section.

Reliability refers to the ability of reproducing consistent results over time, and by different observers. Reliability relies on the function of the instrument, of the population on which it is used, the circumstances it is used in, and on the context. Validity refers to whether the tool measures what it is supposed to measure (Souza et al., 2017).

To establish reliability for the 30-word FWA test, the short list has been tried out on an additional 154 adult L1 speakers of Norwegian, and compared to the responses on the same 30 words from the original 100-word list (Bøyum et al., forthcoming). Similar to the system employed in the current study, two raters each scored all responses, and a third consultant discussed the ratings with the two raters. Alternatively, two raters could have scored either all or a subset of items each, and an inter-rater reliability score could have been calculated.

The results from the comparison between the short and the long tests are slightly different from each other: More meaning-based responses were provided on the short than on the long test, and more form-based responses were given on the long, compared to the short test. This might be due to the different modes of presentation — recall that the long test was presented to participants in writing, and the short test was administered orally. It might be the case that form-based responses are more readily available when a person is presented with written cues than when they are presented with spoken cue words. However, the 30-word list is a reliable substitution for the 100-word list, and can be used in situations where testing time and load need to be reduced.

In addition to the participants included in the current project, the Picture Naming and the word-to-picture matching tests were also tried out on a set of 20 neurologically healthy L1 speakers of Norwegian, aged 25-65. Participants first named the pictures in the picture matching task, before they conducted the word-to-picture matching task. A full norming study of the two tasks was beyond the scope of this project.

However, the results from this small-scale validation study indicate that the Picture Naming test has high reliability, with the participants scoring 98,6% correct on average. For the word-to-picture matching task, all participants scored 100% correct, implying that the comprehension task was easier than the production task.

The results from this small-scale trial of the picture naming and word-picture matching tasks were only scored by one person, and no inter-rater reliability score has been obtained. However, for the larger population who participated in the study reported on in paper II, the inter-rater reliability between the three raters was very high for both verbs and nouns (see 4.3.7.4). High consensus between the raters can indicate that the scores reflect what the test was intended to measure (Mildner, 2013).

The original sentence-to-picture matching task from the VAST has been normed on a sample of both neurologically healthy persons, and persons with aphasia (Bastiaanse et al., 2006; Lind, Moen, & Simonsen, 2006). However, this test has been altered to fit the test paradigm employed in the current study. The version in the current experiment was not validated before testing commenced, but rather piloted on a small set of five volunteers. These volunteers, and all but one of the control participants in the current study scored 100% correct on the sentence-to-picture matching task. This indicates that the task was not particularly difficult for neurologically healthy participants. The task was somewhat more difficult for persons with AD and PPA, but their eye tracking data show less impairment.

An important question is whether this current experiment is in fact measuring sentence comprehension, or even compliance with task instructions. With only two pictures, there is a possibility that the recorded results are purely based on chance performance and guessing. However, both the online data and the offline judgments show that participants do chose the correct picture more than 90% of the time, which would strongly argue that participants are not guessing.

Although all four tests show high stability and reliability (the same results are reproduced for larger groups of participants), this does not automatically imply high validity. Validity is found when the test is proven to measure what it is supposed to measure. In this case, lexical retrieval for both production and comprehension, and sentence comprehension.

Picture naming, word-to-picture and sentence-to-picture matching tasks are some of the most common paradigms in language assessment. Picture naming tests often have high reliability, with test-retest correlations upwards of 0.9 (Howard, Patterson, Franklin, Morton, & Orchard-

Lisle, 1984). According to the Pearson correlation coefficient; 1 is considered perfect reliability, and everything above 0.9 is considered excellent reliability (Kirch, 2008). Regarding the word-to-picture and sentence-to-picture matching paradigms, these have been found to be a reliable tool for research and clinical assessment of comprehension of verbs and sentences (Bastiaanse, Edwards, Mass, & Rispens, 2003). Further research and trialling of the tests included in the current studies should add to the validity of these specific tests.

The aim has been to design a study with high internal, external and ecological validity, and at the same time being suitable for the population in question. Internal validity is strengthened by experiments that are well designed, carefully controlled and accurately measured, to ensure that any alternative explanation for the phenomena under consideration can be excluded (McDermott, 2011). Threats to the internal validity are experimental procedures, or the participants' experiences which influence the researcher's ability to draw correct inferences about the population under investigation based on the collected data (Creswell, 2009). In the current project, I have tried to control for confounding variables which can be a threat to the validity of the studies by keeping the test protocol similar for all participants within groups and collecting as much subject-relevant information as possible. However, there are some variables that it is not possible to control for, such as fatigue, test-wiseness or lack of interest or motivation, which can affect the participants' performance.

One way to assess the external validity of a study, is to replicate the study to see if the same results are obtained in different contexts and settings and with other participants. This strengthens the transferability of the results. The methods and procedures in the different studies have been described in detail in this chapter and in the specific papers. This is important to allow for replication of the studies. The ecological validity refers to the applicability of the experimental findings in the real world. The current project has a high degree of experimental control, which adds to the internal validity. However, this also means that testing at times becomes unnatural. To mimic more naturalistic language environments, production of full sentences and narratives was targeted in studies II and IV.

4.7 Ethical considerations

The different sub-studies were approved by the *Regional committees for medical and health research ethics*[11] with the following information: studies II and III, reference number *2016/1293*; study IV; reference number *2014/1993D*. Study IV was also approved by the Norwegian centre for research data (NSD), with reference number: *40523/3/HIT*.[12] Letters to the participants are found in Appendix A, and all approval notices are found in Appendix B. The data protection services at the University of Oslo also approved the project plan and plans for storage and treatment

[11] Regionale komiteer for medisinsk og helsefaglig forskningsetikk; https://helseforskning.etikkom.no
[12] Norsk senter for forskningsdata; https://nsd.no

of the collected data. In addition, the data protection officials at the Oslo University Hospital approved studies II and III. This was necessary to obtain information about the participants who were recruited from the NorCog register (see 3.4.1).

The participants with AD and PPA in studies II and III were recruited at the memory clinic after they had received their diagnosis, and agreed to be included in a national register for persons with cognitive impairments. The leader of the memory clinic contacted potential participants. Only if they agreed to be contacted to receive information about the project were their contact details given to me. Potential participants were first called to ask if they consented to receive an information package by mail. The information package contained an information letter about the purpose of the project and what participation would entail, a consent form and a response envelope. A follow-up phone call was made approximately one week after the participants were expected to have received the information package. Only the participants who returned the consent form, either spontaneously or after the follow-up phone call, were considered for inclusion.

The information letter was written in clear language, with short and precise sentences to enhance readability. Participants were informed that if they chose to participate, they were allowed to withdraw at any moment without having to provide a reason for this. The information letter had to be updated in the summer of 2018 to account for changes in the European data protection legislation and the introduction of the *General Data Protection Regulation* implemented by the EU (GDPR). All participants who had received an information letter before the implementations of the GDPR were sent a new letter with information about the new data protection guidelines, and a request to sign a new consent form. Updated consent was obtained from all relevant participants.

JJ (study IV) was recruited through the conversation group he attended. The leader of this conversation group approached him with a letter containing information about the purpose of the study, and he subsequently contacted us, agreeing to participate in the study. He signed a consent form at the beginning of the first session at each TP.

Obtaining consent from persons with dementia can be challenging. However, all participants in these studies had been evaluated by their physicians to be capable of giving consent. In early stages of the disease, participants are usually competent enough to give informed consent, but as the cognitive capacities decline, this might no longer be the case. It was therefore important to obtain consent from JJ again before the start of the second round of testing.

All participants included in the studies have been anonymized. However, in a country with a relatively sparse population, and participants recruited from the same geographical area, the question of anonymity is difficult; there is a risk that participants are recognizable based on only few cues, such as the combination of language background, age, gender and diagnosis — particularly for participants with rare diagnoses. Nevertheless, all participants have given their permission for the procedure and the results to be published in this dissertation, in scientific publications, in dissemination to relevant scientific fora and in presentations to students.

5
Main features of the studies

This chapter aims to further introduce the four studies. The main findings from each will be pointed out. After each introduction, there is a section including some commentary to each study. For a comprehensive discussion of the findings and results, see chapter 6.

5.1 Study I

The first study that is included in this dissertation is a systematic review of (psycho-)linguistic variables and lexical retrieval in AD and PPA. The main goals were to investigate which variables most affect production and comprehension of single words. A second aim was to investigate if an analysis of these variables can help distinguish between persons with dementia and control participants, and between AD and PPA. The variables under review were *word class*, *frequency*, *age of acquisition* (AoA) and *imageability*.

A majority of the reviewed studies found a dissociation between nouns and verbs (e.g., Druks et al., 2006; Kim & Thompson, 2004; Hernández et al., 2007), while one did not (Rodríguez-Ferreiro, Davies, González-Nosti, Barbón, & Cuetos, 2009). Among the studies that reported a dissociation between nouns and verbs, both a selective sparing for nouns (e.g., Druks & Weekes, 2013; White-Devine et al., 1996; Hillis et al., 2006; Lind et al., 2018; K. Robinson, Grossman, White-Devine, & D'Esposito, 1996) as well as a selective sparing for verbs (e.g., Hernández et al., 2007; G. Robinson, Rossor, & Cipolotti, 1999; Williamson et al., 1998; Hillis et al., 2004) were found.

There can be several possible reasons for these divergent results. Psycholinguistic variables affect words in different ways, and it is known that the variables *frequency*, *AoA* and *imageability* behave differently for nouns and verbs. There might be an interplay between the different variables that affect naming differences across word classes, rather than the word class membership per se. Another important point to make is that the way the different studies test for naming and comprehension impairment differs across the reviewed studies, which might also influence the results.

AoA was the variable that showed most stable effects across testing modalities and both in comprehension and production. This indicates that the words that are learned early in life are the words that are retained longest. Explanations for this include more entrenchment and plasticity of the neural system. Words that are encountered often will get more entrenched in the mental lexicon, and more entrenchment leads to more new words attaching to this word in later learning. When more words attach to a node, the nodes adapt to again accept more attaching words.

Another finding from this review is that anomia in AD and PPA seems to be primarily a difficulty with access to the phonological form, rather than an underlying semantic impairment. This implies that the semantic links in the mental lexicon are stronger than the phonological ones, and trigger enough excitation for lexical items to reach activation for production or comprehension. This is supported by the results of the majority of the reviewed studies, and by the fact that the main semantic variable under review (imageability) does not seem to affect lexical access in the same manner as AoA and frequency, which are variables that are associated with retrieval.

The results from this study suggest that the language impairments in AD and PPA are quite similar, despite the differences in underlying neural change. This finding is supported by the fact that in most cases, it is not possible to distinguish between AD and PPA on the basis of sensitivity of the studied variables in linguistic data. However, it is possible to use this data to differentiate between the different subtypes of PPA, and between healthy older adults and older adults with AD and/or PPA.

5.1.1 Comments to study I

This paper would have benefited from a more thorough methodological and theoretical discussion. Preferably, I should have included a section including information and discussion about *how* measures of the different variables are obtained. For instance, AoA and imageability are obtained by asking people how old they think they were when they learned a word (for AoA), or how easily they feel that a word gives rise to a sensory mental image (for imageability). Such subjective rating may seem far-fetched, but several studies show that the measures are reliable. Subjective AoA correlates very well with objective AoA as obtained respectively by parental reports and observational studies of children (Hansen, 2016; Łuniewska et al., 2016). Imageability and concreteness correlate to a great degree, but imageability gives more information about the conceptual experience of words that is otherwise missing from judgment of concreteness. There are also many other psycholinguistic variables that affect lexical retrieval, but that were not included in this review. For instance, concreteness, familiarity, word length and phonological variables such as similarity to other words, complexity, etc. to name a few.

Theoretically, a discussion of psycholinguistic processing models, and a closer tie between the findings and a discussion of such models, should have been included. Many theories of language processing are based on evidence from language impaired speakers. Language impairment in dementia is well-suited to use as evidence for psycholinguistic processing models, since it can be assumed that persons with dementia had an intact language system before onset of the disease.

It is quite difficult to make concrete statements about how the different variables influence lexical retrieval in PPA, since I cannot be certain about the type of PPA that is reported on in many cases. In the reviewed studies, there is not much information about which diagnoses the participants have. This is partly due to the large number of studies that were published before the consensus criteria for PPA were in place, but also and in some cases, the PPA subtype is not even reported.

Furthermore, only few articles included participants with lvPPA, either because these participants were left out by the authors (Fraser et al., 2014; Marcotte et al., 2014), or because the majority of the studies on PPA do not give enough information about the type of PPA studied. The most similar results are to be expected between lvPPA and AD, however a proper comparison between the results from the studies including participants with PPA to studies including participants with AD, becomes difficult when the grounds for comparison are lacking.

For a good review to be valid, it should be updated on a regular basis to account for the latest research in the field. A future update of this article should also include a thorough discussion of methodological considerations and theoretical implications.

5.2 Study II

The second study was concerned with how free word associations are affected by cognitive decline in AD and PPA; how different tests of lexical access and association affect naming abilities in PPA and AD; and how results from different tests can inform about lexical processing in general, and for individuals with AD and PPA specifically. The properties that were identified as particularly important for naming success in the literature review (frequency, AoA and imageability) were also taken into consideration in developing a naming and comprehension task included in this study.

Different tests for lexical access can yield different results, as the degree of restriction within the task influences the ease of lexical retrieval. In *picture naming* tasks, the participants are asked to name one specific target word, which makes the word search very restricted. In controlled word association tasks, like *verbal fluency*, the task is slightly less restrictive, as the participants are asked to name as many words belonging to a semantic category or starting with a particular sound; this somewhat restricts the word search. *Free word associations* are far less restrictive, and participant can in theory say any word as a response to a cue. Participants in this study were also assessed with the picture version of the *Pyramids and Palm Trees* test and *De semantische associatie test*, where the semantic associations are purely conceptual; participants have to retrieve information about the concepts and not the words. These tests are also considered restricted, as there is only one answer that is considered correct per trial.

Results from the traditional language assessment tests show that there is generally no dissociation between nouns and verbs for most participants, which is different from the findings regarding word class differences in study I.

Furthermore, the results from this study indicate that there are subtle differences in language behavior between persons with AD and PPA and neurologically healthy controls which cannot be captured by traditional, restrictive tests. In the free word association test, participants with AD provide more responses that are related to the cue words on multiple levels in terms of form, meaning and usage patterns. For instance, their responses can be both form and meaning-related to the cue, or both meaning-related and collocations. They also provide more pure collocations than do controls. These differences are found even in early or mild AD, such as for AD04 and AD06, whose scores on the neuropsychological tests show hardly any impairment.

With only two participants, it is of course not possible to establish a pattern of word association responses that can discriminate PPA from AD, or from neurologically healthy persons for that matter. However, there are marked differences between the two participants with PPA on several tests in this study. This might be related to the fact that PPA01's initial diagnosis was one of AD. Throughout, his scoring profile appears very similar to the participants with AD, but with more word-finding difficulties in his spontaneous speech. He also had more difficulty with the picture naming tests compared to PPA02.

Theoretical implications from this study include further support for an impairment in access rather than impairment in the semantic system per se. The findings also further support a hypothesis that there is interactive activation between nodes in the mental lexicon network (as opposed to serial processing). This study gives support to the *transmission deficit hypothesis* (TDH) (Burke & Shafto, 2004), rather than to the *inhibition deficit hypothesis* (IDH) (Hasher & Zacks, 1988).

The clinical implication found in this study, is that analysis of FWA results can point to some differences between neurologically healthy persons and persons with mild to moderate dementia. The differences between these two participant groups are subtle, and require a thorough analysis of the relationship between words.

5.2.1 Comments to study II

The results from this study show that the FWA task can be a valuable supplement to more traditional assessment tools, but it also provides additional information about the language deficits in dementia when used on its own. However, the scoring categories are broad, and with many categories there will always be a risk of ending up with several categories that receive too few responses for meaningful analysis.

An interesting question is why the participants with AD, and to some degree also the ones with PPA, seem to prefer collocation-based responses. One possible explanation for this could be that collocations are more automated, and are activated in the same way as chunks and prefabs in the mental lexicon.

An interesting observation is that the two participants with PPA behaved quite differently on several of the lexical, but also on the cognitive tasks. Since I do not have access to neuroimaging data from these two participants, I cannot make any assumptions as to whether the observed differences are due to differences in neuropathology. However, in line with current research on the matter (Matias-Guiu et al., 2019; Vandenberghe, 2016) there might be two different logopenic variants of PPA that can be distinguished based on which areas of the brain experience more atrophy, and also on results from different linguistic outcome measures.

Future research should include larger population samples, especially persons with PPA, to investigate if the findings from this study can be attributed to change associated with language impairment because of AD/PPA, or if it is an indication of individual differences.

The scoring system also needs to be discussed in greater detail. Two categories (syntagmatic vs. paradigmatic) are not enough to pick up on the many different relationships between cue and response, but too many categories will make the boundaries between the different categories unclear and confusing.

5.3 Study III

The third study moves beyond the single-word level, and investigates both online and offline sentence comprehension in AD, PPA and healthy aging. The focus of this study is on the comprehension of canonical and non-canonical sentences (i.e., active and clefts) in Norwegian. Sentence comprehension in dementia is far less studied than word production and comprehension, and it is therefore interesting to go more in depth on this issue.

Since lexical knowledge and comprehension appear to be spared in AD and PPA, the sentence comprehension deficits they experience do not stem from difficulties with comprehending the individual words of the sentences. Some previous studies have pointed to (working) memory impairments as the main cause for the sentence comprehension deficit in AD, but here it is contrasted with lexical skills in addition to results on tests for working memory and executive functioning.

Because of poor tracking ratio for a majority of the young participants from study II, the whole group was excluded from analysis. Further, five healthy older controls and two participants with AD had to be omitted for the same reason.

The results from the offline measures showed that the participants with AD in general fared better on this task than the participants with PPA, but neither group seems to indicate an impairment. However, the participants with PPA show better immediate attention to the target in the eye tracking data compared to the participants with AD. Furthermore, there is quite some individual variation between participants.

From the eye tracking data, it is clear that healthy older adults always detect the target picture within the region of interest, or right after, for all three sentence types — which is to be expected. As a group, the participants with AD show a similar gaze pattern as controls, however, they are slower and never reach a very high proportion of looks to target (between 55% and 75%). This indicates that these participants are less certain about which image is the target and which is the distractor, yet they have a preference for the target. The participants with PPA are also slower in detecting the target than controls, but their proportions of looks to target is higher than for participants with AD.

An interesting finding is that all three participant groups seem to prefer an agent-first parsing strategy, as they all initially look to the distractor image (almost 70% looks to the distractor for control and PPA participants, and 60% for AD participants) in the *object cleft* (OC) condition. Once more of the sentence becomes available to the listeners, their gazes shift towards the target, and control participants reach up to 90% looks to target by the end of the region of interest. For participants with PPA and AD, the proportion of looks to target in the OC condition is never over 70%.

Unfortunately, the eye tracking data from a large proportion of the participants were deemed to be unsuited for analysis, and the results can only indicate a trend at best. However, the trend is interesting, because it shows that persons with AD and PPA do comprehend even infrequent, non-canonical sentences, although with some uncertainty and longer response times than healthy controls.

5.3.1 Comments to study III

Just like with the previous study, the number of participants in this study is far too low to make sound generalizations about the results. This is extra problematic when a subset of participants were excluded due to poor eye tracking data, or because they expressed little interest in the task.

The technical difficulties with the eye tracking were in some cases related to poor quality of the calibration, especially for the older controls and the persons with AD. Some researchers have discussed how greater gaze drift for persons with AD may be of clinical relevance (Ko & Ally, 2011). More research on a bigger group of participants is needed to see if this is the same pattern observed in the current study.

There are two further methodological issues related to this study, which should have been taken into account when designing the experiment. First, there is no measure of reaction time for the offline judgment. When this measure is lacking, it is not possible to align the online and offline accuracy results (looks to target and pointing, respectively). Some of the participants with AD and PPA show two peaks of looks to target, one immediate and one delayed. If information about their offline reaction time had been available, this could have be aligned with the eye tracking data to see if the second peak of looks to target is related to looking back to the target picture after making the offline judgement, or if this second peak represents something else entirely.

Second, an addition would have been to include a test for sentence repetition, as this seems to be impaired in PPA and, to some degree, also in AD. Testing sentence repetition in addition to sentence-to-picture matching would have given us an even clearer picture of the sentence impairment for persons with these two diseases, which again would support both clinical and theoretical implications from this study.

Finally, more than half of the control participants were excluded due to poor eye tracking data. This means that when employing eye tracking methodology, one needs to be extra attentive to the huge amounts of potential data loss. In this case, I thought I would be safe with the number of participants that I had recruited for the lexical retrieval study, but should in fact have recruited at least double that number for this substudy.

5.4 Study IV

The last study was initiated to investigate how bilinguals with Alzheimer's disease fare on different naming tests, and how lexical retrieval problems are apparent in different modalities (free conversation, semi-structured narratives and experimental testing) and across languages. Our participant, JJ, was recruited through a conversation group for people with AD. However, after he had gone through two rounds of testing, it was clear that he did not in fact have AD, but rather PPA. The diagnosis is not clearly stated in his medical journal, where AD is listed as his main diagnosis and the doctor has made a note that it might be "a subtype of AD called primary progressive aphasia".

The main focus of this study is on longitudinal aspects of language decline in dementia, and on how one can use different methodologies to study lexical retrieval. Speech samples were elicited through picture naming, narrative production and personal interviews. The results show that JJ had similar issues with singular word activation in spontaneous speech as in confrontation naming.

JJ was tested in both his languages — English and Norwegian — on two separate occasions, approximately a year and a half apart. At the first test point (TP), roughly a year post diagnosis, JJ was very communicative and talked a lot. It was at this point possible to keep a conversation going with him, and he underwent the tests with few difficulties. One year later, his communicative skills had decayed and so had his spontaneous language production. It was no longer easy to guess what he was trying to convey. Confrontation naming and narrative production were also more effortful and difficult for him to perform than at the first TP.

At the first TP, JJ's naming skills were better in English than in Norwegian, and he named nouns better than verbs. His speech was halting with many pauses, and except for a few slips of tongue, always grammatically correct. His halting speech and naming errors made the cartoon description difficult, but the main story was conveyed. There were some instances of code switching to English during the Norwegian assessment, but never the other way around. This might be consistent with his language behavior before diagnosis, as reported by his wife, but there is no other data available on this issue.

At the second TP, 30 months post diagnosis, naming was poor in both languages on the confrontation naming test and even more so in spontaneous speech. At this point, it was almost impossible to keep a conversation going with JJ, especially in Norwegian. Verbs were still far more impaired than nouns in English, but, surprisingly, he performed marginally better on verb naming compared to noun naming in Norwegian. In the cartoon description, he provided very little content, despite producing more words than at the first TP. This was observed for both languages.

This study also supports the views of lexical activation put forward in the transmission deficit hypothesis (Burke & Shafto, 2004). Further, the indications of parallel decline in JJ's two languages is indicative of (at least) partially overlapping lexica for bilingual speakers. This study gives a good account of how language abilities deteriorate during the course of the disease.

5.4.1 Comments to Study IV

The spontaneous speech data in this paper has been analyzed using conversation analysis (CA), which is a good tool to describe the events that take place during conversation. This is a method that I do not have much experience working with personally, and since the main focus in this dissertation has been on psycho-linguistic methodologies, this part of the paper has not received much attention in the previous chapters. However, I see clear benefits of including CA analysis to better capture the word searches that take place in conversation and narrative production.

In this study, the naming data was analyzed only in relation to the words' frequency, and find that for noun naming in his L1, and both noun and verb naming in his L2, JJ experienced a frequency effect (better performance on high-frequency nouns). It would be interesting to also look at effects from other psycholinguistic variables, such as AoA and imageability. As information about these variables is available for the words tested in Norwegian through the

Norwegian Words database (Lind et al., 2013), these variables could have been taken into account in the results. However, it is uncertain how, for instance, AoA — as rated by first-language speakers of Norwegian — affect naming abilities for someone who learned the language as an adult.

JJ was tested in both his L1 and L2 with the same items and same pictures, meaning there could be a risk of familiarity and learning effects at the second session at each TP. However, as seen by the much poorer results in Norwegian compared to English, this does not seem to be the case.[1]

Using the same pictures to assess items in different languages means that items have not been controlled for underlying variables that might differ across languages. Preferably, the English versions of the same tests (PALPA and VAST) should have been used as well, to better assess individual items in each language. But as the Norwegian versions of these tests are based on the original English version (PALPA), and partly on the English translation of the VAST, the tests are, to a large degree, overlapping in the two languages. Norwegian and English also belong to the same language family, and none of the test items were culturally unique to Norwegian. It can therefore be justifiable to assess naming in English by means of tests that were originally adapted from English to Norwegian.

Another major point to discuss regarding this study, is the participant's diagnosis. It was initially communicated that he had AD, but in his medical journal it is noted that it might be PPA. Results on cognitive tests show only a slight cognitive impairment early in the disease, and the main symptom JJ experiences is one of language decline. JJ shows impairment in single-word production, both in confrontation naming and spontaneous speech, but not really impaired comprehension, which is in line with the profile found in lvPPA (see 3.1).

[1] English was tested before Norwegian at both TPs.

6

Discussion and conclusions

In this dissertation, I have studied the language abilities in speakers with Alzheimer's disease and primary progressive aphasia to reach two broad research goals:

1. What can data from persons with an age-related cognitive impairment tell us about how language is organized in the brain?

2. Can language data from persons with dementia aid dementia diagnostics, and can it help to distinguish AD from PPA, and both from healthy aging?

To reach these goals, I put forward research questions addressing three different topics which will be discussed in separate sections below: linguistic aspects in section 6.1, methodological aspects in section 6.2, and clinical aspects in section 6.3. Some limitations are outlined in section 6.4, and ideas for further research are introduced in 6.5. Section 6.6 concludes this dissertation.

6.1 Linguistic aspects

As introduced in chapter 2.1.2 and in chapter 3, there are variables associated with a word's form, usage pattern and meaning which influence how items are stored and processed in the mental lexicon (see section 2.3). In this section, the same issues are revisited, but with a focus on how the findings from the current project relate to these psycholinguistic variables (section 6.1.1) and to linguistic theories (section 6.1.2).

6.1.1 Psycholinguistic variables

Frequency is one of the most studied variables which affect lexical retrieval in naming, together with age of acquisition (AoA), imageability and linguistic variables such as word class and cognate status. The variables can affect language processing in different ways, and on different levels of processing (Vonk et al., 2019).

The results from three of the studies in the current project find frequency effects on different tasks, both word and construction frequency influence processing. In the literature review in study I (Ribu, Under revision), frequency is one of the most stable variables across testing modalities— both for production and comprehension. Frequency effects are also found in study IV (Lind et al., 2018), where JJ's naming is better on high-frequency items than on low-frequency items. In study III (Ribu & Kuzmina, submitted), we see that an agent-first parsing strategy, which is mediated by the frequency of structures where the agent occurs first in the sentence, is preferred for persons with and without dementia on object cleft sentences. These findings suggest that there is a frequency effect on several levels, from single words to sentence-sized constructions.

The literature review also lends support to another variable which is associated with lexical retrieval success, namely age of acquisition (AoA). AoA effects are found in a range of different test modalities — and both for production and comprehension — even when frequency effects are less pronounced. AoA effects can be seen as a consequence of the additive construction of semantic representations, whereby later acquired words are incorporated into a representation already containing earlier acquired words (Chang, Monaghan, & Welbourne, 2019). Early AoA words have richer, more embedded semantic representations than later acquired words (Li, Farkas, & MacWhinney, 2004; Brysbaert & Ghyselinck, 2006). Another view is that AoA effects are due to early plasticity on the learning of mappings between written, spoken and semantic forms of the vocabulary (A. W. Ellis & Lambon Ralph, 2000).

Different psycholinguistic variables were taken into account when developing the Picture Naming and word-to-picture matching tests in study II (Ribu et al., submitted), and this test was used partially to investigate the noun-verb dissociation which is often reported in the literature (e.g., Druks et al., 2006; Kambanaros & Grohmann, 2015). In the literature review selective sparing was found for both nouns and verbs, and it was therefore of interest to look deeper into this issue. On the Picture Naming task and the word-to-picture matching task, no difference between word classes was found (Ribu et al., submitted). However, word class differences were found in naming for both L1 and L2 in study IV (Lind et al., 2018), albeit only when composite scoring of the participant's two languages was employed. The conflicting results regarding word class dissociation in AD and PPA may imply that it is not the word class per se which complicates lexical retrieval, but rather an interplay between several different psycholinguistic variables, such as frequency and AoA, on the different test items.

JJ's naming success in his L2 is greatly influenced by the cognate status of the words in the naming task. He code switches regularly to English when naming pictures in Norwegian, which means that he activates words which are similar in both form and meaning, and that he has both his languages active at all times. Some participants, mostly the control participants, in studies II and III also respond with words in other languages than Norwegian on certain tasks. In most cases, they respond with words which are cognates between Norwegian and the language they switch to, suggesting that words which are similar in both form and meaning are especially easy to retrieve. This indicates that even "monolingual" speakers are in fact not monolingual, but have activation from several languages present at all times.

6.1.2 Theories of language and processing models

The findings from this dissertation fit well with usage-based theories of language, which emphasize the roles usage patterns, semantics and phonology play in language processing. Frequency effects are not only restricted to the single-word level, but are also found in sentence processing. AoA can be seen as a cumulative frequency effect, where words with early AoA will have stronger entrenchment than words with later AoA, which leads to stronger mental representations. Cognates also strengthen the connections between units in the lexicon since these words are similar in both form and meaning, which will influence processing of these words.

Usage-based theories of language presuppose that language is shaped and changed through usage patterns and experiences, and that there are differences across languages, communicative settings (including tasks and modalities), between groups of individuals and also between individuals. This makes it a good candidate to explain the variability across participants and tasks reported on in this dissertation.

The changes found in the three experimental studies looking at impairment in single-word naming and sentence comprehension can be seen as augmented manifestations of changes observed in healthy aging. This means that language changes can be scaled on a continuum from healthy to "pathological" changes, and that there might not be a need for a separate model that explains language impairment in AD/PPA contra healthy aging. A model which accounts for why lexical retrieval is compromised in healthy aging (the transmission deficit hypothesis, TDH, Burke and Shafto (2004), which predicts larger age-related changes in production than in comprehension, see 2.2) should also suffice to explain what happens when cognitive impairment enters the scene. The changes seen in AD and PPA are reinforced manifestations of the changes explained by the TDH.

In studies I, II and IV, I find support for an impairment in access rather than in underlying semantic knowledge. In study I, this is seen by the stronger effect from variables that are associated with retrieval (frequency and AoA), than variables related to semantics (imageability) (Ribu, Under revision). In studies II and IV, the access impairment approach is supported by the responses provided on different tests (Ribu et al., submitted; Lind et al., 2018). Synonyms and other semantically related responses on the Picture Naming task demonstrate that the participants have an understanding of the target object, but the phonological form is inaccessible. Similarly, JJ in study IV code switches between Norwegian and English in the Norwegian testing sessions, which indicates that he can activate the correct intended concept, but not its phonological form. His strategy is to produce an output form which is semantically overlapping, and phonologically similar to the target. Control participants in study II also provide code switching responses on the verbal fluency, picture naming and FWA tests (mainly to English and Scandinavian languages, but there is also one instance of a participant who named one animal in German[1] on the semantic fluency task), indicating phonological activation in a non-target language.

The instances of code switching are most likely not an impairment in inhibiting the non-target language: during the cognitive tasks, participants in the current study —both controls and participants with dementia— perform well on tests of inhibition. Both JJ and the participants in study II produce words in languages that they assume the test administrator knows, and this must be seen as a compensatory strategy, or as a playful display of language knowledge. Code switching as a compensatory strategy has also been found in another study specifically examining code switching patterns in dementia (Svennevig, Hansen, Simonsen, & Landmark, 2019). In studies II and III, the participants were aware of their code switches, and sometimes this production was fully intentional (e.g., pronunciation of 'kangaroo' with an exaggerated Australian accent). The instances of code switching on the Picture Naming task were often activation of cognates with the Norwegian target word (e.g., 'to ride' for 'å ri'). There is one exception which may indicate an inhibition deficit happened for two participants: They responded 'pineapple' instead of the Norwegian 'ananas' on a picture of a pineapple, but these cases were less common than cognate responses. On the FWA task, the non-Norwegian responses were always translation equivalents to the Norwegian cue word; these were scored as synonyms.

JJ's language profile in study IV shows similar change in the two languages over time. This may indicate a shared mental lexicon, especially for languages which are as closely related as Norwegian and English. The mental lexicon is fully capable of handling several languages at the same time, and a bilingual mental lexicon will also incorporate a monolingual mental lexicon. This fits well with the Multilink model for language processing (Dijkstra et al., 2019b), which presupposes that a bilingual mental lexicon is the default. This model also stresses the importance of psycholinguistic variables for processing, such as frequency, which makes it compatible with usage-based theories of language.

On the FWA test from study II, the free generation allows participants to select items which have a broader range of semantic and form-based relationships to the cue word, proving that units in the mental lexicon are connected along both semantic and phonological lines. Further-

[1] 'Kaninchen' ('rabbit'), rather than the Norwegian cognate 'kanin', seemingly to place emphasis on the diminutive in the German word

more, participants with AD produce more collocations, indicating a preference for automated, entrenched chunks of words. This can also be taken as an account of frequency-mediated storage and processing strategies, and is in line with usage-based theories of language. It also fits well with hypotheses of language processing which are interactive rather than serial in nature, since activation flows back and forth between phonological and semantic levels. This is because both form-based and meaning-based, as well as combinations of form and meaning-based responses are preferred for all participants.

6.2 Methodological aspects

The second focus in this project was to show how methodological triangulation can help broaden the knowledge obtained about language impairment in dementia. The three experimental studies in this dissertation employed several different methodologies. In studies II and IV, the same phenomenon, namely lexical retrieval, was approached from different angles using different tools and measures. In study III, both offline and online methods were used to study sentence comprehension.

6.2.1 Triangulation

Both Ribu et al. (submitted) and Lind et al. (2018) contrast restrictive tests of naming with freely produced data. The difference lies in the kinds of freely produced language data which are in focus. On the FWA task in Ribu et al. (submitted), the freely produced language output is still kept at the single-word level, whereas the cartoon description task and conversational data in Lind et al. (2018) were included to elicit connected speech samples. Both studies show that methodological triangulation — the use of more than one method to gather data — can provide a more thorough exploration of the complexity of lexical retrieval impairments, and reveal different patterns of impairments as well as their extent.

For instance, the FWA results suggest that the freely generated associations are different for people with and without dementia, which means that this test can be used to supplement traditional tests, such as picture naming. The test also provides valuable data if used alone, since it captures changes in linguistic behavior that cannot be captured by the means of restrictive tests. Combining both free and restrictive tests can give a comprehensive representation of a persons naming abilities.

Freely produced linguistic data is also crucial for detecting abnormalities in speech patterns and sentence construction. In study IV, the participant's speech was analyzed in terms of idea density and word retrieval in a narrative production task, and contrasted with picture naming tasks. JJ's poor sentence generation abilities may be linked to his low scores on the verb-naming test, since verb production plays an important role in sentence construction (de Diego Balaguer et al., 2006).

Combining online and offline methodologies is also useful, since one measure can identify changes that are not picked up on by another measure. This is found in the sentence comprehension task in study III (Ribu & Kuzmina, submitted), where the participants with AD and PPA are slower at detecting the target image than controls. Participants with dementia also show less automated processing than controls. This difference was not found in the offline data, but can clearly be seen on the eye tracking data. This is elaborated on in the next section.

6.2.2 Eye tracking

In the third study, there is no evidence for a sentence comprehension impairment based on the offline data alone. However, once the offline and the online data are combined, different processing strategies between neurologically healthy participants and participants with dementia become apparent.

In addition to measuring the proportion of looks to target, which is a measure of the participants' preference for the target image, the gaze data can also be used to measure time to detect target. It is on this last measure we find the greatest difference between participants with AD/PPA and healthy controls. This is in contrast to the reports from Tyler, Cobb, and Graham (1992), who state that online tasks tend to reveal more preserved knowledge than suggested by the offline tasks. In this study, four out of the seven participants with dementia show more uncertainties on the online measure than the offline.

The visual world paradigm, where audio and visual stimuli are combined, is recognized for its possibilities to study processing when it happens. The assumption is that it is not possible to control the rapid movements which take place between fixations, and once we are attending to an object on the screen, this is driven by attentional processes which we cannot override or control (Cooper, 1974; Tanenhaus et al., 2000).

By combining online and offline measures, we show that sentence comprehension ability is highly influenced by the measure used. In the offline task, pointing to the correct picture was used as a measure of post-interpretative processing. The task was time restricted, but the interstimulus interval was long enough to allow participants to make a judgment, which in most cases is correct. The online measure, however, demonstrates that most participants with AD, and one with PPA, show more uncertainty about which image is the target during the auditory stimulus presentation. This uncertainty follows an initial preference for the target image, a preference which cannot be captured on the offline task alone.

6.3 Clinical aspects

The first three studies investigate how language data can be used to distinguish AD and PPA from healthy aging, and (lv)PPA from AD (Ribu, Under revision; Ribu et al., submitted; Ribu & Kuzmina, submitted). In this section, the differences between persons with AD and controls will be discussed in relation to language assessment during diagnosis. Furthermore, study IV looks at how language impairments change over time for a person with PPA (Lind et al., 2018).

6.3.1 Assessment and diagnosis

There is often little importance attached to language assessment during screening for dementia, despite the fact that most types of dementia affect language behavior and use. Furthermore, most neuropsychological tests used in dementia assessment rely heavily on verbal information. Insufficient language testing can therefore also have implications for general test performance. If patients have subtle language impairments which are unnoticed due to poor assessment, it may affect their ability to perform well in a demanding testing session.

Traditional language assessment in dementia is often focused on production, and conducted by means of picture naming tests. Language comprehension is often not assessed. By combining tests for production and comprehension, researchers can investigate if a person has an impairment in semantic knowledge, or if there is an impairment in access to the phonological form.

The results from the literature review show that both AD and PPA can be distinguished from healthy aging based on language data. This is especially prominent in tests for lexical retrieval, such as verbal fluency and confrontation naming tasks (Ribu, Under revision). The same pattern was also found in the verbal fluency and confrontation naming tests in Ribu et al. (submitted), where healthy older adults scored better than participants with AD and PPA. There were also differences between persons with AD and controls on the FWA task, which cannot be found by employing (only) traditional, restrictive tasks such as picture naming (Ribu et al., submitted). This indicates that variations in language performance can help differentiate between healthy and pathological aging.

In study IV (Lind et al., 2018), JJ's naming of single words was contrasted with his word-searches in conversation and on semi-spontaneous speech samples. This comparison shows that JJ's poor sentence construction, especially in Norwegian, can be related to his poor verb naming skills on the confrontation naming task. Verbs play an important role in sentence generation, and hence also communication (de Diego Balaguer et al., 2006). This means that it is important to assess verbs as well as nouns to get a comprehensive picture of a person's full language impairment.

The difference between neurologically healthy participants and persons with AD and PPA was less pronounced in the sentence comprehension experiment in study III, especially for offline accuracy judgment. The results from the eye tracking data, on the other hand, show a temporal difference between persons with AD and healthy controls; the participants with AD and PPA are slower in detecting the target picture than the neurologically healthy controls. This indicates that comprehension is slower, especially on low-frequency sentence structures for persons with dementia. This is important to keep in mind when communicating with persons with dementia, or during dementia assessment, since many neuropsychological tests often have complex instructions, which may be harder to follow if not comprehended in their entirety.

Most of the participants with AD, the two with PPA, as well as one of the healthy controls make a few mistakes on the object cleft sentences, but not enough to state that these structures are impossible to comprehend. Some of the mistakes can be attributed to poorly drawn stimuli, for instance in some drawings it can be difficult to clearly see the difference between 'the man' and 'the woman' (see examples in Appendix F). Since the persons with AD and PPA perform relatively similarly to the neurologically healthy participants, this may reflect the assumption noted in the introduction to this dissertation, that language impairments in AD are exaggerated manifestations of the changes observed in healthy aging.

The above-mentioned observations suggest that a thorough assessment of language and language abilities can add to the clinical picture of both AD and PPA, and that linguistic tests should supplement other neurocognitive tests during the diagnostic process more than is the case in current assessment protocols. Results from the review paper also suggest that the patterns of language impairment in AD and PPA are quite similar. However, since only a few of the studies account for which subtype of PPA they studied, it is difficult to say if the different subtypes can be distinguished from AD based on analyses for psycholinguistic variables alone. As an alternative, an FWA task may be a valuable additional test to use for detecting language impairment in dementia.

6.3.2 Longitudinal change

Given the progressive nature of PPA, it is not surprising that there is a gradual decline in JJ's naming abilities over time (Lind et al., 2018). The decline is found in all three measures: confrontation naming, semi-spontaneous picture description and in conversation. The progressive decline in communicative skills were also documented by his wife's reports on the Communicative Effectiveness, CETI Index (Lomas et al., 1989, 2006).

JJ experienced different levels of naming difficulties both between his two languages and between the different tests at the first test point; however, these differences diminished over time. Two years post diagnosis, there was almost no difference between his scores on the picture naming tests in English (L1) and Norwegian (L2). The difference in decline was smallest in the cartoon description task, where his performance was quite poor in both languages already at test point 1. Even though it may seem like JJ's performance in his L1 is better spared than his L2 initially, there were indications in his narrative production and in the dissociation between naming performance on verbs and nouns that he experienced parallel impairments in both languages. JJ's dementia affects his less dominant language faster, but over time both languages became almost equally impaired.

English and Norwegian are closely related languages, and the parallel decline in the two languages may have been a result of deterioration of similar domains across lexically similar languages (e.g., the domains of lexical retrieval and comprehension). However, JJ's data follow a similar pattern found also for a person albeit with non-fluent, agrammatic PPA, who spoke Hungarian and English, two very different languages (Druks & Weekes, 2013).

These results indicate a slow and steady decline in several language functions, and both or all languages over time. However, some studies show that persons with PPA may have benefit from speech-language therapy (Croot et al., 2015; Meyer, Snider, Eckmann, & Friedman, 2015; Beales, Whitworth, & Cartwright, 2018), especially lexical retrieval intervention. Speech-language therapy is currently not a standard offer for people with PPA (or AD) in Norway.

6.3.3 How many types of lvPPA are there?

Since the identification of the three subtypes of PPA was published (Gorno-Tempini et al., 2011), several accounts of unclassifiable types of PPA have been identified. For instance, Sajjadi et al. (2012) found that as many as 41,3% of PPA cases were impossible to assign to any of the three established subtypes. The logopenic variant of PPA has proven especially difficult to correctly diagnose (Machulda et al., 2013; Sajjadi et al., 2012; Vandenberghe, 2016; Marshall et al., 2018).

In two important papers, researchers argue for a split of lvPPA into two different sub-classes. Vandenberghe (2016) proposed that lvPPA can be further divided into a left-hemisphere dominant lvPPA, which is very closely linked to non-amnesic AD with initial language impairment manifestations; and a temporoparietal transition zone lvPPA, with phonological working memory deficits, which is more closely linked to the original diagnostic criteria in Gorno-Tempini et al. (2011). The second type is restricted to phonological working memory impairments, and has a lower likelihood of AD compared to the first type (Vandenberghe, 2016).

Matias-Guiu et al. (2019) also argue for split of lvPPA into two subtypes, which they call Type 1 and Type 2. In their classifications, persons with lvPPA Type 1 performed more poorly on action naming tests than those with lvPPA Type 2. While lvPPA Type 1 is associated with atrophy in the left frontal lobe, Type 2 involves a more posterior region and the right parietotemporal lobe, similar to the regions identified by Vandenberghe (2016). Matias-Guiu et al. (2019) also fond clear gender differences betwen the two types; almost all patients with lvPPA Type 1 were female, while all participants that were classified with lvPPA Type 2 were male (Matias-Guiu et al., 2019).

Based on my own observations from the data collection sessions, I find that PPA01 (papers

II and III) and JJ's (paper IV) language profiles are more similar to each other than either of them are to PPA02 (papers II and III). Unfortunately, the three participants with lvPPA were not assessed with the same protocol, which can make it difficult to compare directly, yet there are some similarities which should be noted.

JJ and PPA01 have the same slow, halting and effortful speech pattern, whereas PPA02's speech is more fluent. PPA02 is equally impaired on action and object naming tests, while JJ and PPA01 perform better on object naming than on action naming. Furthermore, both JJ and PPA01 were initially diagnosed with Alzheimer's disease, whereas PPA02's diagnosis was a more straightforward example of lvPPA. This fits with the view where there is a subtype of PPA which is more associated with AD, and another subtype which is more clearly linked to the original diagnostic criteria proposed by Gorno-Tempini et al. (2011). Of course, it is not possible to state anything with certainty, based on only three participants.

It is further difficult to compare JJ, PPA01 and PPA02's results on tests for cognitive function, as not all the same tests were employed in the different studies. Moreover, both PPA01 and PPA02 were unable to perform some of the tests included in the cognitive test battery (PPA01 did not perform the figure copying test from CERAD,[2] and PPA02 did not perform the Stroop test due to poor color vision). Unfortunately, I do not have access to information about the neural atrophy for the participants with AD and PPA in this study. This means that I cannot conclude that the divergent language patterns are associated with distinct profiles of neural atrophy.

When PPA01 and PPA02 are compared to each other in studies II and III, it is clear that their profiles are quite different, and it is difficult to say if this is only due to individual differences. However, when JJ's scores on the comparable tests are also taken into account, it may look like we find two different patterns. Future studies should include a larger sample of persons with lvPPA to further investigate these patterns.

6.4 Limitations

There are certain methodological limitations related to this project which should be addressed. Specific limitations for each study are outlined in chapter 5 and in each of the four papers.

The main general limitation in this project is the small number of participants, both in terms of participants with dementia and control participants. The number of participants with dementia is especially low, and it is problematic to refer to them as a group. Therefore, in the papers Ribu et al. (submitted) and Ribu and Kuzmina (submitted) their results are mainly discussed on an individual level and compared to the control participants. There is also great variability between the participants, which can make it difficult to make judgments about the validity of the findings.

The low number of control participants is also problematic, especially in study III, where

[2]The test results were unavailable among the results obtained from the national register, and the research nurse who accessed the data could not provide a reason for this.

poor eye tracking measurements from several of the young participants had consequences for the whole group. The poor measurements were due to both technical difficulties with the apparatus and possible confounding factors, such as use of heavy eye make up, reflections from contact lenses and glasses etc., but also because some expressed little interest in the task and several did not want to complete all three test sets.

Small sample sizes are common in neurolinguistic and clinical linguistic research, due to the difficulties with recruiting a representative population. Sampling is often following a so-called *convenience sampling* principle, where participants who are available are included. *Purposive sampling* is another common way of obtaining participants, and it refers to how participants with certain predefined characteristics (e.g., persons with AD and PPA in the current experiment) are approached (Mildner, 2013). The number of participants is often limited also due to the high organizational demands of testing, since only one participant can be tested at a time, and testing sessions are often stretched out over several hours (Mertins, 2016).

A second general limitation is concerned with the use of non-standardized tests in the project. In studies II (Ribu et al., submitted) and III (Ribu & Kuzmina, submitted), tests for lexical retrieval were specifically designed (the FWA test and the Picture Naming test), and in studies III and IV (Lind et al., 2018), different adaptations to standardized tests were implemented. This was done because there is a lack of specifically developed material to test for language impairment in dementia in Norway. As briefly mentioned in 3.4.1, the translations of both the Boston Naming Test (Kaplan et al., 1983) and the word list from the Consortium to Establish a Register for Alzheimer's Disease (Morris et al., 1989; Kirsebom et al., 2019) may be insufficient to use in Norwegian in their current form, as they are uncritically translated and not adapted to Norwegian. One may even ask if they are in fact comparable to the original English versions. The Picture Naming test and the word-to-picture matching test that were specifically developed for this project need to be trialled further, both for persons with different dementia diagnoses and neurologically healthy participants, to establish the validity of these tests (see section 4.6).

6.5 Further research

The data collected during the test sessions includes a lot of material which has not yet been published, and which can be used to further study language in aging and dementia. In this section, I outline three possible research directions, based on the material which is still available from this project.

1. Study of connected speech samples

2. Typicality of words produced in association and verbal fluency tasks

3. Identification of lvPPA subtypes

To begin with, within the scope of this thesis, it was not possible to transcribe and analyze the cartoon description task and spontaneous speech samples collected for studies II and III. Building on the results from JJ in study IV, there might be interesting issues regarding idea density, number of words and kinds of words used in story retelling that might help distinguish between dementia and healthy aging. Furthermore, spontaneous speech data can also be used to investigate pauses, fluency and hesitations — all of which seem to be affected even in early stages of the different diseases (Mack et al., 2015). Samples of connected speech can also be used to study the relationship between word searches in discourse and in confrontation naming (Lind et al., 2018).

Next, the data from the word association test was, for the purpose of this study included here analyzed per cue word, rather than per participant. Some research suggests that persons with AD provide less typical responses than neurologically healthy participants (Eustache et al., 1990; Gollan, Salmon, & Paxton, 2006). Bøyum (2016) created norms lists from the participants in her study, giving a good account for which words can be expected as response words to a particular cue. For instance, more than 20% of the participants over 60 years old in her study responded 'factory' to the cue 'industry'. Among the participants with AD and PPA in this study not one provided that same response word. In fact, only one person responded with a word which is also found in the norms list from the 100-word test, namely 'virksomhet ('enterprise'), which might also have been a perseveration of a cue word prompted just before the cue 'industry' (see Appendix E). A possible next step with this test can be to look at each participant's responses in relation to the norms lists and investigate the typicality of responses, and also to look at the number of perseverations the participants with AD and PPA produce, which may be an indication of reduced working memory.

Lastly, based on the current trends in PPA research and in light of the divergent profiles found among the three participants with lvPPA included in this project, a natural next step would be to study these patterns in more detail. Future studies should include more participants with lvPPA, and also follow these participants over time to see if the different patterns are only visible initially and will converge over time, or if there are two different subtypes with a different underlying pathology and different progress throughout.

Since all three participants with PPA reported on in the studies in this dissertation are male, it is impossible to corroborate the gender difference found in Matias-Guiu et al. (2019) with data from my own research. Further investigation of the gender differences in lvPPA would be interesting, and more research on larger population samples is needed.

6.6 Conclusion

The findings from the studies included in this dissertation can be used for a discussion of both theoretical and clinical issues related to language impairment in dementia, and methodological issues regarding assessment and research.

The results from this dissertation are compatible with usage-based theories of language, and with interactive models of language processing. Language changes throughout the life span due to experiences, frequency and usage patterns. This can, for instance, be exemplified by the findings that older adults have larger vocabularies than younger adults, which leads to a higher number of semantically related, and especially synonymous, responses on the Picture Naming and word association tests.

Of the models of language and aging, the transmission deficit hypothesis (TDH) (Burke & Shafto, 2004) receives a great deal of support. This hypothesis suggests that the semantic knowledge is still intact, even if access to the phonological form is impaired. This is in line with the findings that older adults, and persons with dementia provide semantically related responses in lexical retrieval tasks.

Furthermore, responses in the FWA test show that there is interaction between phonological and semantic levels in production, fitting with interactive models of language processing, and not to serial models. Multilink (Dijkstra et al., 2019b) is an interesting model in this respect, since it also assumes a multilingual mental lexicon as the default, and there is indication of bilingual activation on a range of tests, even from the "monolingual" participants. This means that persons who think they are monolingual are in fact bilinguals, at least this is true for most Norwegians.

In this project, I have also shown the importance of employing different tools and methods to study the same topic. Methodological triangulation gives a more comprehensive account of the studied phenomena. For instance, without the combination of online and offline measures of sentence comprehension in study III, the difference between participants with and without dementia would not have been visible to the same degree.

There are subtle differences between persons with and without dementia on several tasks, which may be of clinical relevance. There is not one specific linguistic test which can distinguish healthy aging from dementia, or AD from PPA, but some linguistic markers can be very beneficial as additional clinical cues, and should not be neglected. For instance, the results from the FWA test show that this test can supplement traditional tests for lexical access to give a more nuanced picture of the differences between healthy aging and AD/PPA.

It can be difficult to distinguish between AD and lvPPA, and the differential diagnosis is further complicated by the possibility of two different subtypes as lvPPA. In some cases, persons with lvPPA are at risk of being misdiagnosed with AD, which might need different treatment and management. The studies included in this dissertation lend support to the hypothesis that there might be two types of lvPPA, and signify the importance of more research on language impairment in dementia.

The changes in language in dementia are similar to the changes observed in healthy aging, albeit more pronounced. Older adults experience difficulties on confrontation naming tests, and to some degree also with sentence comprehension. These difficulties are even more evident for persons with dementia. This means that language impairment in later life can be seen as a continuum, from "non-pathological" aging at one end, to decline associated with cognitive impairment at the other (De Bot & Makoni, 2005). It is important to be aware of these changes when evaluating persons who seek help for cognitive complaints.

With this dissertation, I have tried to find ways of learning more about language impairment in AD and PPA, and shown how studying linguistic data can help shed light on the clinical manifestations of the two diseases. The studies show that there are linguistic markers — such as different response patterns on the FWA test — which can distinguish persons with AD and PPA from their neurologically healthy peers. This implies that language screening during the diagnostic process should not be taken lightly. Furthermore, there is a need for proper diagnostic tools which can capture the fine-grained differences which are associated with dementia.

References

Adlam, A. L., Patterson, K., Bozeat, S., & Hodges, J. R. (2010). The Cambridge Semantic Memory Test Battery: Detection of semantic deficits in semantic dementia and Alzheimer's disease. *Neurocase*, 193–207. doi:10.1080/13554790903405693

Almor, A., Kempler, D., MacDonald, M. C., Andersen, E. S., & Tyler, L. K. (1999). Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease. *Brain and Language*, *67*(3), 202–227.

Arbuckle, T. Y. & Gold, D. P. (1993). Aging, inhibition, and verbosity. *Journal of Gerontology*, *48*(5), 225–232.

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

Au, R., Albert, M. L., & Obler, L. K. (1989). Language in normal aging: Linguistic and neuropsychological factors. *Journal of Neurolinguistics*, *4*(3–4), 347–364. doi:10.1016/0911-6044(89)90026-2

Au, R., Joung, P., Nicholas, M., Obler, L. K., Kass, R., & Albert, M. L. (1995). Naming ability across the adult life span. *Aging, Neuropsychology, and Cognition*, *2*(4), 300–311.

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The Word-Frequency mirror effect in young, old, and early-stage Alzheimer's Disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, *46*(1), 199–226. doi:10.1006/jmla.2001.2803

Balota, D. A. & Chumbley, J. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357.

Bastiaanse, R., Edwards, S., Mass, E., & Rispens, J. (2003). Assessing comprehension and production of verbs and sentences: The verb and sentence test (vast). *Aphasiology*, *17*(1), 49–73.

Bastiaanse, R., Lind, M., Moen, I., & Simonsen, H. G. (2006). *Verb- og Setningstesten*. Novus.

Bastiaanse, R. & van Zonneveld, R. (2006). Comprehension of passives in broca's aphasia. *Brain and language*, *96*(2), 135–142.

Bayles, K. A., Kaszniak, A. W., & Tomoeda, C. K. (1987). *Communication and cognition in normal aging and dementia*. Google-Books-ID: RnZxz38BORgC. Little, Brown.

Bayles, K. A. & Tomoeda, C. K. (1991). Caregiver report of prevalence and appearance order of linguistic symptoms in Alzheimer's patients. *The Gerontologist*.

Bayles, K. A. & Tomoeda, C. K. (2007). *Cognitive-Communication Disorders of Dementia* (1 edition). San Diego: Plural Pulishing Inc.

Beales, A., Whitworth, A., & Cartwright, J. (2018). A review of lexical retrieval intervention in primary progressive aphasia and alzheimer's disease: Mechanisms of change, generalisation, and cognition. *Aphasiology*, *32*(11), 1360–1387.

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*(2), 459–464.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240–250. doi:10.1016/j.tics.2012.03.001

Bickel, C., Pantel, J., Eysenbach, K., & Schröder, J. (2000). Syntactic comprehension deficits in Alzheimer's Disease. *Brain and Language*, *71*(3), 432–448. doi:10.1006/brln.1999.2277

Bird, H., Howard, D., & Franklin, S. (2000). Why is a verb like an inanimate object? Grammatical category and semantic category deficits. *Brain and Language*, *72*(3), 246–309. doi:10.1006/brln.2000.2292

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: The importance of being imageable. *Objects vs. actions and nouns vs. verbs*, *16*(2–3), 113–149. doi:10.1016/S0911-6044(02)00016-7

Boone, K. B., Miller, B. L., Lesser, I. M., Hill, E., & D'Elia, L. (1990). Performance on frontal lobe tests in healthy, older individuals. *Developmental Neuropsychology*, *6*(3), 215–223.

Boye, K. & Harder, P. (2017). Konstruktionsgrammatik, regler og funktionsbaseret struktur i neurokognitiv belysning. *Ny forskning i grammatik*, *25*(24). doi:10.7146/nfg.v25i24.97244

Bøyum, M. (2016). *A New Word Association Test for Norwegian: Implications for theories on the mental lexicon, and on language and ageing* (Master's thesis, University of Oslo). Retrieved from https://www.duo.uio.no/handle/10852/52231

Bøyum, M., Hansen, P., & Ribu, I. S. (forthcoming). Assessing lexical processing with a free word association test: The roles of age, language proficiency and cognitive decline.

Braak, H., Braak, E., & Hohl, J. (1993). Staging of alzheimer-related cortical destruction. *European Neurology*.

Brysbaert, M. & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual cognition*, *13*(7-8), 992–1011.

Burke, D. M. & Laver, G. D. (1990). Aging and word retrieval: Selective age deficits in language. In Eugene A. Lovelace (Ed.), *Advances in Psychology* (Vol. Volume 72, pp. 281–300). North-Holland.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the Tip of the Tongue: What Causes Word Finding Failures in young and Older Adults. *Journal of Memory and Language*.

Burke, D. M. & Shafto, M. A. (2004). Aging and Language Production. *Current Directions in Psychological Science*, *13*(1), 21–24. doi:10.1111/j.0963-7214.2004.01301006.x

Burke, D. M. & Shafto, M. A. (2008). Language and aging. In *The handbook of aging and cognition, 3rd ed* (pp. 373–443). New York, NY, US: Psychology Press.

Bybee, J. (2001). *Phonology and language use*. Cambridge University Press.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711–733.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Caplan, D. (1996). Syntactic Processing in Sentence Comprehension Under Dual-task Conditions in Aphasic Patients. *Language and Cognitive Processes*, *11*(5), 525–551. doi:10.1080/016909696387141

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and cognition*, *5*(1), 41–66.

Chang, Y.-N., Monaghan, P., & Welbourne, S. (2019). A computational model of reading across development: Effects of literacy onset on language processing. *Journal of Memory and Language*.

Charles, D., Olm, C., Powers, J., Ash, S., Irwin, D. J., McMillan, C. T., … Grossman, M. (2014). Grammatical comprehension deficits in non-fluent/agrammatic primary progressive aphasia. *J Neurol Neurosurg Psychiatry*, *85*(3), 249–256.

Connor, L. T., Spiro, A. I., & Obler, L. K. (2004). Change in object naming ability during adulthood. *The Journals of Gerontology*.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*.

Creswell, J. W. (2009). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications.

Croot, K., Hodges, J. R., & Patterson, K. (1999). Evidence for impaired sentence comprehension in early Alzheimer's disease. *Journal of the International Neuropsychological Society: JINS*, *5*(5), 393–404.

Croot, K., Taylor, C., Abel, S., Jones, K., Krein, L., Hameister, I., … Nickels, L. (2015). Measuring gains in connected speech following treatment for word retrieval: A study with two participants with primary progressive aphasia. *Aphasiology*, *29*(11), 1265–1288.

Crowe, S. F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the trail making test. *journal of clinical psychology*.

Cuetos, F., Arce, N., Martínez, C., & Ellis, A. W. (2017). Word recognition in Alzheimer's disease: Effects of semantic degeneration. *Journal of Neuropsychology*, *11*(1), 26–39. doi:10.1111/jnp.12077

Cuetos, F., Gonzalez-Nosti, M., & Martínez, C. (2005). The picture-naming task in the analysis of cognitive deterioration in Alzheimer's disease. *Aphasiology*, *19*(6), 545–557. doi:10.1080/02687030544000010

Cuetos, F., Herrera, E., & Ellis, A. W. (2010). Impaired word recognition in Alzheimer's disease: The role of age of acquisition. *Neuropsychologia*, *48*(11), 3329–3334. doi:10.1016/j.neuropsychologia.2010.07.017

Cuetos, F., Rodríguez-Ferreiro, J., & Menéndez, M. (2009). Semantic Markers in the Diagnosis of Neurodegenerative Dementias. *Dementia and Geriatric Cognitive Disorders*, *28*(3), 267–274.

Cuetos, F., Rodríguez-Ferreiro, J., Sage, K., & Ellis, A. W. (2012). A fresh look at the predictors of naming accuracy and errors in Alzheimer's disease. *Journal of Neuropsychology*, *6*(2), 242–256. doi:10.1111/j.1748-6653.2011.02025.x

Dąbrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*.

Damico, J. S., Müller, N., & Ball, M. J. (2010). *The Handbook of Language and Speech Disorders*. Google-Books-ID: zQcdTCy0tm0C. John Wiley & Sons.

Davies, M. (2008). The corpus of contemporary american english: 450 million words, 1990-present.

De Bot, K. & Makoni, S. (2005). *Language and Aging in Multilingual Contexts*. MultiLingual Matters.

de Diego Balaguer, R., Rodrıguez-Fornells, A., Rotte, M., Bahlmann, J., Heinze, H.-J., & Münte, T. F. (2006). Neural circuits subserving the retrieval of stems and grammatical features in regular and irregular verbs. *Human Brain Mapping*, *27*(11), 874–888.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis Kaplan Executive Function System: Examiner's manual*. San Antonio (Tex.): Pearson.

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*.

Dick, F., Bates, E., Wilfeck, B., Utman, J. A., Dronkers, N., & Gernsbacker, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological review*.

Diessel, H. (2017). Usage-based linguistics. In *Oxford research encyclopedia of linguistics*.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019a). Modelling bilingual lexical processing: A research agenda and desiderabilia. *Bilingualism: Language and Cognition*, *22*(4), 703–713.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019b). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679.

Domoto-Reilly, K., Sapolsky, D., Brickhouse, M., & Dickerson, B. C. (2012). Naming impairment in Alzheimer's disease is associated with left anterior temporal lobe atrophy. *NeuroImage*, *63*(1), 348–355. doi:10.1016/j.neuroimage.2012.06.018

Druks, J., Masterson, J., Kopelman, M., Clare, L., Rose, A., & Rai, G. (2006). Is action naming better preserved (than object naming) in Alzheimer's disease and why should we ask? *Brain and Language*, *98*(3), 332–340. doi:10.1016/j.bandl.2006.06.003

Druks, J. & Weekes, B. S. (2013). Parallel deterioration to language processing in a bilingual speaker. *Cognitive Neuropsychology*, *30*(7-8), 578–596. doi:10.1080/02643294.2014.882814

Dyvik, H., Meurer, P., Rosén, V., Smedt, K. D., Haugereid, P., Losnegaard, G. S., ... Thunes, M. (2016). Norgrambank: A 'deep' treebank for norwegian. In N. C. (Chair), K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, ... S. Piperidis (Eds.), *Proceedings of the tenth international conference on language resources and evaluation (lrec 2016)*. Portorož, Slovenia: European Language Resources Association (ELRA).

Egeland, J., Landrø, N. I., Tjemsland, E., & Walbækken, K. (2006). Norwegian norms and factor-structure of phonemic and semantic word list generation. *The Clinical Neuropsychologist*, *20*(4), 716–728.

Ellis, A. W. & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(5), 1103–1123.

Ellis, A. W. & Young, H. D. (1977). Age-of-acquisition and recognition of nouns presented in the left and right visual fields: A failed hypothesis. *Neuropsychologia*, *15*(6), 825–828. doi:10.1016/0028-3932(77)90015-X

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(02). doi:10.1017/S0272263102002024

Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonserach task. *Perception and psychophysics*.

Eustache, F., Cox, C., Brandt, J., Lechevalier, B., & Pons, L. (1990). Word-association responses and severity of dementia in Alzheimer disease. *Psychological Reports*, *66*(3 Pt 2), 1315–1322. doi:10.2466/pr0.1990.66.3c.1315

Farkas, T., Ferris, S. H., Wolf, A. P., De Leon, M. J., CHristman, D. R., Reisberg, B., ... Reivich, M. (1982). 18F-2-deoxy-2-fluoro-D-glucose as a tracer in the positron emission tomographic study of senile dementia. *The American Journal of Psychiatry*.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164–203. doi:10.1016/S0010-0285(03)00005-7

Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. J. (2015). Establishing the Reliability of Word Association Data for Investigating Individual and Group Differences. *Applied Linguistics*, *36*(1), 23–50. doi:10.1093/applin/amt020

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.

Frank, S. L. & Christiansen, M. H. [Morten H]. (2018). Hierarchical and sequential processing of language: A response to: Ding, melloni, tian, and poeppel (2017). rule-based and word-level statistics-based processing of language: Insights from neuroscience. language, cognition and neuroscience. *Language, Cognition and Neuroscience*, *33*(9), 1213–1218.

Frank, S. L. & Yang, J. (2018). Lexical representation explains cortical entrainment during speech comprehension. *PloS one*, *13*(5), e0197304.

Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., & Rochon, E. (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Language, Computers and Cognitive Neuroscience*, *55*, 43–60. doi:10.1016/j.cortex.2012.12.006

Fraser, K. C., Meltzer, J. A., & Rudzicz, F. (2015). Linguistic features identify Alzheimer's disese in narrative speech. *Journal of Alzheimer*. doi:10.3233/JAD-150520

Gahl, S. & Menn, L. (2016). Usage-based approaches to aphasia. *Aphasiology*, *30*(11), 1361–1377. doi:10.1080/02687038.2016.1140120

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Gollan, T. H., Salmon, D. P., Montoya, R. I., & da Pena, E. (2010). Accessibility of the nondominant language in picture naming: A counterintuitive effect of dementia on bilingual language production. *Neuropsychologia*, *48*(5), 1356–1366. doi:10.1016/j.neuropsychologia.2009.12.038

Gollan, T. H., Salmon, D. P., & Paxton, J. L. (2006). Word association in early Alzheimer's disease. *Brain and Language*. Language Comprehension across the Life Span, *99*(3), 289–303. doi:10.1016/j.bandl.2005.07.001

Gómez-Ruiz, I., Aguilar-Alonso, Á., & Espasa, M. A. (2012). Language impairment in catalan-spanish bilinguals with alzheimer's disease. *Journal of Neurolinguistics*, *25*(6), 552–566.

Goral, M. (2004). First-language decline in healthy aging: Implications for attrition in bilingualism. *Journal of Neurolinguistics*. Attrition, *17*(1), 31–52. doi:10.1016/S0911-6044(03)00052-6

Goral, M. (2013). Bilingualism, language and aging. In J. Altarriba & L. Isurin (Eds.), *Memory, Language and Bilingualism*. Cambridge University Press.

Goral, M. (2019). The bilingual mental lexicon beyond dutch-english written words. *Bilingualism: Language and cognition*.

Goral, M., Clark-Cotton, M. R., Spiro, A. I., Obler, L. K., Verkuilen, J., & Albert, M. L. (2011). The contribution of set shifting and working memory to sentence processing in older adults. *Experimental aging research*.

Goral, M., Spiro, A. I., Martin, L. A., Obler, L. K., & Connor, L. T. (2007). Change in lexical retrieval skills in adulthood. *The Mental Lexicon*.

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., … Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014.

Gorno-Tempini, M. L. & Miller, B. L. (2013). Primary Progressive Aphasia as a model to study the neurobiology of language. *Brain and language*, *127*(2), 105. doi:10.1016/j.bandl.2013.10.008

Grossman, M., Mickanin, J., Onishi, K., Hughes, E., D'Esposito, M., Ding, X.-S., … Reivich, M. (1996). Progressive Nonfluent Aphasia: Language, Cognitive, and PET Measures Contrasted with Probable Alzheimer's Disease. *Journal of Cognitive Neuroscience*, *8*(2), 135–154.

Guevara, E. (2010). NoWaC: A large web-based corpus for Norwegian. In *Proceedings of the NAACL HLT 2010 Sixth Web as Corpus Workshop* (pp. 1–7).

Hansen, P. (2016). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language*, 0142723716679956.

Happ, M., Bathke, A. C., & Brunner, E. (2019). Optimal sample size planning for the wilcoxon-mann-whitney test. *Statistics in medicine*, *38*(3), 363–375.

Hardy, C. J. D., Johnson, J. C. S., & Warren, J. D. (2019). Auditory symptoms in primary progressive aphasia: A commentary on Utiansky et al. (2019). *Cortex*.

Hasher, L. & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In *Psychology of learning and motivation* (Vol. 22, pp. 193–225). Elsevier.

Henry, M. L. & Gorno-Tempini, M. L. (2010). The logopenic variant of primary progressive aphasia. *Current opinion in neurology*.

Hernández, M., Costa, A., Sebastián-Gallés, N., Juncadella, M., & Reñé, R. (2007). The organisation of nouns and verbs in bilingual speakers: A case of bilingual grammatical category-specific deficit. *Journal of Neurolinguistics*, *20*(4), 285–305. doi:10.1016/j.jneuroling.2006.10.002

Hillis, A. E., Heidler-Gary, J., Newhart, M., Chang, S., Ken, L., & Bak, T. (2006). Naming and comprehension in primary progressive aphasia: The influence of grammatical word class. *Aphasiology*, *20*(2-4), 246–256.

Hillis, A. E., Oh, S., & Ken, L. (2004). Deterioration of naming nouns versus verbs in primary progressive aphasia. *Annals of Neurology*, *55*(2), 268–275. doi:10.1002/ana.10812

Hillis, A. E., Tuffiash, E., & Caramazza, A. (2002). Modality-Specific Deterioration in Naming Verbs in Nonfluent Primary Progressive Aphasia. *Journal of Cognitive Neuroscience*, *14*(7), 1099–1108. doi:10.1162/089892902320474544

Hirsh, K. W. & Funnell, E. (1995). Those old, familiar things: Age of acquisition, familiarity and lexical access in progressive aphasia. *Journal of Neurolinguistics*, *9*(1), 23–32. doi:10.1016/0911-6044(95)00003-8

Hodges, J. R. & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia*, *33*(4), 441–459. doi:10.1016/0028-3932(94)00127-B

Hodges, J. R., Patterson, K., Graham, N., & Dawson, K. (1996). Naming and knowing in dementia of Alzheimer's type. *Brain and Language*, *54*(2), 302–325.

Hoffman, P., Sajjadi, S. A., Patterson, K., & Nestor, P. J. (2017). Data-driven classification of patients with primary progressive aphasia. *Brain and Language*, *174*, 86–93. doi:10.1016/j.bandl.2017.08.001

Hoffmann, I., Németh, D., Dye, C., Pákáski, M., Irinyi, T., & Kálmán, J. (2010). Temporal features of spontaneous speech in Alzheimer's disease. *International Journal of Speech and Language*.

Howard, D. & Patterson, K. (1992). *The pyramids and palm trees test*. Thames Valley Test Company.

Howard, D. & Patterson, K. (2005). *Pyramide- og palmetesten*. Novus forlag.

Howard, D., Patterson, K., Franklin, S., Morton, J., & Orchard-Lisle, V. (1984). Variability and consistency in picture naming by aphasic patients. *Advances in neurology*, *42*, 263.

Ishkhanyan, B., Sahraoui, H., Harder, P., Mogensen, J., & Boye, K. (2017). Grammatical and lexical pronoun dissociation in French speakers with agrammatic aphasia: A usage-based account and REF-based hypothesis. *Journal of Neurolinguistics*.

Ivanova, I. & Kleinman, D. (2019). Multilink for bilingual language production. *Bilingualism: Language and Cognition*, *22*(4), 687–688.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological tests' norms above age 55: Cowat, bnt, mae token, wrat-r reading, amnart, stroop, tmt, and jlo. *The Clinical Neuropsychologist*, *10*(3), 262–278.

Jap, B. A., Martinez-Ferreiro, S., & Bastiaanse, R. (2016). The effect of syntactic frequency on sentence comprehension in standard indonesian broca's aphasia. *Aphasiology*, *30*(11), 1325–1340.

Jarema, G. & Libben, G. (2007). *The Mental Lexicon: Core perspectives*. Elsevier.

Juhasz, B. J. (2005). Age-of-Acquisition Effects in Word and Picture Identification. *Psychological Bulletin*, *131*(5), 684–712.

Juncos-Rabadán, O., Pereiro, A. X., & Rodríguez, M. S. (2005). Narrative speech in aging: Quantity, information content and cohesion. *Brain and Language*.

Kambanaros, M. & Grohmann, K. K. (2015). Grammatical Class Effects Across Impaired Child and Adult Populations. *Frontiers in Psychology*, *6*, 1670. doi:10.3389/fpsyg.2015.01670

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Kavé, G. & Goral, M. (2016a). Do age-related word retrieval difficulties appear (or disappear) in connected speech? *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, 1–20. doi:10.1080/13825585.2016.1226249

Kavé, G. & Goral, M. (2016b). Word retrieval in connected speech in Alzheimer's disease: A review with meta-analyses: Aphasiology: Vol 32, No 1. *Aphasiology*. Retrieved from https://www.tandfonline.com/doi/full/10.1080/02687038.2017.1338663

Kay, J., Lesser, R., & Coltheart, M. (1996). *Psycholinguistic Assessment of Language Processing in Aphasia*. Psychology Press.

Kay, J., Lesser, R., & Coltheart, M. (2009). *Psykolingvistisk kartlegging av språkprosessering hos afasirammede – norsk version*. Novus forlag.

Kemper, S. (2012). The interaction if linguistic constraints, working memory, and aging on language production and comprehension. In Naveh-Benjamin & Ohta (Eds.), *Memory and Aging*. CRS Press.

Kempler, D., Almor, A., & MacDonald, M. C. (1998). Teasing Apart the Contribution of Memory and Language Impairments in Alzheimer's DiseaseAn Online Study of Sentence Comprehension. *American Journal of Speech-Language Pathology*, *7*(1), 61–67. doi:10.1044/1058-0360.0701.61

Kempler, D., Almor, A., Tyler, L. K., Andersen, E. S., & MacDonald, M. C. (1998). Sentence Comprehension Deficits in Alzheimer's Disease: A Comparison of Off-Line vs. On-Line Sentence Processing. *Brain and Language*, *64*(3), 297–316. doi:10.1006/brln.1998.1980

Kim, M. & Thompson, C. K. (2004). Verb deficits in Alzheimer's disease and agrammatism - Implications for lexical organization. *Brain and language*, *88*(1), 1–20.

Pearson's Correlation Coefficient. (2008). In W. Kirch (Ed.), *Encyclopedia of public health* (pp. 1090–1091). Dordrecht: Springer Netherlands. doi:10.1007/978-1-4020-5614-7_2569

Kirsebom, B., Espenes, R., Hessen, E., Waterloo, K., Johnsen, S. H., Gundersen, E., ... Fladby, T. (2019). Demographically adjusted cerad wordlist test norms in a norwegian sample from 40 to 80 years. *The Clinical Neuropsychologist*, 1–13.

Knoph, M. (2010). The Bilingual Aphasia Test – På norsk. *Norsk tiddskrift for logopedi*. Retrieved from https://www.mcgill.ca/linguistics/files/linguistics/Norwegian_BAT.pdf

Ko, P. C. & Ally, B. A. (2011). Visual Cognition in Alzheimer's Disease and Its Functional Implications. In S. De La Monte (Ed.), *The clinical spectrum of alzheimer's sisease: The charge toward comprehensive diagnostic and therapeutic strategies* (Chap. Visual Cognition in Alzheimer's Disease and Its Functional Implications). InTech.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical prerequisites. Volume 1*.

Lee, J., Yoshida, M., & Thompson, C. K. (2015). Grammatical Planning Units During Real-Time Sentence Production in Speakers With Agrammatic Aphasia and Healthy Speakers. *Journal of Speech Language and Hearing Science*.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.

Leyton, C. E., Ballard, J., Kirrie, Piguet, O., & Hodges, J. R. (2014). Phonological errors as a clinical marker of the logopenic variant of ppa. *Neurology*.

Leyton, C. E., Savage, S., Irish, M., Schubert, S., Piguet, O., Ballard, K. J., & Hodges, J. R. (2014). Verbal repetition in primary progressive aphasia and in Alzheimer's disease. *Journal of Alzheimer's disease*.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.) (M. D. Lezak, Ed.). Oxford: Oxford University Press.

Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural networks*, *17*(8-9), 1345–1362.

Libben, G. & Goral, M. (2015). How bilingualism shapes the mental lexicon. In J. W. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing*. Cambridge University Press.

Libben, M., Goral, M., & Libben, G. (2017). The dynamic lexicon. In M. Libben, M. Goral, & G. Libben (Eds.). John Benjamins Publishing Company.

Lind, M., Moen, I., & Simonsen, H. G. (2006). Verb- og setningstesten (VOST): Et nytt redskap i den logopediske verktøykassa (del 1). *Logopeden*. Retrieved from https://www.mysciencework.com/publication/show/b13a47ac1ba048b21e7841109b5cd4f8

Lind, M., Simonsen, H. G., Hansen, P., Holm, E., & Mevik, B.-H. (2013). Norwegian Words: A lexical database for clinicians and researchers. *Clinical Linguistics & Phonetics*, 276–290. Retrieved from http://tekstlab.uio.no/ordforradet

Lind, M., Simonsen, H. G., Ribu, I. S., Svendsen, B. A., Svennevig, J., & de Bot, K. (2018). Lexical access in a bilingual speaker with dementia: Changes over time. *Clinical Linguistics & Phonetics*, *32*(4), 353–377. doi:10.1080/02699206.2017.1381168

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). The Communicative Effectiveness Index: Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of speech and hearing disorders*.

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (2006). *Spørreskjema for nærpersoner – CETI norsk versjon*. Bredtvet kompetansesenter.

Louwersheimer, E., Keulen, M. A., Steenwijk, M. D., Wattjes, M. P., Jiskoot, L. C., Vrenken, H., … Pijnenburg, Y. A. L. (2016). Heterogeneous Language Profiles in Patients with Primary Progressive Aphasia due to Alzheimer's Disease. *Journal of Alzheimer's disease: JAD*, *51*(2), 581–590. doi:10.3233/JAD-150812

Łuniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., Anđelković, D., … Ünal-Logacev, Ö. (2016). Ratings of age of acquisition of 299 words across 25 languages: Is there a cross-linguistic order of words? *Behavior Research Methods*, *48*(3), 1154–1177. doi:10.3758/s13428-015-0636-6

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a "new view". *Inhibition in cognition*, *17*, 145–162.

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., & Andersen, E. S. (2001). Assessing Working Memory and Language Comprehension in Alzheimer's Disease. *Brain and Language*, *78*(1), 17–42. doi:10.1006/brln.2000.2436

MacDonald, M. C. & Christiansen, M. H. [Morten H.]. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54, discussion 55–74.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review*, *101*(4), 676.

Machulda, M. M., Whitwell, J. L., Duffy, J. R., Strand, E. A., Dean, P. M., Senjem, M. L., … Josephs, K. A. (2013). Identification of an atypical variant of logopenic progressive aphasia. *Brain and Language*, *127*(2), 139–144. doi:10.1016/j.bandl.2013.02.007

Mack, J. E., Chandler, S. D., Meltzer-Asscher, A., Rogalski, E., Weintraub, S., Mesulam, M.-M., & Thompson, C. K. (2015). What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA? *Neuropsychologia*, *77*, 211–222. doi:10.1016/j.neuropsychologia.2015.08.019

MacKay, D. G. & Burke, D. M. (1990). Cognition and Aging: A Theory of New Learning and the Use of Old Connections. In F.G.A. Stone and Robert West (Ed.), *Advances in Psychology* (Chap. 5, Vol. Volume 71, pp. 213–263). North-Holland. doi:10.1016/S0166-4115(08)60159-4

Malcolm, T., Lerman, A., Korytkowska, M., Vonk, J. M. J., & Obler, L. K. (2019). Primary Progressive Aphasia in Bilinguals and Multilinguals. In J. W. Schwieter & M. Paradis (Eds.). Wiley Blackwell.

Marcotte, K., Graham, N. L., Black, S. E., Tang-Wai, D., Chow, T. W., Freedman, M., ... Leonard, C. (2014). Verb production in the nonfluent and semantic variants of primary progressive aphasia: The influence of lexical and semantic factors. *Cognitive Neuropsychology*, *31*(7-8), 565–583. doi:10.1080/02643294.2014.970154

Marczinski, C. A. & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language*, *97*(3), 258–265. doi:10.1016/j.bandl.2005.11.001

Marshall, C. R., Hardy, C. J., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D., ... Rossor, M. N., et al. (2018). Primary progressive aphasia: A clinical approach. *Journal of neurology*, *265*(6), 1474–1490.

Marslen-Wilson, W. D. (1980). Speech understanding as a psychological process. In *Spoken language generation and understanding* (pp. 39–67). Springer.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102.

Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, *10*(1), 29–63.

Martínez-Ferreiro, S., Bastiaanse, R., & Boye, K. (2019). Functional and usage-based approaches to aphasia: The grammatical-lexical distinction and the role of frequency. *Aphasiology*.

Martínez-Sánchez, F., Meilán, J. J. G., García-Sevilla, J., Carro, J., & Arana, M. J. (2013). Oral reading fluency analysis in patients with Alzheimer's disease and asymptomatic control subjects. *Neurologia (English Edition*.

Matias-Guiu, J. A., Días-Álvarez, J., Cuetos, F., Cabrera-Martín, M. N., Segovia-Ríos, I., Pytel, V., ... Ayala, J. L. (2019). Machine learningin the clinical and language characterisation of primary progressive aphasia variants. *Cortex*.

McClelland, J. L. & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McDermott, R. (2011). Internal and external validity. (pp. 27–40). Cambridge University Press Cambridge.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*.

McKhann, G. M., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*(7), 939–939. doi:10.1212/WNL.34.7.939

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., … Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(3), 263–269. doi:10.1016/j.jalz.2011.03.005

Mertins, B. (2016). The use of experimental methods in linguistic research: Advantages, problems and possible pitfalls. *Slavic Languages in Psycholinguistics*, 15–33.

Mesulam, M.-M. (2001). Primary progressive aphasia. *Annals of Neurology*, *49*(4), 425–432.

Mesulam, M.-M. (2007). Primary progressive aphasia: A 25-year retrospective. *Alzheimer Disease and Associated Disorders*. doi:10.1097/WAD.0b013e31815bf7e1

Meuter, R. (2009). Neurolinguistic contributions to understanding the bilingual mental lexicon. In A. Pavlenko (Ed.). Multilingual Matters.

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, *25*(1), 31–43. doi:10.1016/j.jneuroling.2011.08.001

Meyer, A. M., Snider, S. F. [Sarah F.], Campbell, R. E., & Friedman, R. B. (2015). Phonological short-term memory in logopenic variant Primary Progressive Aphasia and mild Alzheimer's disease. *Cortex*.

Meyer, A. M., Snider, S. F. [Sarah F], Eckmann, C. B., & Friedman, R. B. (2015). Prophylactic treatments for anomia in the logopenic variant of primary progressive aphasia: Cross-language transfer. *Aphasiology*, *29*(9).

Mildner, V. (2013). Experimental and quasi-experimental research in clinical linguistics and phonetics. In N. Müller & M. J. Ball (Eds.).

Molitor, R. J., Ko, P. C., & Ally, B. A. (2015). Eye movements in Alzheimer's disease. *Journal of Alzheimer's disease: JAD*, *44*(1), 1–12. doi:10.3233/JAD-141173

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., … Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165.

Obler, L. K., Albert, M. L., Spiro, A. I., Goral, M., Rykhlevskaia, E., Huyn, J., … Schnyer, D. (2010). Language Changes Associated with Aging. In M. Albert & J. Knoefel (Eds.), *Clinical Neurology of Aging* (Chap. 42). Oxford University Press. Retrieved from http://oxfordmedicine.com/view/10.1093/med/9780195369298.001.0001/med-9780195369298-chapter-042

Obler, L. K. & Gjerlow, K. (1999). *Language and the Brain*. Google-Books-ID: nhc36DgVM4UC. Cambridge University Press.

Obler, L. K. & Pekkala, S. (2008). Language and Communication in Aging. In *Handbook of the Neuroscience of Language* (Chap. 26, pp. 351–358). San Diego: Elsevier. doi:10.1016/B978-008045352-1.00034-3

Oldfield, R. C. & Wingfield, A. [A.]. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281.

Ortega, A., Gómez-Ariza, C. J., Román, P., & Bajo, M. T. (2012). Memory inhibition, aging, and the executive deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(1), 178.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1p2), 1.

Paradis, M. (1998). *The Bilingual Aphasia Test*. McGil University. Retrieved from https://www.mcgill.ca/linguistics/research/bat

Park, D. C. & Reuter-Lorenz, P. (2009). The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology*, *60*, 173–196.

Peelle, J. (2018). Language and aging. In de Zubicaray & Schiller (Eds.), *The Oxford Handbook of Neurolinguistics*. Oxford University Press.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child development*, *78*(4), 1255–1264.

Pereira, M. F., Camargo, M. v. Z. A., Aprahamian, I., & Forlenza, O. V. (2014). Eye movement analysis and cognitive processing: Detecting indicators of conversion to Alzheimer's disease. *Neuropsychiatric Disease and Treatment*, *10*, 1273–1285. doi:10.2147/NDT.S55371

Pistono, A., Pariente, J., Bézy, C., Lemesle, B., Le men, J., & Jucla, M. (2019). What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease. *Neuropsychologia*.

R: A Language and Environment for Statistical Computing. (2019). R Core Team. Vienna, Austria. Retrieved from https://www.R-project.org/

Ribu, I. S. (Under revision). First in, last out: A systematic review of the effects of (psycho-)linguistic variables on lexical retrieval in Alzheimer's disease and primary progressive aphasia.

Ribu, I. S. & Kuzmina, E. (submitted). Tracking non-canonical sentence comprehension in Alzheimer's disease and Primary progressive aphasia.

Ribu, I. S., Norvik, M. I., Lehtonen, M., & Simonsen, H. G. (submitted). Free word association in Alzheimer's disease and Primary progressive aphasia.

Robinson, G., Rossor, M., & Cipolotti, L. (1999). Selective sparing of verb naming in a case of severe Alzheimer's disease. *Cortex*, *35*(3), 443–450.

Robinson, K., Grossman, M., White-Devine, T., & D'Esposito, M. (1996). Category-specific difficulty naming with verbs in Alzheimer's disease. *Neurology*, *47*(1), 178–182.

Rochon, E., Waters, G. S., & Caplan, D. (1994). Sentence Comprehension in Patients with Alzheimer's Disease. *Brain and Language*, *46*(2), 329–349.

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. doi:10.1016/j.jneuroling.2008.05.003

Rohrer, J. D., Caso, F., Mahoney, C., Henry, M., Rosen, H. J., Rabinovici, G., … Gorno-Tempini, M. L. (2013). Patterns of longitudinal brain atrophy in the logopenic variant of primary progressive aphasia. *Brain and Language*, *127*(2), 121–126. doi:10.1016/j.bandl.2012.12.008

Rosén, V., De Smedt, K., Meurer, P., & Dyvik, H. (2012). An open infrastructure for advanced treebanking. In *Meta-research workshop on advanced treebanking at lrec2012* (pp. 22–29). Hajič, Jan.

Rowland, J. (2009). *Rowland Universal Dementia Assessment Scale*.

Sajjadi, S. A., Patterson, K., Arnold, R. J., Watson, P. C., & Nestor, P. J. (2012). Primary progressive aphasia: A tale of two syndromes and the rest. *Neurology*, *78*(21), 1670–1677. doi:10.1212/WNL.0b013e3182574f79

Salthouse, T. A. (2010). Influence of age on practice effects in longitudinal neurocognitive change. *Neuropsychology*.

Santo Pietro, M. J. & Goldfarb, R. (1985). Characteristic patterns of word association responses in institutionalized elderly with and without senile dementia. *Brain and Language*, *26*(2), 230–243.

Schmid, M. S. & De Bot, K. (2004). Language attrition. *The handbook of applied linguistics*, 210.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime 2.0*. Psychology Software Tools Inc. Pittsburgh.

Seckin, M., Mesulam, M.-M., Voss, J. L., Huang, W., Rogalski, E. J., & Hurley, R. S. (2016). Am I looking at a cat or a dog? Gaze in the semantic variant of primary progressive aphasia is subject to excessive taxonomic capture. *Journal of Neurolinguistics*, *37*, 68–81. doi:10.1016/j.jneuroling.2015.09.003

Silveri, M. C., Pravatà, E., Brita, A. C., Improta, E., Ciccarelli, N., Rossi, P., & Colosimo, C. (2014). Primary progressive aphasia: Linguistic patterns and clinical variants. *Brain and Language*, *135*, 57–65. doi:10.1016/j.bandl.2014.05.004

Simonsen, H. G., Lind, M., Hansen, P., Holm, E., & Mevik, B.-H. (2013). Imageability of Norwegian nouns, verbs and adjectives in a cross-linguistic perspective. *Clinical Linguistics & Phonetics*, *27*(6-7), 435–446.

Sloot, P. & Jonkers, R. (2011). Benoemen van acties door Alzheimerpatiënten: Effect van instrumentaliteit en naamsverwantschap. *Stem-, Spraak- en Taalpathologie*, *17*(2). Retrieved from https://sstp.nl/article/view/3277

Souza, A. C. d., Alexandre, N. M. C., & Guirardello, E. d. B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, *26*(3), 649–659.

Stilwell, B. L., Dow, R. M., Lamers, C., & Woods, R. T. (2016). Language changes in bilingual individuals with Alzheimer's disease. *International Journal of Language & Communication Disorders*, *51*(2), 113–127.

Street, J. A. & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua*. Asymmetries in Language Acquisition, *120*(8), 2080–2094. doi:10.1016/j.lingua.2010.01.004

Svennevig, J., Hansen, P., Simonsen, H. G., & Landmark, A. M. D. (2019). Code-switching in multilinguals with dementia: Patterns across speech contexts. *Clinical linguistics & phonetics*, 1–22.

Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J., & Pakaski, M. (2015). Speaking in Alzheimer's Disease, is That an Early Sign? Importance of Changes in Language Abilities in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, *7*. doi:10.3389/fnagi.2015.00195

Taler, V. & Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology*, *30*(5), 501–556.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of psycholinguistic research*, *29*(6), 557–580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science (New York, N.Y.) 268*(5217), 1632–1634.

Taylor, J. R. (2002). *Cognitive Grammar*. Oxford: Oxford University Press.

Teichmann, M., Kas, A., Boutet, C., Ferrieux, S., Nogues, M., Samri, D., … Migliaccio, R. (2013). Deciphering logopenic primary progressive aphasia: A clinical, imaging and biomarker investigation. *Brain, Journal of neurology*.

Thompson, C., Ballard, K., Tait, M., Weintraub, S., & Mesulam, M.-M. (1997). Patterns of language decline in non-fluent primary progressive aphasia. *Aphasiology*, *11*(4-5), 297–321.

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203–214. doi:10.1016/S0887-6177(03)00039-8

Tomoeda, C. K., Bayles, K. A., Boone, D. R., Kaszniak, A. W., & Slauson, T. J. (1990). Speech rate and syntactic complexity effects on the auditory comprehension of alzheimer patients. *Journal of Communication Disorders*, *23*(2), 151–161. doi:10.1016/0021-9924(90)90019-U

Trueswell, J. C. & Tanenhaus, M. K. (1994). Toward a lexicalist framework of constraint-based syntactic ambiguity resolution.

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. doi:10.1016/j.archger.2013.03.011

Tyler, L. K., Cobb, H., & Graham, N. (1992). *Spoken language comprehension: An experimental approach to disordered and normal processing.* The MIT Press.

Tyler, L. K., Russell, R., Fadili, J., & Moss, H. E. (2001). The neural representation of nouns and verbs: PET studies. *Brain: A Journal of Neurology*, *124*(Pt 8), 1619–1634.

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving Syntactic Processing across the Adult Life Span: The Modulation of the Frontotemporal Language System in the Context of Age-Related Atrophy. *Cerebral cortex*.

Ukita, H., Abe, K., & Yamada, J. (1999). Late Acquired Words in Childhood Are Lost Earlier in Primary Progressive Aphasia. *Brain and Language*, *70*(2), 205–219. doi:10.1006/brln.1999.2152

Utianski, R. L., Botha, H., Martin, P. R., Schwartz, C. G., Duffy, J. R., Clark, H. M., … Josephs, K. A. (2019). Clinical and neuroimaging characteristic of clinically unclassifiable Parimary Progressive Aphasia. *Brain and Language*.

Van Hell, J. G. (2019). Words only go so far: Linguistic context affects bilingual word processing. *Bilingualism: Language and Cognition*, *22*(4), 689–690.

Vandenberghe, R. (2016). Classification of the primary progressive aphasias: Principles and review of progress since 2011. *Alzheimer's Research and Therapy*.

Visch-Brink, E., Stronks, D., & Denes, G. (2005). *Semantische Associatie test (SAT)*. Pearson.

Vita, M. G., Marra, C., Spinelli, P., Caprara, A., Scaricamazza, E., Castelli, D., … Quaranta, D. (2014). Typicality of words produced on a semantic fluency task in amnesic mild cognitive impairment: Linguistic analysis and risk of conversion to dementia. *Journal of Alzheimer's disease: JAD*, *42*(4), 1171–1178. doi:10.3233/JAD-140570

Vitevitch, M. S. & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological Science*, *9*(4), 325–329. ArticleType:

research-article / Full publication date: Jul., 1998 / Copyright © 1998 Association for Psychological Science. Retrieved from http://www.jstor.org/stable/40063346

Vonk, J. M., Jonkers, R., Hubbard, H. I., Gorno-Tempini, M. L., Brickman, A. M., & Obler, L. K. (2019). Semantic and lexical features of words dissimilarly affected by non-fluent, logopenic, and semantic primary progressive aphasia. *Journal of the International Neuropsychological Society*, 1–12.

Waters, G. S. & Caplan, D. (1997). Working Memory and On-Line Sentence Comprehension in Patients with Alzheimer's Disease. *Journal of Psycholinguistic Research*, *26*(4), 377–400. doi:10.1023/A:1025073902269

Waters, G. S., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence comprehension in patients with alzheimer's disease. *Cognitive Neuropsychology*, *12*(1), 1–30. doi:10.1080/02643299508251990

Wechsler, D. (2008). Wechsler adult intelligence scale–fourth edition (wais–iv). *San Antonio, TX: NCS Pearson*, *22*, 498.

White-Devine, T., Grossman, M., Robinson, K., Onishi, K., Biassou, N., & D'Esposito, M. (1996). Verb confrontation naming and word-picture matching in Alzheimer's disease. *Neuropsychology*, *10*(4), 495–503.

WHO. (2016a). *The international classification of diseases, clinical modification: ICD-CM*. DHHS publication ; Washington, D.C.: U.S. Dept. of Health and Human Services, Public Health Service, Health Care Financing Administration : For sale by the Supt. of Docs., U.S. G.P.O.

WHO. (2016b). WHO Dementia Fact Sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs362/en/

Williamson, D. J., Adair, J. C., Raymer, A. M., & Heilman, K. M. (1998). Object and Action Naming in Alzheimer's Disease. *Cortex*, *34*(4), 601–610. doi:10.1016/S0010-9452(08)70517-3

Wilson, S. M., Dronkers, N. F., Ogar, J. M., Jang, J., Growdon, M. E., Agosta, F., … Gorno-Tempini, M. L. (2010). Neural correlates of syntactic processing in the non-fluent variant of primary progressive aphasia. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(50), 16845–16854. doi:10.1523/JNEUROSCI.2547-10.2010

Wilson, S. M., Galantucci, S., Tartaglia, M. C., & Gorno-Tempini, M. L. (2012). The neural basis of syntactic deficits in primary progressive aphasia. *Brain and language*, *122*(3), 190–198. doi:10.1016/j.bandl.2012.04.005

Wingfield, A. [Arthur] & Stine-Morrow, E. A. L. (2000). Language and speech. In F. I. M. Craik & T. A. Salthouse (Eds.). Lawrence Erlbaum Associates Publishers.

Wong, L., Martin-Khan, M., Rowland, J., Varghese, P., & Gray, L. C. (2012). The Rowland Universal Dementia Assessment Scale (RUDAS) as a reliable screening tool for dementia when administered via videoconferencing in elderly post-acute hospital patients. *Journal of Telemedicine and Telecare*.

Zacks, R. T. & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory.

Zannino, G. D., Perri, R., Pasqualetti, P., Caltagirone, C., & Carlesimo, G. A. (2006). (Category-specific) semantic deficit in Alzheimer's patients: The role of semantic distance. *Neuropsychologia*, *44*(1), 52–61. doi:10.1016/j.neuropsychologia.2005.04.008

Zhu, Z., Hou, X., & Yang, Y. (2018). Reduced Syntactic Processing Efficiency in Older Adults During Sentence Comprehension. *Frontiers in Psychology*.

Part II

Papers

Ribu, I. S. (Under revision). First in, last out: A systematic review of the effects of (psycho-)linguistic variables on lexical retrieval in Alzheimer's disease and primary progressive aphasia

I

First in, last out: A systematic review of the effects of (psycho-)linguistic variables on lexical retrieval in Alzheimer's disease and primary progressive aphasia

Ingeborg Sophie Ribu
Department of linguistics and Scandinavian studies, University of Oslo, Po.box. 1102, Blindern, 0317 Oslo, Norway.
i.s.b.ribu@iln.uio.no

Abstract

This systematic review explores how psycholinguistic variables of words affect lexical retrieval in Alzheimer's disease (AD) and primary progressive aphasia (PPA) for monolingual and bilingual speakers. Word class, frequency, age of acquisition (AoA) and imageability, affect lexical retrieval in different ways across pathologies. In this review the effects of these variables are studied in naming and comprehension in different tasks and modalities. Forty-eight studies are reviewed; 32 studies focus on AD, 14 on PPA and 6 on both AD and PPA. The results show that AoA is the variable with the strongest effect on lexical retrieval in both production and comprehension, which suggests a primary retrieval impairment. PPA and AD can be successfully distinguished from healthy controls based on the AoA, frequency and imageability of words produced during testing. Bilingual speakers with AD and PPA show similar patterns of deterioration in both languages, which indicates a shared bilingual mental lexicon.

Keywords: Alzheimer's disease; Primary Progressive Aphasia; frequency; age of acquisition; imageability; word class; lexical retrieval; bilingualism

1. Introduction

Most dementia diseases will affect language behaviour and use, in addition to other cognitive domains. Alzheimer's disease (AD) and Primary Progressive Aphasia (PPA) are two different kinds of dementia, which are known to affect language already early in the course of the disease. AD affects memory and other cognitive domains, including, but not limited to disturbances in executive functions, impairments in language, orientation and judgment ("DSM-5," 2016; "ICD-10 Version:2016," 2016). In PPA language impairments are the first symptoms, and other cognitive functions and domains are affected later (Gorno-Tempini et al., 2011). Changes in linguistic behaviour can help distinguish between different subtypes of PPA, and aid diagnosis. Impairments in lexical access and anomia, or word finding difficulties, are the most prominent language symptoms for both AD and PPA (Bayles & Tomoeda, 2007; Gorno-Tempini et al., 2011; McKhann et al., 2011; Weintraub, Rubin, & Mesulam, 1990). However, as the two syndromes are clinically and behaviourally different, there is no reason to assume that the anomia is identical in the two syndromes.

This review investigates different variables that may affect lexical access in production and comprehension for mono- and bilingual [1] persons with AD and PPA. The variables are frequency (how often a word occurs in spoken or written text), (subjective) age of acquisition[2] (henceforth AoA; how early a word is acquired in childhood), and imageability (how easily a word gives rise to a sensory mental image). Furthermore, the lexical variable word class is also included.[3] Research suggests that verbs/actions and nouns/objects are differentially affected by AD and PPA (Druks et al., 2006; Hillis, Oh, & Ken, 2004; G. Robinson, Rossor, & Cipolotti, 1999; Rodríguez-Ferreiro, Davies, González-Nosti, Barbón, & Cuetos, 2009). It is therefore of interest to look at word class differences in the two pathologies. These variables are frequently investigated in studies of lexical access of healthy and clinical populations, as they can help shed light on access to and structuring of the mental lexicon. One can learn a lot about how language is organized in the healthy brain by studying language impairments from different clinical populations.

1.1.Aims and research questions

This systematic review examines studies that focus on lexical access and retrieval in patients with AD and PPA as well as on underlying psycholinguistic factors that affect lexical access and retrieval. The main goal is to identify which variables most affect lexical retrieval in these two patient groups, and whether they differ across pathologies and languages.

To investigate these issues, the articles included in this review will be used to answer the following research question with sub-questions:

How do the different (psycho-)linguistic variables (word class, frequency, AoA and imageability) affect lexical access in (oral) production and (auditory) comprehension for persons with AD and PPA?

---

[1] Bilingualism is used in a broad sense in the remainder of this review to refer to speakers of two or more languages, regardless of the number of languages spoken by the individual.
[2] Subjective AoA is measured by asking adults how old they think they were when they learned a word. This measure correlates well with when children actually learn words (Hansen, 2016; Łuniewska et al., 2016).
[3] While frequency, AoA, and imageability are psycholinguistic variables, word class is not. In the following sections the term *(psycho-)linguistic variables* will be used to refer to all four variables together.

- Can the different variables be used to discriminate between pathologies, and different subtypes of PPA?
- How do the (psycho-)linguistic variables affect naming and comprehension across languages for bilingual persons with AD or PPA, and what does this tell us about language organization in bilinguals?

A better understanding of the mechanisms underlying naming difficulties in different populations can have both clinical and theoretical implications. Knowledge about how language is affected by different neurodegenerative diseases can help shed light on the relationship between language and the brain, for instance to inform on what parts of the language system or access to language is impaired. Clinically it can help develop clearer diagnostic criteria, and aid diagnosis.

1.2. Language in Alzheimer's disease and primary progressive aphasia

Although the main symptoms of AD are problems with episodic memory and learning capacity, anomia is also reported as an early symptom (Morris, 1996). Proper names are usually affected earlier than common names (Semenza, Mondini, Borgo, Pasini, & Sgaramella, 2003). Typically, comprehension and knowledge of word meaning are preserved in the initial stages of the disease. Other language functions will deteriorate over time, such as auditory comprehension, reading and writing. In the latest stages of the disease language expression becomes limited to a working vocabulary of few words, or even mutism (Altmann & McClung, 2008; Kirshner, 2012). Spoken discourse is often grammatical early on, although grammatical errors can be found earlier in writing (Bayles & Tomoeda, 2007).

A unified diagnosis of PPA and its three subtypes were first agreed upon in 2011(Gorno-Tempini et al., 2011). Articles before this time may use different names to refer to the same disease. In the literature, PPA has been known as "slowly progressing aphasia", "progressive non-fluent aphasia" or "semantic dementia" before, and to some extent also after, 2011.

PPA is a degenerative disease with, in most cases, semantic degeneration as the core symptom. Three subtypes are identified based on their language manifestations: a logopenic variant (lvPPA), a semantic variant (svPPA) and a non-fluent/agrammatic variant (nfavPPA). Despite these well-defined subtypes, as much as 40% of persons with PPA are not classifiable into either of the variants; either because they do not fit the subtypes, or because they show a mixed pattern (Kertesz, 2003; Sajjadi, Patterson, Tomek, & Nestor, 2012). There are no reliable numbers on the distribution of the subtypes, but svPPA is the most frequently reported subtype of PPA, followed by nfavPPA, and then lvPPA. Some of the studies included in this review do not specify which type of PPA is under investigation, either because they were written before the sub-classification system was properly agreed upon, or because there is uncertainty about the diagnosis. Unless the subtype is clearly stated in the original paper, there will be made no attempt of classification here.

lvPPA is identified by impaired single-word retrieval in spontaneous speech and impaired repetition, as well as phonological errors in spontaneous speech and naming, but spared single-word comprehension and object knowledge. In the nfavPPA single-word comprehension and object knowledge is spared, but language production is in most cases clearly agrammatic and effortful with inconsistent speech sound errors and distortions. Finally, in svPPA object knowledge and single-word comprehension and confrontation naming is impaired, especially for low-frequent items, but repetition, motor-speech and grammar are spared (Gorno-Tempini et al., 2011).

AD and PPA are closely related and easily compared, yet there are some core differences: In AD there is decline in at least two areas of cognition and a progressive worsening of memory and other cognitive functions ("DSM-5," 2016; "ICD-10 Version:2016," 2016; McKhann et al., 2011). Whereas a clinical diagnosis of PPA requires a prominent and isolated language decline, with impairments in language production, object naming, grammar, or word comprehension apparent in conversation and during language testing, but without any decline in other cognitive domains during the first two years of the disease (Mesulam, 2001).

1.3.(Psycho-)linguistic variables and lexical access

All words have variables that are linked to form, meaning, and usage patterns that affect how they are stored and processed in the mental lexicon. For instance, word frequency was early recognized as an important factor in studies of lexical retrieval, processing and use (Oldfield & Wingfield, 1965). However, frequency is not the only variable that affects retrieval. Morrison, Ellis, & Quinlan, (1992) re-analysed the data from Oldfield and Wingfield (1965) and found that age of acquisition (AoA) was a better predictor than frequency for naming accuracy for healthy adults. Furthermore, imageability is a semantic variable that affects how easily words are accessed from the mental lexicon: words that more easily conjure a mental sensory image are easier to access than words with low imageability (Bird, Howard, & Franklin, 2003; Paivio, Yuille, & Madigan, 1968).[4]

These variables correlate to some extent. AoA is related to frequency, as more frequent words are often learned earlier in childhood than less frequent words (Juhasz, 2005). AoA also correlates with imageability (Bird et al., 2003; Paivio et al., 1968), as earlier learned words often denote concrete objects, which tend to be more imageable than later acquired and abstract words (Andrew. W. Ellis, 2012).

Finally, several studies have shown that verbs and nouns can be differentially affected in both comprehension and production for persons with language impairments (see Kambanaros & Grohmann, 2015 for an overview). Studies of verb-noun dissociation in AD seem to show little consistency; some studies find that persons with AD are more impaired on verb retrieval than on noun retrieval (Druks et al., 2006), whereas others find the opposite pattern (G. Robinson et al., 1999). Noun-verb dissociations can be indicative of PPA subgroup, as verbs seem to deteriorate first in nfavPPA, and nouns first in svPPA (Hillis et al., 2004; Kim & Thompson, 2004; Thompson, Ballard, Tait, Weintraub, & Mesulam, 1997). A majority of studies into lexical retrieval focus in particular on noun retrieval, and verb retrieval is far less studied. This review investigates how different word classes are affected by dementia, if there is a difference in production and comprehension, across tasks and modalities, and languages.

2. Methods

2.1.Literature search

Four searches were conducted in the ScienceDirect and PubMed electronic databases for full text articles, the first one on November 16th 2016. The following keywords were entered into the “Title”, “Abstract” and “key words” fields in the advanced search option, as well as MeSH terms on PubMed[5]: ‘Alzheimer’s disease’, ‘Primary Progressive Aphasias’, combined with ‘Lexical Access’, ‘Imageability’, ‘Age of Acquisition’, ‘Word frequency’ and ‘Word Class’.

On December 12th 2016, a second search was carried out in the most recent issues of the journals that had published one or more study included after the first search, to identify any

[4] Other variables will also affect use and processing, including, but not limited to, familiarity, name agreement, semantic categories, concreteness, phonological neighborhood density and word length amongst others. A review of all the possible underlying variables that affect lexical access in AD and PPA is beyond the scope of this article. Factors that are proven to affect both verbs and nouns are included.

[5] Medical Subject Headings; not available for ScienceDirect.

additional studies. On March 16th 2018 a new search with the same search terms was conducted to identify new titles in the same databases. The final search was carried out on November 9th 2018 specifically to identify further papers on lexical access in bilinguals.

Only full-text articles written in English were included, and there were no restrictions with regard to publication year.

2.2.Study selection and inclusion criteria

The database searches returned 433 results, but only 48 articles were included in the final review. The selection process is shown in the flow chart in Figure 1, based on a model from the PRISMA group (Moher, Liberati, Tetzlaff, Altmann, & The PRISMA Group, 2009).

For a study to be eligible for inclusion in the review it had to be original research published in a peer-reviewed journal; abstracts, book chapters, short communications, editorial notes and other reviews were excluded. The articles had to report on either AD or PPA or both. Four studies were excluded because they did not specify type of dementia. Articles reporting on lexical access in AD/PPA and other pathologies were also included.[6] Articles reporting on test development, or norming of tests in different languages were excluded.

Among the original 433 titles there were 59 articles written in other languages than English (including duplicated titles) and another 31 duplicates. These were removed and a thorough screening of titles, abstracts and methodology gave another 302 titles for exclusion, as they did not investigate the effect of the variables on lexical access. A majority of these studies (n=158) were non-linguistic studies (medical studies, treatment studies, memory studies etc.) that relied partly on linguistic information for disease classification. 11 studies investigated bilingual advantage, the hypothesis that bilinguals are better protected against cognitive decline than monolinguals, these were also excluded as they do not report on how the variables affect lexical access.

Only studies that discussed behavioural data were included, as the methodologies and results from structural and functional imaging studies, eye-tracking, ERP and EEG studies are difficult to compare with studies using behavioural methods.[7]

[6] The results regarding other pathologies than AD or PPA will not be discussed in this review.
[7] Two cases (Mack et al., 2015; Venneri et al., 2008) included both behavioral and fMRI data. In those cases only the behavioral data were included in this review.

Studies reporting solely on reading and writing rather than oral production and auditory comprehension were excluded, as reading and writing may be differently affected than naming and comprehension in AD and PPA (Bayles & Tomoeda, 2007; Hillis, Tuffiash, & Caramazza, 2002; Matías-Guiu et al., 2017). This final screen left 41 articles eligible for inclusion. The articles cited in the references of these articles were also screened for title, abstract and methodology and considered for inclusion, which gave 19 new titles.

[Figure 1: flow chart in here]

Figure 1. Overview of the study selection process

The articles that were eligible for review reported on the effects of some or all of the following variables on lexical access: word class, word frequency, AoA or imageability. Studies that focused on other variables in addition to at least one of these were also included. Studies that reported on semantic distinctions (object vs. action) were included, as objects are always nouns and actions typically verbs (Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). Finally, studies equating familiarity with frequency and concreteness with imageability were excluded, to ensure that the variables were the same across all studies. AoA in all studies are based on subjective AoA, which correlates strongly with objective AoA (Hansen, 2016; Łuniewska et al., 2016).

Studies from before 2011 that included persons with “semantic dementia”, “slowly progressive aphasia” or similar, were also included, as these persons would be classified as having a variant of PPA according to the current diagnostic criteria (Gorno-Tempini et al., 2011).

Many studies originate from the same research labs and it can therefore not be excluded that the same participants are reported on in multiple papers. 36 of the studies came from in total eleven different research groups, and in two articles (Hillis et al., 2004, 2002) the researchers reported on (partly) the same participants in the different studies. These studies were treated as separate studies, and the results have not been collapsed. Table 1 gives an overview of the articles that are included in the review.

Study		**Participants**			
First author	Year	AD	PPA	Controls	Language
Shuttleworth	1988	20		25	English
Goldstein	1992	41		14	English
Wallesch	1994	11			German
Hirsch	1995		2		English
Montanes	1995	25		25	French
Robinson	1996	20		18	English
White-Devine	1996	21		14	English
Glosser	1997	20 + 21 + 18[a]		17	English

Thompson	1997		4	5	English
Laiacona	1998	26			Italian
Williamson	1998	10		10	English
Robinson	1999	1			English
Ukita	1999		1		Japanese
Bird	2000		3	20	English
Kremin	2001	8	8		French
Hillis	2002		3	5	English
Silveri	2002	39		12 + 20	Italian
Hillis	2004		15 + 7		English
Kim	2004	14		10 + 10	English
Cuetos	2005	10			Spanish
Forbes-MacKay	2005	96		40	English
Druks	2006	19		19	English
Hillis	2006		27 + 16		English
Holmes	2006	22		22	English
Marczinski	2006	20	8	20	English
Albanese	2007	48		40 + 40	Italian
Hernandez	2007	1			Catalan-Spanish
Silveri	2007		22 + 7		Italian
Tippet	2007	58		59	English
Cuetos	2008	2			Spanish
Hernandez	2008		1		Spanish-Catalan
Venneri	2008	25		25	English
Martinaud	2009	50		56	French
Peters	2009	16		16 + 16	French
Rodríguez-Ferreiro	2009	20		20	Spanish
Sailor	2011	34		36	English
Costa	2012	47			Spanish-Catalan
Cuetos	2012	20 + 20		20 + 20	Spanish and English
Druks	2013		1		Hungarian-English
Foster	2013	25		20	English
Ivanova	2013	68+18		44+11	English and Spanish-English
Fraser	2014		10 + 14	16	English
Marcotte	2014		23	15	English
Vita	2014	20		45	Italian
Beber	2015	35			Brazilian Portuguese
Mack	2015		12 + 11 + 12	12	English
Rumiati	2016	14	9	30	Italian
Lind	2018		1		English-Norwegian

[a] A '+' indicates that the studies include several subgroups of participants

Table 1. Studies included in the review

3. Results

3.1.Study characteristics

Thirty-two studies focused only on AD, 14 on PPA, and three on both AD and PPA. Six studies investigated lexical retrieval in bilinguals; three focusing on AD and three on PPA. Another 33 studies compared persons with AD or PPA to neurologically healthy participants. In general, the controls were matched on age and education, and in four cases the studies also included young controls (Albanese, 2007; Peters, Majerus, De Baerdemaeker, Salmon, & Collette, 2009; Silveri, Cappa, Mariotti, & Puopolo, 2002; Vita et al., 2014).

All but 14 studies included word frequency as a measure, and five studies included no other measure than frequency. Word class differences was the most studied variable (n=21), followed by AoA (n=19). One study investigated verb production alone (Beber, da Cruz, & Chaves, 2015). Ten studies investigated the influence of imageability on lexical retrieval. Furthermore, 27 of the studies included other variables in addition to the ones of interest. Results from these other variables will not be discussed.

Most of the studies were conducted in English with English-speaking participants (n=26), followed by Italian (n=6), Spanish and French (n=4 each), and one study in each of the languages Brazilian Portuguese, German, and Japanese. One study included both Spanish and English-speaking participants. Three studies included Spanish-Catalan bilinguals; other language pairs in the studies that included bilingual participants were Spanish-English, Hungarian-English and English-Norwegian. That the majority of the studies were conducted in English has implications for the noun vs. verb measure, which will be discussed below.

3.2.Alzheimer's disease

AD diagnosis was mainly based on NINCDS-ADRA criteria (McKhann et al., 1984, 2011). Mental status in all studies was based on the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). Subgrouping based on severity was reported on in only six studies, usually by means of MMSE scores. However, the cut-off scores used for sub-grouping based on severity is not the same across studies.

Many of the studies had a main focus on category specific impairments; either in terms of semantic distinctions between action- and object naming/comprehension, or categories involving living/animate and non-living/inanimate objects, however, this last distinction will not be discussed here. An overview of the studies with a focus on AD is found in table 2 below.

STUDY	PRODUCTION / COMPREHENSION	WORD CLASS	FREQUENCY	AOA	IMAGEABILITY	OTHER
Shuttleworth, 1988	P		X			
Goldstein, 1992	P		X			
Wallesch, 1994	P + C		X			X
Montanes, 1995	P		X			X
Robinson, 1996	P	X				
White-Devine, 1996	P + C	X				X
Glosser, 1997	P		X			
Laiacona, 1998	P		X			X
Williamson, 1998	P	X	X			
Robinson, 1999	P + C	X				
Kremin, 2001	P + C		X	X		X
Silveri, 2002	P		X	X		X
Kim, 2004	P + C	X				

Cuetos, 2005	P		X	X	X	
Forbes-MacKay, 2005	P		X	X		X
Druks, 2006	P	X			X	X
Holmes, 2006	P			X		X
Marczinski, 2006	P		X			X
Albanese, 2007	P		X	X	X	X
Hernández, 2007	P	X	X			
Tippet, 2007	P		X	X		X
Cuetos, 2008	P		X	X	X	X
Venneri, 2008	P		X	X		X
Martinaud, 2009	P		X	X		X
Peters, 2009	C				X	
Rodríguez-Ferreiro, 2009	P	X	X	X		
Sailor, 2011	P		X	X		
Costa, 2012	P		X			
Cuetos, 2012	P		X	X	X	X
Foster, 2013	P		X			
Ivanova, 2013	P		X		X	X
Vita, 2014	P		X			X
Beber, 2015	P	X	X			
Rumiati, 2016	P + C		X	X		X

Table 2. Overview of studies focusing on AD and the different variables they investigate

Nouns and verbs in AD

Nine studies investigated word class differences in AD, but the results from these are not conclusive. Both a selective sparing of nouns (Druks et al., 2006; Kim & Thompson, 2004; K. M. Robinson, Grossman, White-Devine, & D'Esposito, 1996; White-Devine et al., 1996) as well as a selective sparing of verbs (Hernández, Costa, Sebastián-Gallés, Juncadella, & Reñé, 2007; G. Robinson et al., 1999; Williamson, Adair, Raymer, & Heilman, 1998) were found. Interestingly, one study found no selective sparing for word class (Rodríguez-Ferreiro et al., 2009). The authors partly ascribe this to the fact that studies that find a noun/object or verb/action deficit mainly have been conducted in English, and that the high number of words that can be either verbs or nouns (*a fish* vs. *to fish*) have different frequencies and AoA, and can therefore not be compared directly. Their study was conducted in Spanish, and items that could belong to different word classes were excluded, which might in part explain their results. However, this does not explain why Hernandez and colleagues found a noun naming deficit for their participant, in both Spanish and Catalan (Hernández et al., 2007).

Both Kim and Thompson (2004), and Beber et al. (2015) investigated verb deficit in more detail and found that verb confrontation naming was more impaired than verb fluency (Beber et al., 2015), and that verb comprehension was better preserved than verb production (Kim &

Thompson, 2004). Verb production seems to be more demanding if the verbs require several arguments. Persons with AD also seem to use general **--** (i.e. *to clean*) rather than specific verbs (i.e. *to scrub*) (Kim & Thompson, 2004). However, these items are often also of higher frequency than the more specific verbs. Verb naming, but not verb fluency was found to be more impaired depending on disease severity (Beber et al., 2015).

Robinson (1999) found that verbs were better preserved in naming than nouns. As this was a case study, it might be an example of individual variation. However, similar results are reported by Hernández et al. (2007); selective sparing for verbs compared to nouns was found in both L1 Catalan and L2 Spanish (Hernández et al., 2007). Williamson et al., report similar findings for the 10 participants in their study: object naming was more impaired than action naming (Williamson et al., 1998). Table 3 gives an overview of the main results from the studies with a focus on word class differences in AD.

STUDY	TASK/MODALITY	FINDINGS
Robinson, 1996	Confrontation naming	nouns > verbs
White-Devine, 1996	Confrontation naming	nouns > verbs
	Word-to-picture matching	nouns > verbs (approach significance)
Williamson, 1998	Confrontation naming	verbs > nouns
Robinson, 1999	Oral naming	verbs > nouns
Kim, 2004	Naming	nouns > verbs
	Comprehension	no difference between verb and noun comprehension
	For verbs only	comprehension of verbs > naming of verbs
	Narrative production	one-place verbs > three-place verbs
	Sentence completion	Higher rate of general verbs compared to specific verbs.
Druks, 2006	Confrontation naming	Objects > actions
Hernandez, 2007	Confrontation naming	Disproportionate impairment for nouns in L1 (Catalan) and L2 (Spanish)
Rodríguez-Ferreiro, 2009	Confrontation naming	Word class not important for naming success for persons with AD.
Beber, 2015	Confrontation naming vs. verb fluency	Verbs more impaired in confrontation naming than in fluency task

Table 3. Overview of papers focusing on AD and word class differences

Frequency in AD

Frequency was the most studied variable for AD, with 27 articles reporting on frequency effects. A majority of the studies find an effect of frequency on naming performance, with high-frequent words being named faster and more accurately than low-frequent words in confrontation naming (Beber et al., 2015; Costa et al., 2012; Cuetos, Gonzalez-Nosti, & Martínez, 2005; Cuetos, Rodríguez-Ferreiro, Sage, & Ellis, 2012; Cuetos, Rosci, Laiacona, & Capitani, 2008; Goldstein, Green, Presley, & Green, 1992; Hernández et al., 2007; Laiacona, Barbarotto, &

Capitani, 1998; Montanes, Goldblum, & Boller, 1995; Rodríguez-Ferreiro et al., 2009; Rumiati, Foroni, Pergola, Rossi, & Silveri, 2016; Shuttleworth & Huber, 1988; Tippett, Meier, Blackwood, & Diaz-Asper, 2007; Wallesch & Hundsalz, 1994; Williamson et al., 1998). Martinaud, Opolczynski, Gaillard, & Hannequin, (2009) used logic regression to see if there was an effect of frequency on an individual level and found that frequency predicted naming success for a small number of their participants (8 out of 50) on a confrontation naming task. There was no effect of frequency on group level. Ivanova et al. (2013) found that frequency predicted naming accuracy on a picture naming task when no other variables were included in the analysis. Once other variables, including imageability, were taken into account, frequency had no predictive power on naming accuracy (Ivanova, Salmon, & Gollan, 2013).

Frequency affected fluency tasks differently: On semantic fluency tasks persons with AD in general provide words with higher frequency than controls (Forbes-McKay, Ellis, Shanks, & Venneri, 2005; Foster et al., 2013; Venneri et al., 2008). Furthermore, higher-frequency items were generated on an animal fluency task compared to the items generated on a grocery fluency task (Marczinski & Kertesz, 2006).

On phonological fluency tasks, the results are less conclusive. Foster et al., (2013) report that the frequency for words produced in phonological fluency are lower for persons with AD than for controls, whereas Sailor, Zimmerman, & Sanders (2011) report the opposite pattern. Furthermore, the frequency of the words produced in the phonological fluency task is higher than the words produced in the semantic fluency task reported by Sailor et al. No such difference was found in the semantic and phonological fluency scores reported by Foster and colleagues. Frequency effects were also found in a repetition task (Glosser & Friedman, 1991).

In contrast, a few studies found no effect of frequency in confrontation naming tasks (Albanese, 2007; Kremin et al., 2001; Silveri et al., 2002), fluency tasks (Beber et al., 2015; Vita et al., 2014), or comprehension (Costa et al., 2012; Hernández et al., 2007; Wallesch & Hundsalz, 1994).

Silveri and colleagues (2002) suggest that the lack of a frequency-effect in their sample may stem from a correlation with AoA, and that AoA is a better predictor for naming success in AD than frequency. The main results from the studies focusing of frequency in AD can be found in table 4 below.

STUDY	TASK/MODALITY	FINDINGS

Shuttleworth, 1988	Confrontation naming Response time production	high frequency items = medium frequency items > low frequency items high frequency items = medium frequency items > low frequency items
Goldstein, 1992	Response time production	High frequency items > low frequency items
Wallesch, 1994	Production Comprehension	greater naming accuracy for high-frequent items no significant effect of word frequency
Montanes, 1995	Production accuracy	high frequent items > low frequent items
Glosser, 1997	Repetition task	High-frequent words > low frequent words > pseudo words
Laiacona, 1998	Naming	Frequency significantly affected naming success.
Williamson, 1998	Naming	Frequency important for naming success, but not as much as action/object distinction
Kremin, 2001	Picture naming	No effect of frequency
Silveri, 2002	Naming	No effect of frequency
Cuetos, 2005	Naming	Frequency significantly affected naming success
Forbes-MacKay, 2005	Semantic fluency	More high-frequent words than healthy controls
Marczinski, 2006	Animal fluency vs. grocery fluency	Higher frequency on animal fluency task Compared to patients with PPA: lower frequency on animal fluency, but higher frequency on grocery fluency.
Albanese, 2007	Naming	No effect of frequency
Hernandez, 2007	Confrontation naming Word-picture matching	Frequency effect on noun and verb naming in L1 and L2 No effect of frequency
Tippet, 2007	Naming	Frequency significantly affected naming success.
Cuetos, 2008	naming	Frequency significantly affected naming success.
Venneri, 2008	Semantic fluency	Higher frequency items than healthy controls
Martinaud, 2009	Confrontation naming	Frequency effects for 8/50 persons with AD, no group effect
Rodríguez-Ferreiro, 2009	Naming	Frequency significantly affected naming success
Sailor, 2011	Verbal fluency	Higher frequency for words in a letter fluency task than in semantic fluency
Costa, 2012	Picture naming Word translation Picture-word matching	Frequency significantly affected naming success in both L1 and L2 No effect of frequency in either L1 or L2 No effect of frequency in either L1 or L2
Cuetos, 2012	Naming	Effect of frequency for both English and Spanish participants
Foster, 2013	Verbal fluency	Lower frequency items on phonological fluency, and higher frequency items on semantic fluency compared to controls
Ivanova, 2013	Picture naming	No effect of frequency when other variables are included
Vita, 2014	Fluency	No effect of frequency
Beber, 2015	Verb confrontation naming Verb fluency task	Frequency significantly affect naming performance No effect of frequency
Rumiati, 2016	Confrontation naming	Frequency significantly affected naming success.

Table 4. Overview of studies focusing on AD and frequency

AoA in AD

All but one of the 16 studies (Rumiati et al., 2016) found an effect of AoA, but the results vary for different tasks. Martinaud and colleagues could only identify AoA effects for 19 out of 50 participants with AD with logic regression; no effect was found on group level (Martinaud et al., 2009). In all the other studies included here, early acquired words were processed more accurate and faster in confrontation naming tasks (Albanese, 2007; Cuetos et al., 2005, 2012, 2008; Holmes & Ellis, 2006; Kremin et al., 2001; Martinaud et al., 2009; Rodríguez-Ferreiro et al., 2009; Silveri et al., 2002; Tippett et al., 2007), semantic fluency tasks (Forbes-McKay et al., 2005; Sailor et al., 2011; Venneri et al., 2008), and in object recognition (Holmes & Ellis, 2006).

AoA also seems to be an important factor as the disease progresses: In a longitudinal study, Cuetos et al. (2005) found that words that were named correctly two years apart were words with early AoA. Many of the words that were not named correctly at the second time point were words with later AoA than the ones that had been named correctly at both time points. (Cuetos et al., 2005). Another study by Cuetos and colleagues (2008), and by Tippet and colleagues (2007) confirmed this finding: The AoA of words produced in picture naming tasks were a predictor for naming success as the disease progressed (Cuetos et al., 2008; Tippett et al., 2007).

It seems that the AoA of words generated on verbal fluency tasks can distinguish persons with AD from controls: Persons with AD produce words with earlier AoA than do controls, but it cannot successfully discriminate between disease severity (Forbes-McKay et al., 2005; Sailor et al., 2011; Venneri et al., 2008). Furthermore, the AoA of words generated on semantic fluency tasks had earlier AoA than the words on a letter fluency task (Sailor et al., 2011).

Tippet et al. (2007) investigated category-specific naming deficits for living vs. non-living things, controlling for different variables including frequency and AoA. When items were matched for frequency, persons with AD named living things better than non-living things. This category deficit disappeared when the items were further matched for AoA in addition to frequency (Tippett et al., 2007). In general, AoA seems to have a strong influence on lexical access across tasks. Table 5 sums up the main findings of the studies focusing on AoA.

STUDY	TASK/MODALITY	FINDINGS
Kremin, 2001	Confrontation naming	AoA significantly affected naming success.
Silveri, 2002	Confrontation naming	AoA significantly affected naming success.
Cuetos, 2005	Confrontation naming	AoA significantly affected naming success.
Forbes-MacKay, 2005	Semantic fluency	Persons with AD produced words with earlier AoA than controls.
Holmes, 2006	Object recognition Object naming	Early AoA > late AoA Early AoA > late AoA
Albanese, 2007	Confrontation naming	AoA significantly affected naming success.
Tippet, 2007	Confrontation naming	AoA significantly affected naming success.
Cuetos, 2008	Confrontation naming	AoA significantly affected naming success.

Venneri, 2008	Semantic fluency	Persons with AD produced words with earlier AoA than controls.
Martinaud, 2009	Confrontation naming	AoA effects for 19/50 persons with AD, but no group effect.
Rodríguez-Ferreiro, 2009	Confrontation naming	AoA significantly affected naming success.
Sailor, 2011	Verbal fluency	AoA was lower for the words produced in a semantic fluency task compared to those produced in a letter fluency task.
Cuetos, 2012	Confrontation naming	AoA significantly affected naming success in both Spanish and English.
Rumiati, 2016	Confrontation naming	No effect of AoA

Table 5. Overview of studies focusing on AD and Age of Acquisition

Imageability in AD

Four out of seven studies that investigated how lexical access in AD is influenced by imageability, found no effect (Albanese, 2007; Cuetos et al., 2005, 2012, 2008). These four studies all focused on confrontation naming, and also reported on effects of frequency and AoA, which seemed to have a stronger predictive power on naming success than imageability.

If imageability is the only variable included in analysis, it shows a positive effect on lexical access; high-imageable words are processed faster and more accurately than low-imageable words. This is found both for naming accuracy (Druks et al., 2006; Ivanova et al., 2013), and on serial recall and synonym judgment tasks (Peters et al., 2009).

In one case, imageability was found to have an effect on naming success when frequency did not (Ivanova et al., 2013). The authors claim that meaning is important for object naming success in AD because high-imageable words have conceptual representations with more semantic features than low-imageable words. Naming deficits in AD stem from a damage to the integrity of semantic memory, rather than an impairment to the lexical system (Ivanova et al., 2013). However, this is the only study in the sample that identifies a beneficial effect of a semantic variable (imageability) over variables that are more associated with retrieval (AoA and frequency).

One reason why imageability shows a less clear effect may be that the range of the variable is too small, or the mean of the variable is too high (Cuetos et al., 2005). This is a result of using high-imageable words in picture naming tasks. Items that are easy to depict are often high in imageability. Another reason might be that other variables, such as AoA and frequency mask the effect of imageability. A summary of the main findings from the studies that investigated imageability can be found in table 6 below.

STUDY	TASK/MODALITY	FINDINGS
Cuetos, 2005	Confrontation naming	No effect of imageability.
Druks, 2006	Object and action naming accuracy	Imageability significant predictor
	Object and action naming latency	Imageability significant predictor
Albanese, 2007	Confrontation naming	No effect of imageability.
Cuetos, 2008	Confrontation naming	No effect of imageability.
Peters, 2009	Serial recall task	High imageability > low imageability
	Synonym judgment	High imageability > low imageability
Cuetos, 2012	Confrontation naming	No effect of imageability in either Spanish or English
Ivanova, 2013	Picture naming	Imageability predict naming success

Table 6. Overview of the studies focusing on AD and imageability

3.3.Intermediate discussion

The fact that there is no general pattern when it comes to word class differences can be ascribed to individual variation among the participants or the items under investigation. The studies included in the current review show both a selective sparing for nouns and for verbs and even no difference between word classes at all.

When several variables are studied in the same sample, AoA seems to be the best predictor for naming success (accuracy and response time), and in different fluency tasks. This same pattern also holds true for comprehension and word recognition tasks. Other factors, such as frequency and imageability seem to have less effect when the sample also includes AoA. This is more pronounced for imageability than for frequency. One explanation might be that the other variables strongly correlate with AoA, and therefore do not reach significance in statistical analysis. Furthermore, this points towards an impairment in access to lexical items rather than an underlying semantic impairment.

3.4.Primary Progressive Aphasia

Seventeen studies in this review focus on PPA. Most of these were published before 2011, and do not mention any subtype of PPA. This means that the actual diagnoses or subtypes of PPA reported on in these studies might be overlapping or inadequately distinguished from each other, due to the lack of uniform diagnostic criteria. A majority of the studies from before 2011 report on what they call non-fluent PPA (n=6), which in most cases corresponds to nfavPPA post 2011. Three studies contrasted nfavPPA with Semantic Dementia (SD).[8] Of the studies included here, only two include results from persons with lvPPA (Lind et al., 2017; Mack et al., 2015), although the diagnosis in Lind (2018) is not certain. Two other studies had participants from

[8] Which is similar to the svPPA reported on after 2011.

this subgroup, but did not report on them as the number of participants were too few (Fraser et al., 2014; Marcotte et al., 2014). In the following sections, specification of PPA subtype will only be made where there is a clear diagnosis reported in the original paper. Table 7 gives an overview of the papers focusing on PPA and the variables they include.

STUDY	PRODUCTION/ COMPREHENSION	WORD CLASS	FREQUENCY	AOA	IMAGEABILITY	OTHER
Hirsch, 1995	P + C		X	X		X
Thompson, 1997	P	X				
Ukita, 1999	P		X	X		X
Bird, 2000	P	X	X		X	
Kremin, 2001	P + C		X	X		X
Hillis, 2002	P + C	X				
Hillis, 2004	P + C	X				
Hillis, 2006	P	X				
Marczinski, 2006	P		X			X
Silveri, 2007	P	X				X
Hernández, 2008	P	X				
Druks, 2013	P	X	X		X	
Fraser, 2014	P	X	X	X	X	X
Marcotte, 2014	P	X	X	X		X
Mack, 2015	P	X	X			X
Rumiati, 2016	P + C		X	X		X
Lind, 2018	P	X	X			X

Table 7. Overview of studies with focus on PPA and the different variables included in the study

Nouns and verbs in PPA

The general picture of noun/verb dissociation in PPA is syndrome-dependent. For the nfavPPA and svPPA, there is a clear division between naming and comprehension abilities in relation to word class. Persons with nfavPPA perform better on tasks for nouns than on verbs; verbs seem to deteriorate earlier, whereas nouns are more spared. In svPPA, the opposite pattern is found; nouns are more difficult to name and comprehend than verbs, and this is seen across modalities.

Verb naming is better than noun naming for persons with svPPA on confrontation naming tasks (Hillis et al., 2004; Silveri & Ciccarelli, 2007), and noun naming is better spared than verb naming for persons with nfavPPA (Druks & Weekes, 2013; Hillis et al., 2006, 2004, 2002; Silveri & Ciccarelli, 2007), and in lvPPA (Lind et al., 2017). Verb naming is also more impaired for one participant where the PPA subtype is not reported (Hernàndez et al., 2008). For the bilingual participants this pattern was consistent across languages (Druks & Weekes, 2013; Hernàndez et al., 2008; Lind et al., 2017).

In the study by Silveri and Ciccarelli (2007) they also included a group of persons with fluent PPA, who did not show any significant dissociation between nouns and verbs. The authors operate with a very broad definition of fluent and non-fluent PPA, that does not clearly distinguish between lvPPA and some cases of svPPA. They criticize their own classification system, and state that this broad definition of fluent PPA might explain the null-results for this group (Silveri & Ciccarelli, 2007).

Several of the studies report on different samples of connected speech, ranging from story completion, picture descriptions and other narratives, to discourse samples and conversational data. These samples are analysed for noun to verb ratio and pauses before verbs and nouns as a measure of retrieval difficulties. The noun to verb ratio is higher for persons with PPA than for healthy controls in some cases (Bird, Lambon Ralph, Patterson, & Hodges, 2000; Thompson et al., 1997), and lower in one other case (Fraser et al., 2014).

Persons with lvPPA produced far fewer nouns than did controls, which might be of clinical relevance. These subjects also paused more before nouns than before verbs in connected speech than participants from the other subtypes did. This might indicate that persons with lvPPA have a small noun naming deficit that is not directly apparent in confrontation naming. An analysis of pauses in connected speech shows that PPA participants in general (all three subtypes) made more pauses than controls (Mack et al., 2015). Both Mack and colleagues and Thompson and colleagues find that participants with PPA have difficulties with verb argument structure, and make more errors on two- and three place verbs (Mack et al., 2015; Thompson et al., 1997). For one bilingual participant low-frequent nouns seem to be the most problematic words to retrieve in the L1 (English), whereas both verbs and nouns are problematic in the weaker language (Norwegian) (Lind et al., 2018).

The frequency and imageability of verbs and nouns produced in spontaneous speech can be used to separate persons with svPPA from healthy controls, as controls produce significantly more nouns than persons with svPPA, and use verbs with lower frequency and imageability than persons with svPPA (Fraser et al., 2014). Table 8. gives an overview of the findings from the included studies focusing on verb/noun differences.

Study	**Task/modality**	**Findings**	**Subtype (if reported)**
Thompson, 1997	Connected speech	Higher noun to verb ratio than healthy controls	
Bird, 2000	Connected speech	Higher noun to verb ratio	svPPA

Hillis, 2002	Naming	Nouns > verbs	nfavPPA
Hillis, 2004	Naming	Nouns > verbs Verbs > nouns	nfavPPA svPPA
Hillis, 2006	Naming	Nouns > verbs Verbs > nouns	nfavPPA svPPA
Silveri, 2007	Naming	Noun deficit Verb deficit No significant dissociation between nouns and verbs	Semantic dementia Non-fluent PPA Fluent PPA
Hernández, 2008	Naming	Nouns > verbs	
Druks, 2013	Object and action naming test	Objects named better than actions in L1 and L2. L1 performance better than L2 performance.	nfavPPA
Fraser, 2014	Connected speech	Fewer nouns than controls, lower noun-verb ratio than controls	svPPA
Marcotte, 2014	Naming	nfavPPA > svPPA on noun naming	
Mack, 2015	Connected speech	More pauses before verbs than before nouns fewer nouns than healthy controls; more pauses before nouns than before verbs	lvPPA
Lind, 2018	Picture naming Connected speech	Nouns better preserved than verbs in both L1 and L2 Difficulties with verb retrieval in L1, both verb and noun retrieval in L2	lvPPA

Table 8. Summary of findings regarding word class differences in the included studies

Frequency in PPA

Most of the studies found a beneficial effect of frequency in naming, across modalities. There are also some differences between subtypes, and as the disease progresses.

In confrontation naming frequency influences naming accuracy and response time in most cases (Bird et al., 2000; Kremin et al., 2001; Lind et al., 2017; Mack et al., 2015; Marcotte et al., 2014). However, in the study by Marcotte and colleagues this effect was most pronounced for the participants with svPPA. The nfavPPA participants made significantly more mistakes than svPPA participants on high-frequency items (Marcotte et al., 2014) Kremin and colleagues used multiple regression and found an effect of frequency on the naming performance for one participant with PPA (Kremin et al., 2001).

Hirsch and Funnell (1995) compared frequency and familiarity in naming and comprehension. They claim that a subjective rating of frequency,[9] word familiarity, is a much better predictor for word retention, at least for one patient with svPPA (Hirsh & Funnell, 1995).

On semantic fluency tasks, Marczinski et al., find that participants with PPA provide more high-frequent items on the animal fluency task compared to the grocery fluency task (Marczinski & Kertesz, 2006).

In narrative production persons with svPPA produced significantly more high-frequent nouns than lvPPA and nfavPPA participants and controls. The trend was less clear for verb frequency, but there was a trend towards more use of high-frequent verbs for persons with nfavPPA than for controls (Fraser et al., 2014; Mack et al., 2015). Participants with PPA pause more before low-frequency words than high-frequency words (Mack et al., 2015). Main findings from the studies focusing on frequency is found in table 9 below.

Study	**Task/modality**	**Findings**	**Subtype (if reported)**
Hirsch, 1995	Confrontation naming	Word familiarity > frequency	svPPA
Bird, 2000	Confrontation naming	Significant effect of frequency. Even as disease progress.	
Kremin, 2001	Confrontation naming	Significant effect of frequency for one participant.	
Marczinski, 2006	Semantic fluency	Higher frequency than on animal fluency task than grocery fluency task.	nfavPPA
Fraser, 2014	Connected speech	Less frequent nouns than svPPA, more high-frequent verbs than controls	nfavPPA
		More high-frequent nouns than nfavPPA, and more high-frequent nouns **and** verbs than controls.	svPPA
Marcotte, 2014	Confrontation naming	High frequency items > low frequency items	svPPA
		Significantly more mistakes than controls on high-frequency words	nfavPPA
Mack, 2015	Naming	Tendency for more use of high-frequency verbs compared to healthy controls	nfavPPA
	Connected speech	More pauses before low- than for high-frequent words	
	Naming	Significantly more high-frequent nouns than nfavPPA and lvPPA, no difference in verb production.	svPPA
	Connected speech	More pauses before low- than for high-frequent words	
	Connected speech	More pauses before low- than for high-frequent words	lvPPA
Rumiati, 2016	Confrontation naming	Significant effect of frequency	

[9] Word familiarity is a measure where people are asked to subjectively rate how frequent a word is in their language, and is often used when frequency data is unavailable. The measures of familiarity is assumed to correlate with frequency, but this correlation has not been extensively proven. This is a vague measure, as speakers are often not aware of how frequent words are in their language.

Lind, 2018	Confrontation naming	Better naming success on high-frequent nouns and verbs in both L1 and L2	lvPPA
	Connected speech	Difficulties with low-frequency noun retrieval in L1, difficulties with both verbs and nouns in L2.	

Table 9. Overview of results from studies on frequency and PPA

AoA in PPA

Of the six studies that investigated AoA, only two found no effect (Fraser et al., 2014; Rumiati et al., 2016). Both of these studies focused on connected speech, and included other variables than AoA as well. In the study by Fraser and colleagues, the null-effect of AoA was only found for the participants with svPPA. Their findings suggest that the AoA of nouns used in connected speech can be used to distinguish persons with nfavPPA from controls, but not nfavPPA from svPPA (Fraser et al., 2014).

In confrontation naming AoA is a significant predictor of naming accuracy. Hirsh and Funnel (1995) included AoA, frequency and familiarity in confrontation naming. AoA was the only significant predictor of naming performance for their participant with a fluent variant of PPA (Hirsh & Funnell, 1995).

In a longitudinal study of a Japanese man with unspecified PPA, Ukita and colleagues (1999) found that AoA significantly correlated with naming performance (Ukita, Abe, & Yamada, 1999). The patient showed significantly better naming performance on early acquired words, and late acquired words seemed to be more difficult to access even early in the course of the disease. Kremin and colleagues (2001) investigated amongst others the effects of AoA and frequency on picture naming for both AD and PPA patients, and found only AoA to be a significant predictor for naming success in both patient groups. The main findings from the studies investigating AoA effects are summarized in table 10.

Study	**Task/modality**	**Findings**	**Subtype (if reported)**
Hirsch, 1995	Naming	Significant effect of AoA	
Ukita, 1999	Naming	Significant effect of AoA	
Kremin, 2001	Naming	Significant effect of AoA for the majority of their participants.	
Fraser, 2014	Connected speech	AoA of nouns can be used to distinguish this group from healthy controls, but not between nfavPPA svPPA. No effect of AoA	nfavPPA svPPA
Marcotte, 2014	Naming	Significant effect of AoA	
Rumiati, 2016	Confrontation naming	No effect of AoA.	

Table 10: overview of results from studies on AoA and PPA

Imageability in PPA

Only three studies in the review include imageability as a factor in naming and comprehension in PPA (Bird et al., 2000; Druks & Weekes, 2013; Fraser et al., 2014). All studies find that imageability has a smaller effect on naming performance than frequency. Bird et al. find that for persons with svPPA general nouns with low(er) imageability are the words that are best preserved until late stages of the disease (Bird et al., 2000). However, these words are often high in frequency in addition to being high in imageability, which will aid retrieval.

Imageability was also the only significant predictor for accuracy scores in an auditory synonym judgement task in the L2 of a Hungarian-English person with PPA (Druks & Weekes, 2013). This task was part of a bigger battery to test language performance in the two languages of this participant, and it was the only auditory or oral task that showed an effect of imageability.

Samples of connected speech show that the imageability of verbs can be used to separate persons with svPPA from healthy controls. No similar effect is found for other word classes in connected speech (Fraser et al., 2014). Table 11 below summarizes the main findings from the studies on imageability in PPA.

Study	**Task/modality**	**Findings**	**Subtype (if reported)**
Bird, 2000	Connected speech	Words that are high in frequency, but with low imageability are best preserved	
Druks, 2013	Auditory synonym judgment	Imageability affected performance in L2	nfavPPA
Fraser, 2014	Connected speech	Verb imageability significantly different between controls and persons with svPPA (verbs with higher imageability produced by persons with svPPA than controls).	

Table 11: Overview of studies of imageability and PPA

3.5.Intermediate discussion

Depending on the PPA subtype, the different variables affect naming and comprehension differently. The effect of the variables is also different in testing and in analysis of samples of connected speech.

Similar to the results from the AD data, AoA is more important for naming success than frequency and imageability. Frequency effects are found in the majority of the studies (as it is the most studied variable), but once a sample includes AoA as a variable in addition to frequency, this seems to have a stronger effect. This mirrors the results found for AD above, and also a study by Bastiaanse and colleagues for persons with agrammatic aphasia (Bastiaanse, Wieling, & Wolthuis, 2016).

The results concerning imageability are difficult to generalize, as only three studies in the current review investigate the effect of imageability on lexical retrieval, and the effect is studied in very different ways in all studies. Some authors point to the difference in imageability observed between nouns and verbs, but highlight that this difference cannot account for the differences observed in naming and comprehension performance of different word classes (Bird et al., 2000). Again, these word class differences might also be influenced by the language under study, as lexical items that can belong to different word classes often have been included in the studies conducted in English, and as lexical variables seem to have a stronger influence on retrieval than semantic variables.

Verbs and nouns are differentially affected depending on PPA subtype, which might have clinical implications: Early in the course of the disease, when anomia is the most persistent deficit, an analysis of verb-noun ratio might separate the different variants of PPA from each other, and PPA from healthy controls. In all PPA subtypes, the noun-verb ratio is significantly different than for controls, and persons with svPPA have a higher noun-verb ratio than persons with nfavPPA.

4. General discussion

In this review, 48 studies of AD and PPA were reviewed to investigate how different (psycho-)linguistic variables affect naming and comprehension. The variables of interest were word class (noun-verb dissociation), frequency, age of acquisition (AoA) and imageability.

The main research question asks which (psycho-)linguistic variables affect lexical retrieval in AD and PPA. The findings from this review show that these variables all affect lexical retrieval in both AD and PPA, and that similar main patterns are observed in both pathologies.

AoA seems to be the best predictor for naming success (accuracy). If AoA is not studied, frequency greatly affects both accuracy and response time in naming and also comprehension. Imageability is the variable that shows fewer stable effects across modalities and pathologies. As AoA and frequency are more associated with access to- and imageability with the underlying semantics of lexical items, these finding might point to an access impairment rather than an underlying semantic impairment in AD and PPA.

AoA and frequency are strongly correlated, but when studied together AoA seems to be a better predictor for naming success than frequency. AoA can be seen as a cumulative frequency effect as words that are acquired early are easier to retrieve because we have encountered them more often than words that are acquired later in life. Another view is that AoA is related to the plasticity of the lexical network; when there are fewer items in the network, the connections between them will be stronger than for more items. This results in an advantage for earlier learned items over later acquired items. (A. W. Ellis & Lambon Ralph, 2000; H. D. Ellis & Young, 1977; Juhasz, 2005).

The studies included here show divergent patterns concerning noun/verb dissociations in AD and PPA; both a selective sparing for nouns and a selective sparing for verbs have been found for participants with AD. Participants with svPPA are more impaired on nouns than verbs, whereas the opposite is seen for nfavPPA. No conclusion can be drawn on the noun-verb dissociation for lvPPA from the samples included in this review.

A majority of the studies reviewed include English-speaking participants (n = 30). English has many homonyms that can belong to different word classes (*to whistle*/*a whistle)*. However, similarity in sound/form does not necessarily mean similarity in underlying linguistic variables. Frequency, AoA and imageability can vary for the same word depending on its word class. Some of the studies reporting on naming and comprehension differences between nouns and verbs do not take this difference into account, which makes it difficult to know if they have found a word class effect or an effect of any of the underlying variables. In the other languages represented in this review, the word-similarity across word classes is not as pronounced as for English, and certain studies state that they have avoided items that can belong to different word classes.

The second research question addressed how the different variables can be used to distinguish AD from PPA and the different subtypes of PPA from each other. Word class differences can be used diagnostically to separate persons with PPA from healthy controls, but the present study shows that it is not possible to distinguish persons with PPA from persons with AD based on word class differences. There are subtle differences in verb and noun use between persons with svPPA, nfavPPA and lvPPA, but the results reported on in the studies included here are not conclusive.

In addition to the number of words generated on fluency tasks, the AoA and frequency of those words can be used to distinguish persons with AD from healthy controls. Persons with AD provide words with higher frequency and earlier AoA than healthy controls.

The third research question addressed concerns how the variables affect lexical access in bilingual speakers with respect to their different languages. Six studies in this review included speakers of more than one language, three for AD and three for PPA. In these studies, frequency seemed to be the variable that had the largest effect on naming success as measured by accuracy across testing modalities and languages. For the Spanish-Catalan participants with AD, language impairment was similar in the two languages, so if nouns were more severely impaired in the L1, this was also true for the L2. The results from the other language pairs show that the deterioration of naming was similar in both languages, although the less dominant language is affected more severely by naming impairments than the dominant language.

In all cases regarding bilingual PPA the participants were more impaired in their L2 early on in the disease, but as the disease progressed similar patterns of deterioration were also visible in the L1. This was seen both in confrontation naming tasks and connected speech samples, which indicates a parallel deterioration of the bilinguals' languages.

A range of other variables that can affect lexical retrieval that have not been reviewed here. In total 27 studies in the current sample included other variables than the ones of interest in this review. The most studied factors of those are familiarity, word length, animacy, and (proto)typicality to name a few.

Many of the studies reported on investigate the impact of different variables in connected speech in PPA. However, these do not look at the variables in natural interaction and conversation, but rather uses excerpts of semi-spontaneous narratives. These speech samples usually elicit a specific vocabulary that is not necessarily representative for the person's natural speech. The transferability between these findings and language use in every-day interaction is not clear, however the insight gained from these narratives can be applied to testing situations to understand more about the underlying deficits in naming, and inform on what kinds of items should be used for assessment. Similar samples were not available for the AD participants in this review, and it is therefore not possible to make any comparisons between AD and PPA participants' speech.

5. Conclusions

In many cases, AD and PPA can be distinguished from neurologically healthy participants by analysing language data. Naming and comprehension tasks can give some insight as to which aspects of language is impaired, and this should be taken into account when assessing for cognitive decline. Results from this review show that it is difficult to separate AD from PPA based on sensitivity of these variables in linguistic data, which indicates that the language impairments in the two diseases are not very different. A thorough analysis of the different variables can be used to distinguish the different subtypes of PPA from each other, and the variables studied here are sensitive enough to identify naming and comprehension deficits in AD and PPA even early in the course of the diseases.

AoA is the variable that is most critical for naming and comprehension success across testing modalities. Frequency is the second variable that affects lexical retrieval in both diseases and in different test modalities. That these two variables are the ones with best predictive power for naming and comprehension success suggest that the lexical retrieval difficulties observed in AD and PPA are caused by an impairment in access rather than an underlying semantic impairment.

For speakers of more than one language, the lexical retrieval difficulties will affect the less dominant language more and earlier than the dominant language, possibly because the more dominant language will have more stable lexical representations overall. The dominant language will also be affected, and as the disease progress, the bilinguals' languages will deteriorate following a similar pattern as in the less dominant language, which points towards a shared mental lexicon for the two languages.

The anomia observed in AD and PPA seems to be further affected by an impairment in access to the semantic system. Semantic variables such as imageability do affect naming and comprehension, but not as much as the lexical variables frequency and AoA. Future studies should include AoA as a variable when investigating naming deficits in AD and PPA, and the AoA of words from different word classes should be taken into account to investigate if the word class dissociations observed can be ascribed to differences in word class or in underlying variables.

Acknowledgements:

I wish to thank Minna Lehtonen, Monica Norvik, Hanne Gram Simonsen, Valantis Fyndanis, Marianne Lind and Roelien Bastiaanse for valuable feedback on earlier drafts of this paper, and Pernille Hansen for suggesting ideas for a title.

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References:

Albanese, E. (2007). The “hidden” semantic category dissociation in mild-moderate Alzheimer’s disease patients. *Neuropsychologia*, *45*(4), 639–643. https://doi.org/10.1016/j.neuropsychologia.2006.07.018

Altmann, L. J. P., & McClung, J. S. (2008). Effects of semantic impairment on language use in Alzheimer’s disease. *Seminars in Speech and Language*, *29*(1), 18–31. https://doi.org/10.1055/s-2008-1061622

Bastiaanse, R., Wieling, M., & Wolthuis, N. (2016). The role of frequency in the retrieval of nouns and verbs in aphasia. *Aphasiology*, *30*(11), 1221–1239. https://doi.org/10.1080/02687038.2015.1100709

Bayles, K. A., & Tomoeda, C. K. (2007). *Cognitive-Communication Disorders of Dementia* (1 edition). San Diego: Plural Pulishing Inc.

Beber, B. C., da Cruz, A. N., & Chaves, M. L. (2015). A behavioral study of the nature of verb production deficits in Alzheimer’s disease. *Brain and Language*, *149*, 128–134. https://doi.org/10.1016/j.bandl.2015.07.010

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: the importance of being imageable. *Objects vs. Actions and Nouns vs. Verbs*, *16*(2–3), 113–149. https://doi.org/10.1016/S0911-6044(02)00016-7

Bird, H., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2000). The Rise and Fall of Frequency and Imageability: Noun and Verb Production in Semantic Dementia. *Brain and Language*, *73*(1), 17–49. https://doi.org/10.1006/brln.2000.2293

Costa, A., Calabria, M., Marne, P., Hernández, M., Juncadella, M., Gascón-Bayarri, J., … Reñé, R. (2012). On the parallel deterioration of lexico-semantic processes in the bilinguals’ two languages: evidence from Alzheimer’s disease. *Neuropsychologia*, *50*(5), 740–753. https://doi.org/10.1016/j.neuropsychologia.2012.01.008

Cuetos, F., Gonzalez-Nosti, M., & Martínez, C. (2005). The picture-naming task in the analysis of cognitive deterioration in Alzheimer's disease. *Aphasiology*, *19*(6), 545–557. https://doi.org/10.1080/02687030544000010

Cuetos, F., Rodríguez-Ferreiro, J., Sage, K., & Ellis, A. W. (2012). A fresh look at the predictors of naming accuracy and errors in Alzheimer's disease. *Journal of Neuropsychology*, *6*(2), 242–256. https://doi.org/10.1111/j.1748-6653.2011.02025.x

Cuetos, F., Rosci, C., Laiacona, M., & Capitani, E. (2008). Different variables predict anomia in different subjects: a longitudinal study of two Alzheimer's patients. *Neuropsychologia*, *46*(1), 249–260. https://doi.org/10.1016/j.neuropsychologia.2007.07.011

Druks, J., Masterson, J., Kopelman, M., Clare, L., Rose, A., & Rai, G. (2006). Is action naming better preserved (than object naming) in Alzheimer's disease and why should we ask? *Brain and Language*, *98*(3), 332–340. https://doi.org/10.1016/j.bandl.2006.06.003

Druks, J., & Weekes, B. S. (2013). Parallel deterioration to language processing in a bilingual speaker. *Cognitive Neuropsychology*, *30*(7–8), 578–596. https://doi.org/10.1080/02643294.2014.882814

DSM-5. (2016). Retrieved April 18, 2017, from https://www.psychiatry.org/psychiatrists/practice/dsm

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: insights from connectionist networks. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(5), 1103–1123.

Ellis, Andrew. W. (2012). The Acquisition, Retention, and Loss of Vocabulary in Aphasia, Dementia, and Other Neuropsychological Conditions. In Faust, Miriam (Ed.), *The*

Handbook of the Neuropsychology of Language (Vol. 2). Retrieved from https://www.researchgate.net/publication/284480946_The_Acquisition_Retention_and_Loss_of_Vocabulary_in_Aphasia_Dementia_and_Other_Neuropsychological_Conditions

Ellis, H. D., & Young, A. W. (1977). Age-of-acquisition and recognition of nouns presented in the left and right visual fields: A failed hypothesis. *Neuropsychologia*, *15*(6), 825–828. https://doi.org/10.1016/0028-3932(77)90015-X

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.

Forbes-McKay, K. E., Ellis, A. W., Shanks, M. F., & Venneri, A. (2005). The age of acquisition of words produced in a semantic fluency task can reliably differentiate normal from pathological age related cognitive decline. *Neuropsychologia*, *43*(11), 1625–1632. https://doi.org/10.1016/j.neuropsychologia.2005.01.008

Foster, P. S., Drago, V., Yung, R. C., Pearson, J., Stringer, K., Giovannetti, T., … Heilman, K. M. (2013). Differential lexical and semantic spreading activation in Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, *28*(5), 501–507. https://doi.org/10.1177/1533317513494445

Fraser, K. C., Meltzer, J. A., Graham, N. L., Leonard, C., Hirst, G., Black, S. E., & Rochon, E. (2014). Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Language, Computers and Cognitive Neuroscience*, *55*, 43–60. https://doi.org/10.1016/j.cortex.2012.12.006

Glosser, G., & Friedman, R. B. (1991). Lexical but not semantic priming in Alzheimer's disease. *Psychology and Aging*, *6*(4), 522–527. https://doi.org/10.1037/0882-7974.6.4.522

Goldstein, F. C., Green, J., Presley, R., & Green, R. C. (1992). Dysnomia in Alzheimer's disease: An evaluation of neurobehavioral subtypes. *Brain and Language*, *43*(2), 308–322. https://doi.org/10.1016/0093-934X(92)90132-X

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., … Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Hansen, P. (2016). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language*, 0142723716679956. https://doi.org/10.1177/0142723716679956

Hernàndez, M., Caño, A., Costa, A., Sebastián-Gallés, N., Juncadella, M., & Gascón-Bayarri, J. (2008). Grammatical category-specific deficits in bilingual aphasia. *Brain and Language*, *107*(1), 68–80. https://doi.org/10.1016/j.bandl.2008.01.006

Hernández, M., Costa, A., Sebastián-Gallés, N., Juncadella, M., & Reñé, R. (2007). The organisation of nouns and verbs in bilingual speakers: A case of bilingual grammatical category-specific deficit. *Journal of Neurolinguistics*, *20*(4), 285–305. https://doi.org/10.1016/j.jneuroling.2006.10.002

Hillis, A. E., Heidler-Gary, J., Newhart, M., Chang, S., Ken, L., & Bak, T. H. (2006). Naming and comprehension in primary progressive aphasia: The influence of grammatical word class. *Aphasiology*, *20*(2–4), 246–256. https://doi.org/10.1080/02687030500473262

Hillis, Oh, S., & Ken, L. (2004). Deterioration of naming nouns versus verbs in primary progressive aphasia. *Annals of Neurology*, *55*(2), 268–275. https://doi.org/10.1002/ana.10812

Hillis, Tuffiash, E., & Caramazza, A. (2002). Modality-Specific Deterioration in Naming Verbs in Nonfluent Primary Progressive Aphasia. *Journal of Cognitive Neuroscience*, *14*(7), 1099–1108. https://doi.org/10.1162/089892902320474544

Hirsh, K. W., & Funnell, E. (1995). Those old, familiar things: age of acquisition, familiarity and lexical access in progressive aphasia. *Journal of Neurolinguistics*, *9*(1), 23–32. https://doi.org/10.1016/0911-6044(95)00003-8

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition*, *13*(7–8), 884–910. https://doi.org/10.1080/13506280544000093

ICD-10 Version:2016. (2016). Retrieved April 18, 2017, from http://apps.who.int/classifications/icd10/browse/2016/en

Ivanova, I., Salmon, D. P., & Gollan, T. H. (2013). The multilingual naming test in Alzheimer's disease: clues to the origin of naming impairments. *Journal of the International Neuropsychological Society: JINS*, *19*(3), 272–283. https://doi.org/10.1017/S1355617712001282

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, *131*(5), 684–712. https://doi.org/10.1037/0033-2909.131.5.684

Kambanaros, M., & Grohmann, K. K. (2015). Grammatical Class Effects Across Impaired Child and Adult Populations. *Frontiers in Psychology*, *6*, 1670. https://doi.org/10.3389/fpsyg.2015.01670

Kertesz, A. (2003). Anomia A2 - Aminoff, Michael J. In R. B. Daroff (Ed.), *Encyclopedia of the Neurological Sciences* (pp. 193–198). New York: Academic Press. https://doi.org/10.1016/B0-12-226870-9/00375-0

Kim, M., & Thompson, C. K. (2004). Verb deficits in Alzheimer's disease and agrammatism: Implications for lexical organization. *Brain and Language*, *88*(1), 1–20.

Kirshner, H. S. (2012). Aphasia A2 - Ramachandran, V.S. In *Encyclopedia of Human Behavior (Second Edition)* (pp. 177–186). San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-375000-6.00029-X

Kremin, H., Perrier, D., Wilde, M. D., Dordain, M., Bayon, A. L., Gatignol, P., … 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

Laiacona, M., Barbarotto, R., & Capitani, E. (1998). Semantic category dissociations in naming: is there a gender effect in Alzheimer's disease? *Neuropsychologia*, *36*(5), 407–419.

Lind, M., Simonsen, H. G., Ribu, I. S. B., Svendsen, B. A., Svennevig, J., & de Bot, K. (2017). Lexical access in a bilingual speaker with dementia: Changes over time. *Clinical Linguistics & Phonetics*, *32*(4), 353–377. https://doi.org/10.1080/02699206.2017.1381168

Łuniewska, M., Haman, E., Armon-Lotem, S., Etenkowski, B., Southwood, F., Anđelković, D., … Ünal-Logacev, Ö. (2016). Ratings of age of acquisition of 299 words across 25 languages: Is there a cross-linguistic order of words? *Behavior Research Methods*, *48*(3), 1154–1177. https://doi.org/10.3758/s13428-015-0636-6

Mack, J. E., Chandler, S. D., Meltzer-Asscher, A., Rogalski, E., Weintraub, S., Mesulam, M.-M., & Thompson, C. K. (2015). What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA? *Neuropsychologia*, *77*, 211–222. https://doi.org/10.1016/j.neuropsychologia.2015.08.019

Marcotte, K., Graham, N. L., Black, S. E., Tang-Wai, D., Chow, T. W., Freedman, M., … Leonard, C. (2014). Verb production in the nonfluent and semantic variants of primary progressive aphasia: the influence of lexical and semantic factors. *Cognitive Neuropsychology*, *31*(7–8), 565–583. https://doi.org/10.1080/02643294.2014.970154

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language*, *97*(3), 258–265. https://doi.org/10.1016/j.bandl.2005.11.001

Martinaud, O., Opolczynski, G., Gaillard, M.-J., & Hannequin, D. (2009). Relevant Category-Specific Effect on Naming in Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, *28*(5), 398–403. https://doi.org/10.1159/000256255

Matías-Guiu, J. A., Cuetos, F., Cabrera-Martín, M. N., Valles-Salgado, M., Moreno-Ramos, T., Carreras, J. L., & Matías-Guiu, J. (2017). Reading difficulties in primary progressive aphasia in a regular language-speaking cohort of patients. *Neuropsychologia*, *101*, 132–140. https://doi.org/10.1016/j.neuropsychologia.2017.05.018

McKhann, G. M., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*(7), 939–939. https://doi.org/10.1212/WNL.34.7.939

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., … Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(3), 263–269. https://doi.org/10.1016/j.jalz.2011.03.005

Mesulam, M. M. (2001). Primary progressive aphasia. *Annals of Neurology*, *49*(4), 425–432.

Moher, Liberati, Tetzlaff, Altmann, & The PRISMA Group. (2009). PRISMA. Retrieved November 14, 2016, from http://prisma-statement.org/PRISMAStatement/FlowDiagram.aspx

Montanes, P., Goldblum, M. C., & Boller, F. (1995). The naming impairment of living and nonliving items in Alzheimer's disease. *Journal of the International Neuropsychological Society: JINS*, *1*(1), 39–48.

Morris, R. (1996). *The cognitive neuropsychology of Alzheimer-type dementia*. Oxford University Press.

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*(6), 705–714. https://doi.org/10.3758/BF03202720

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281. https://doi.org/10.1080/17470216508416445

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1p2), 1. https://doi.org/10.1037/h0025327

Peters, F., Majerus, S., De Baerdemaeker, J., Salmon, E., & Collette, F. (2009). Impaired semantic knowledge underlies the reduced verbal short-term storage capacity in Alzheimer's disease. *Neuropsychologia*, *47*(14), 3067–3073. https://doi.org/10.1016/j.neuropsychologia.2009.07.002

Robinson, G., Rossor, M., & Cipolotti, L. (1999). Selective sparing of verb naming in a case of severe Alzheimer's disease. *Cortex*, *35*(3), 443–450.

Robinson, K. M., Grossman, M., White-Devine, T., & D'Esposito, M. (1996). Category-specific difficulty naming with verbs in Alzheimer's disease. *Neurology*, *47*(1), 178–182.

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

Rumiati, R. I., Foroni, F., Pergola, G., Rossi, P., & Silveri, M. C. (2016). Lexical-semantic deficits in processing food and non-food items. *Food for Thought: The Functional and Neural Mechanisms of Food Perception and Choice*, *110*, 120–130. https://doi.org/10.1016/j.bandc.2016.08.007

Sailor, K. M., Zimmerman, M. E., & Sanders, A. E. (2011). Differential impacts of age of acquisition on letter and semantic fluency in Alzheimer's disease patients and healthy older adults. *Quarterly Journal of Experimental Psychology (2006)*, *64*(12), 2383–2391. https://doi.org/10.1080/17470218.2011.596660

Sajjadi, S. A., Patterson, K., Tomek, M., & Nestor, P. J. (2012). Abnormalities of connected speech in the non-semantic variants of primary progressive aphasia. *Aphasiology*, *26*(10), 1219–1237. https://doi.org/10.1080/02687038.2012.710318

Semenza, C., Mondini, S., Borgo, F., Pasini, M., & Sgaramella, M. T. (2003). Proper names in patients with early Alzheimer's disease. *Neurocase, 9*(1), 63–69. https://doi.org/10.1076/neur.9.1.63.14370

Shuttleworth, E. C., & Huber, S. J. (1988). The naming disorder of dementia of Alzheimer type. *Brain and Language*, *34*(2), 222–234.

Silveri, M. C., Cappa, A., Mariotti, P., & Puopolo, M. (2002). Naming in Patients With Alzheimer's Disease: Influence of Age of Acquisition and Categorical Effects. *Journal of Clinical and Experimental Neuropsychology*, *24*(6), 755–764. https://doi.org/10.1076/jcen.24.6.755.8407

Silveri, M. C., & Ciccarelli, N. (2007). Naming of grammatical classes in frontotemporal dementias: linguistic and non linguistic factors contribute to noun-verb dissociation. *Behavioural Neurology*, *18*(4), 197–206.

Thompson, C. K., Ballard, K. J., Tait, M. E., Weintraub, S., & Mesulam, M. (1997). Patterns of language decline in non-fluent primary progressive aphasia. *Aphasiology*, *11*(4–5), 297–321.

Tippett, L. J., Meier, S. L., Blackwood, K., & Diaz-Asper, C. (2007). Category Specific Deficits in Alzheimer's Disease: Fact or Artefact? *Cortex*, *43*(7), 907–920. https://doi.org/10.1016/S0010-9452(08)70690-7

Ukita, H., Abe, K., & Yamada, J. (1999). Late Acquired Words in Childhood Are Lost Earlier in Primary Progressive Aphasia. *Brain and Language*, *70*(2), 205–219. https://doi.org/10.1006/brln.1999.2152

Venneri, A., McGeown, W. J., Hietanen, H. M., Guerrini, C., Ellis, A. W., & Shanks, M. F. (2008). The anatomical bases of semantic retrieval deficits in early Alzheimer's disease. *Neuropsychologia*, *46*(2), 497–510. https://doi.org/10.1016/j.neuropsychologia.2007.08.026

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, *35*(3), 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007

Vita, M. G., Marra, C., Spinelli, P., Caprara, A., Scaricamazza, E., Castelli, D., … Quaranta, D. (2014). Typicality of words produced on a semantic fluency task in amnesic mild cognitive impairment: linguistic analysis and risk of conversion to dementia. *Journal of Alzheimer's Disease: JAD*, *42*(4), 1171–1178. https://doi.org/10.3233/JAD-140570

Wallesch, C. W., & Hundsalz, A. (1994). Language Function in Delirium: A Comparison of Single Word Processing in Acute Confusional States and Probable Alzheimer′s Disease. *Brain and Language*, *46*(4), 592–606. https://doi.org/10.1006/brln.1994.1032

Weintraub, S., Rubin, N. P., & Mesulam, M.-M. (1990). Primary Progressive Aphasia: Longitudinal Course, Neuropsychological Profile, and Language Features. *Archives of Neurology*, *47*(12), 1329–1335. https://doi.org/10.1001/archneur.1990.00530120075013

White-Devine, T., Grossman, M., Robinson, K. M., Onishi, K., Biassou, N., & D'Esposito, M. (1996). Verb confrontation naming and word-picture matching in Alzheimer's disease. *Neuropsychology*, *10*(4), 495–503. https://doi.org/10.1037/0894-4105.10.4.495

Williamson, D. J. G., Adair, J. C., Raymer, A. M., & Heilman, K. M. (1998). Object and Action Naming in Alzheimer's Disease. *Cortex*, *34*(4), 601–610. https://doi.org/10.1016/S0010-9452(08)70517-3

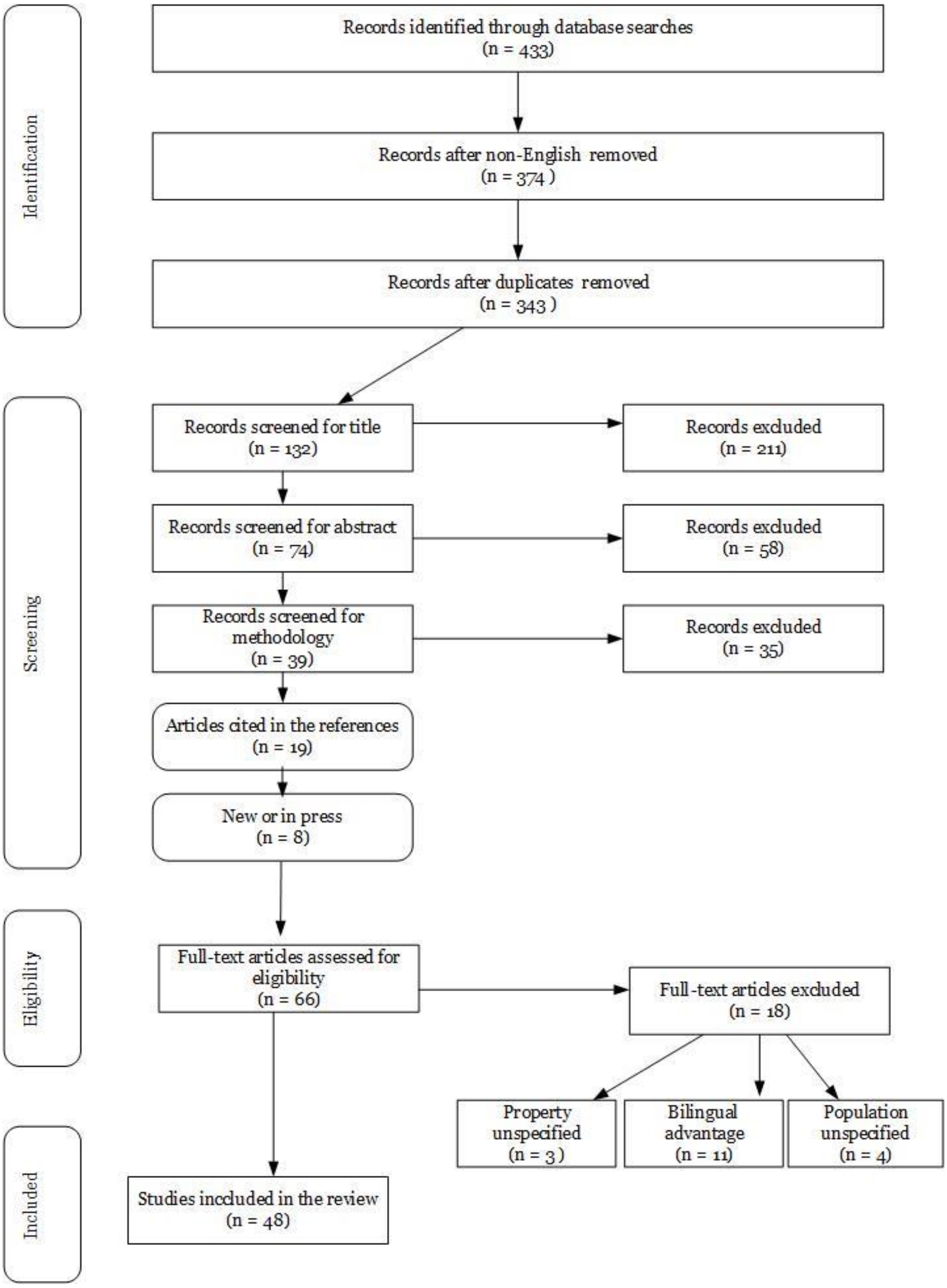


Figure 1. Overview of the study selection process

Ribu, I. S., Norvik, M. I., Lehtonen, M., & Simonsen, H. G. (submitted). Free word association in Alzheimer's disease and Primary progressive aphasia

II

Free word associations in Alzheimer's disease, Primary Progressive Aphasia and healthy aging

Ingeborg Sophie Ribu (i.s.b.ribu@iln.uio.no), Monica I. Norvik, Minna Lehtnonen, Hanne Gram Simonsen

Abstract

Background: Different kinds of tests to assess anomia can reveal different difficulties with lexical retrieval. Using multiple tests that tap different lexical processing mechanisms can give a nuanced picture of the lexical retrieval problems in Alzheimer's disease (AD) and primary progressive aphasia (PPA). There are a few studies that use free word associations (FWA) on participants with AD, but so far none on PPA.

Aims: In this study, we investigate whether free word associations are useful as a supplement to traditional language assessment tasks to assess language impairment in AD and PPA.

Methods & procedures: We administered tests for semantic knowledge, picture-naming, verbal fluency an a FWA test to seven participants with AD, two with PPA and 29 neurologically healthy controls. The responses from all participants were analysed into different response categories depending on their relation to the cue words.

Outcome & results: Response types to cue words differed between participants with AD, PPA and neurologically healthy participants. PPA participants gave more null-responses than AD and control participants, and AD participants gave more collocation responses than participants with PPA and controls.

Conclusions: Responses from FWA tests capture differences between persons with AD and healthy controls that traditional tests do not, indicating that FWA tests can be a useful supplement to more traditional language tests when assessing language impairment in dementia.

Key words: Free word association; Alzheimer's disease; primary progressive aphasia; mental lexicon; lexical retrieval;

Introduction

Language skills are generally well preserved for healthy speakers across the life span (Goral, 2004; Wingfield & Stine-Morrow, 2000), however, lexical access during word production seems to present difficulties for older adults, in particular in experimental settings (Chertkow, Whatmough, Saumier, & Duong, 2008; Cuetos, Arce, Martínez, & Ellis, 2017; Cuetos, Rodríguez-Ferreiro, Sage, & Ellis, 2012). These retrieval difficulties have been defined as reflecting either an underlying semantic impairment, or an impairment in access to the lexical items (Balthazar, Cendes, & Damasceno, 2008; Kempler & Goral, 2008). Lexical access deficits are more pronounced for persons with Alzheimer's disease (AD) and primary progressive aphasia (PPA) than for healthy older adults (Azuma, Sabbagh, & Connor, 2013; Bowles, Obler, & Albert, 1987; Cuetos et al., 2012; Gorno-Tempini et al., 2011). These findings are often based on assessing language skills by means of picture-naming and picture identification tests. Different tests can yield different results even though they assess the same domain (i.e. lexical retrieval). The aim of this study is to investigate how free word associations (FWA) can be used as a supplement to more conventional tests to assess language impairment in AD and PPA.

Alzheimer's disease and primary progressive aphasia

Both AD and PPA are neurodegenerative diseases that affect several cognitive domains, including language. The main initial symptom of AD involves decline in episodic memory, but language impairments, including word-finding difficulties (anomia), are also reported as early symptoms, and lexical access can be impaired even in mild cases (Bayles & Tomoeda, 2007; McKhann et al., 2011). The language deficits in AD are fundamentally semantic, not orthographic or lexical (Barresi, Nicholas, Connor, Obler, & Albert, 2000; Hodges, Patterson, Graham, & Dawson, 1996; Obler & Gjerlow, 1999).

The core feature of PPA is an isolated and progressive decline in language functions in the initial couple of years of the disease (Gorno-Tempini et al., 2011). Three sub-types of PPA are identified, where the logopenic variant of PPA (lvPPA) is the one that is most similar to AD pathology, and has in some cases been described as an atypical variant of AD (Henry & Gorno-Tempini, 2010). lvPPA is characterized by impaired sentence repetition and impaired single-word retrieval in both spontaneous speech and confrontation naming. Single-word comprehension and object knowledge are intact (Gorno-Tempini et al., 2011). A semantic variant (svPPA) and a non-fluent agrammatic variant (nfavPPA) have also been described. The first one is recognised by impaired confrontation naming and single-word comprehension, but spared repetition and speech production; the latter by agrammatism, effortful, haltering speech, impaired comprehension of syntactically complex sentences, but spared single-word comprehension and object knowledge (Gorno-Tempini et al., 2011).

The mental lexicon

The mental lexicon can be seen as an organized network of semantic, phonological, orthographic, morphological and other kinds of linguistic information (Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017). These networks grow continuously through the lifespan, giving older adults larger lexica than younger adults and children. In spite of their larger lexicon, older adults show a general slowing on lexical access tasks. This can be attributed to a general slowing of cognitive processing (Salthouse, 1996), or because the search processes take longer when the lexicon is larger (Vitevitch & Luce, 1998). Including both younger and older participants in the current study can help shed light on age-related changes in lexical access.

Most models of language processing presuppose two levels of processing; one semantic level for word meaning, and one phonological level for word form. Different lexical retrieval models vary in the explanation of the relationship between these levels; whether there is interactive activation between form and meaning, or one-way from form to meaning or vice-versa.

According to one theory of lexical access, the *Transmission Deficit Hypothesis* (TDH), aging weakens the connections between the semantic and the phonological representations of words. Weak connections between the representations transmit too little excitation for a given representation to reach the threshold for activation, which leads to production failure (Burke & Shafto, 2004, 2008; MacKay & Burke, 1990). The TDH assumes that there are multiple links between semantic nodes in the network, but only one link between concepts and phonological form, which makes phonological forms more vulnerable to breakdown when the connections between nodes weaken. According to TDH, semantic skills remain intact in older age, but production is affected due to the weaker connections between the conceptual word form and the phonological form. This may result in multi-word responses, mainly descriptions of the target, when participants have been instructed to respond with single-words, as the links to the precise word forms are weakened or lost (Burke & Shafto, 2008).

Another model, the *Inhibition Deficit Hypothesis* (IDH), suggest that aging weakens the inhibitory processes associated with task-irrelevant information (Zacks & Hasher, 1997). This leads to an inability to suppress competing lexical items during lexical access. This could also explain why older adults and persons with dementia provide more multi-word responses, often unrelated to the target, on naming tasks, as they are unable to inhibit irrelevant information and it is difficult to suppress thoughts that digress from the task at hand. This would result in production of personal and unrelated information in their responses.

Assessment of lexical retrieval

There are several ways to assess lexical retrieval. In this paper we discuss how three levels of test restriction, *bound*, *controlled*, and *free* tests, affect lexical access. Traditionally, lexical retrieval is assessed with confrontation naming tests for production (such as the Boston

Naming Test (BNT): (Kaplan, Goodglass, Weintraub, & Goodglass, 1983)), or tests for semantic knowledge for comprehension (e.g. the Pyramids and Palm Trees Test (PPT): (Howard, Patterson, & Thames Valley Test Company, 1992)). These kinds of tests are considered to be *bound*, or restricted, as only one specific response is considered correct. Lower scores on such tests for persons with AD and PPA can indicate that access to the precise items under investigation may be impaired.

Another common test for lexical retrieval is verbal fluency (VF), where participants are instructed to name as many words they can, belonging to a given phonological or semantic category within a specific time slot. VF tests are less bound, but still *controlled* as restrictions are imposed on the participant. These restrictions may increase retrieval difficulties for persons with a language impairment.

Persons with AD and PPA are impaired on such tasks compared to neurologically healthy persons; for instance, they have longer response latencies and fewer correct responses on confrontation naming, and fewer items in total on verbal fluency tests tasks than neurologically healthy persons (Barresi et al., 2000; Chertkow & Bub, 1990; J. D. Henry, Crawford, & Phillips, 2004; Kempler & Goral, 2008; Nicholas, Obler, Albert, & Goodglass, 1985).

For a deeper insight into the organization of concepts in the mental lexicon, we introduce a *free* word association (FWA) test, since FWA tests are considered to tap into the structuring of the mental lexicon, rather than providing a reflection of normal language production (Deyne & Storms, 2015). Responses to cues on FWA tests mirror the links between concepts in the mental lexicon, whether the relationship is based on semantics, phonology, or other linguistic information. Free naming tasks can trigger responses that show an understanding of the tested item, which are not dichotomously scored as correct or incorrect as in bound tasks.

The current study employed these three levels of restriction of tests for semantic memory and knowledge. Including both comprehension and production tests, especially tasks that target the same items, may help persons with dementia show whether they have knowledge of the concepts although the precise lexical items may be difficult to activate in a bound production task.

Free word associations

Free word association data (FWA) differ from other kinds of linguistic data as they are freely produced, yet consist of single items for analysis. Participants are required to respond with the first word that comes to mind when presented with a cue word, which may reveal which words that are more closely connected to each other in the lexicon. Nodes in the lexicon are considered to be linked on different levels (phonological, semantic etc.), and one can expect responses to a cue on FWA tasks to originate in different levels too.

The inclusion of FWA in addition to more traditional tests for naming and comprehension has both theoretical and clinical implications. Analyses of results from different kinds of tests can inform on theories of structuring of the mental lexicon and can give a more nuanced picture of semantic impairment and decline in AD and PPA. The response patterns from these tests may reveal something about the underlying problems of anomia for these participant groups; if participants provide predominantly phonologically related responses to cues, this may indicate that the phonological connections are stronger than the semantic ones, and vice versa. FWA can thus be used to supplement traditional naming tasks, as responses consist of single items that are produced without any restrictions. This provides insight into processes involved in lexical organisation that goes beyond the information that is captured in spoken language production (Deyne & Storms, 2015).

The clinical implications include the possibility for providing earlier dementia diagnosis based on speech output, and a possible differentiation between AD and PPA, if the response patterns from different participant groups significantly differ from each other. FWA tests have traditionally been scored looking at either the typicality of the responses provided (i.e. responses that are similar across a larger group of participants), or by analysing the responses as being either paradigmatic or syntagmatic in relation to the cue. Paradigmatic responses come from the same word class e.g. '*dog*' as a response to '*cat*', whereas syntagmatic responses do not belong to the same grammatical class, but often co-occur in normal speech; e.g. '*book*' as response to '*read*'. Previous studies of FWA in persons with AD show that they give less typical responses to cue words (Eustache, Cox, Brandt, Lechevalier, & Pons, 1990; Gollan, Salmon, & Paxton, 2006) and more multi-word responses despite instructions to reply with a single word (Santo Pietro & Goldfarb, 1985). The number of paradigmatic responses has been seen to decrease as a function of AD severity (Gewirth, Shindler, & Hier, 1984). This two-way scoring system does not capture the finer relationship between cue words and responses, and we propose an alternative scoring system in the Analysis section below.

To our knowledge, there are no studies that employ FWA tests in the study of lexical retrieval impairments in PPA. In the current study we include two participants with the logopenic variant of PPA to see if their responses are markedly different from persons with AD and healthy controls.

Research questions and hypotheses

We investigated how patterns of word association responses can inform on the semantic impairments in AD and PPA. Furthermore, we employed different tests for assessment of naming and semantic knowledge to get insight into the mechanisms that affect anomia. We wanted to investigate if the inclusion of FWA can help us answer these questions:

1. How do results from FWA, in combination with traditional tasks for lexical access, broaden our knowledge about the semantic impairment in AD and PPA?

2. How can responses from FWA help inform on theories of organisation of the mental lexicon?
3. Can response patterns from FWA tests be diagnostically relevant for early identification of AD and PPA?

We hypothesised that participants with AD and PPA in general would perform more poorly than healthy controls on traditional tasks for lexical retrieval. If persons with AD and PPA have impaired access to specific items in production, but spared semantic representations, they should still be able to recognise items in comprehension tasks, even if they cannot name the items in production tasks.

In contrast, if the lexical retrieval impairments observed in AD and PPA are due to underlying semantic deficits, we would see responses that show a lack of understanding of the tested items, also on the comprehension tasks.

Descriptions and semantically related responses on the production tasks reflect an impairment in lexical access instead of the semantic system per se. Similarly, multi-word responses on the FWA task can also be indicative of difficulties with access rather than an underlying semantic deficit.

Furthermore, the results from the FWA task will differ from the results on the controlled and bound tasks, as no restrictions are imposed on the participant, and responses are not scored as correct or incorrect but rather seen in relation to the cue. The responses may stem from a larger array of categories in relation to the cue word than the words generated on the controlled and bound tests.

The current paper reports on results from participants with AD, PPA and healthy controls on different tests for lexical knowledge; bound tests for comprehension and production, controlled production, and free word associations.

Methods

Participants

Seven persons with AD, two with lvPPA and 29 neurologically healthy controls took part in this study. The participants with AD and PPA were recruited from the memory clinic at Oslo University Hospital. One participant with PPA was initially diagnosed with AD, but the diagnosis was later changed to lvPPA. Diagnoses for all participants are based on scores on several standardised tests, and supported by physiological findings such as biomarkers in cerebrospinal fluid and MRI. Persons with other neurological disorders or diseases (history of strokes, epilepsy, and other dementias than AD or PPA, etc.), premorbid language disorders, or persons who had a history of drug or alcohol abuse were excluded from participation.

Fourteen neurologically healthy, elderly control participants (HCE), and 15 healthy younger controls (HCY) were also included in the study. The HCE group was recruited through personal networks and online information posts. The younger control group came mainly from the student body at the University of Oslo. Participant demographics are summarised in Table 1.

Participants with AD, PPA and HCE did not differ significantly in terms of age on a Wilcoxon sum rank test. There HCE group had slightly higher education than the participants with AD ($w = 22.5$, $p = 0.051$), but there were no significant differences between the other groups.

	AD (N = 7)		**PPA** (N=2)	**HCE** (N = 14)		**HCY** (N = 15)	
Female : Male	5 : 2		0 : 2	9 : 5		12 : 3	
	Mean	*SD*		*Mean*	*SD*	*Mean*	*SD*
Age	74,85	4,3	75; 81	73,14	6,3	24,8	2,9
Education (years)	14,28	2,49	15; 14	16,6	2,37	16,3	1,84
MMSE	27,28	2,18	24; 29	29,14	0,63	29,06	0,99

Table 1: Demographics of participants included in the study, the age, years of education and MMSE scores are given for both the participants with PPA, whereas means and standard deviations are given for the other participant groups.

All participants were functionally monolingual, using Norwegian only in their daily lives. However, all participants reported exposure to at least one other language on a regular basis (mainly English through TV/radio, or literature).

The study was approved by the Regional committees for medical and health research ethics, and the data protection officials at the University of Oslo and Oslo University Hospital, Norway. All participants gave their written informed consent. Participants with AD and PPA were all considered by their primary doctors to be able to give consent, so no additional consent was needed from guardians or next-of-kin.

Materials and procedures

Cognitive test battery

All participants went through a neuropsychological screening with standardised tests for general cognitive functioning, in addition to a language battery presented in the next section. The cognitive test battery included the *Mini-Mental State examination* (MMSE) (Folstein, Folstein, & McHugh, 1975), the word-list learning, recall and recognition, and figure copying and recall from the *Consortium to Establish a Registry for Alzheimer's Disease* (CERAD, (Morris et al., 1989)) for verbal episodic memory and visuo-constructional skills and memory. Executive functioning was assessed with *Trail Making Test A and B* (TMT (Tombaugh, 2004)) and the *Stroop task* (Delis, Kaplan, & Kramer, 2001). Finally, working memory was assessed with three digit span tasks; digit forward and backward from *WAIS-IV* (Wechsler, 2008) and digit ordering (MacDonald, Almor, Henderson, Kempler, & Andersen, 2001).

Linguistic test battery

The linguistic test battery included tests with different level of restriction, *bound* tests for semantic memory (*Pyramids and Palm Trees* (Howard et al., 1992) and *De Semantische Associatie Test* (Visch-Brink, Stronks, & Denes, 2005)) in addition to word-picture matching and picture-naming; *controlled* verbal fluency and *free* word association tasks. Each test is presented in the following sections.

The Pyramids and Palm Trees Test (PPT) and the Semantische Associatie Test (SAT) are bound tasks used to assess semantic comprehension and knowledge for objects and actions respectively. PPT consists of 52 items where participants have to match one of two competing object drawings to a third, based on their knowledge of how things occur together; e.g. '*glasses*' should be matched to '*eyes*' and not to '*ears*'. The same principle is employed in SAT, but for actions rather than objects; '*playing piano*' should for example be matched to '*playing drums*', and not to '*baking*' or '*working on a computer*'. The SAT consists of 20 items with three competing drawings. The link between the cue and target item is only apparent if the participant is able to correctly access complete semantic information about all test items.

To assess lexical production and comprehension, a picture-naming test and a word-picture matching test were created specifically for this study. The tests include 30 nouns and 30 verbs, and were constructed using Norwegian norms for frequency, age of acquisition and imageability – variables that are known to affect lexical retrieval for persons with AD and PPA (Almor et al., 2009; Bird, Lambon Ralph, Patterson, & Hodges, 2000; Cuetos et al., 2012; Ribu, under revision). Twenty adult first-language speakers of Norwegian participated in a norming study for these tests.

Controlled production was assessed by means of both semantic and phonological verbal fluency. The participants were instructed to name as many words they could, belonging to a specific category (*animals*) or starting with a certain sound (/f/, /a/, /s/), in 60 seconds. These tests are called *controlled associations*, as both time and semantic or phonological restrictions are imposed on the participant.

A free word association test consisting of 30 items (18 nouns and 12 verbs) was included to tap semantic processing (Bøyum, Hansen, & Ribu, in preparation). The test administrator read a cue word from a word-list and instructed the participants to respond with the first word that came to mind. No restrictions were given with regard to time, and participants were instructed to respond with one word only. If participants did not provide single-word responses, they were reminded to do so for the remaining trials. These instructions were repeated after every multi-word response.

Participants were tested individually either at home (four participants with AD) or at the University of Oslo (all other participants). One participant with AD was tested on two separate days, but for all other participants testing was conducted in one session of approximately two and a half to three hours. All tests were pen-and-paper based, and the whole session was audio recorded.

Participants with AD or PPA were assessed with MMSE, CERAD and TMT A+B at the memory clinic at Oslo University hospital as part of their dementia evaluation. The remaining cognitive and all language tests were conducted at the University of Oslo.

Analysis

Bound tests

Results from the PPT/SAT and the word-picture matching tasks were analysed by accuracy, whereas an error analysis was implemented for the picture-naming task where non-target responses were analysed into nine response categories:[1] 1) synonyms, 2) hyponyms, 3) hypernyms, 4) description, 5) code switch, 6) wrong word class that still includes the target word, 7) visual confusion, 8) semantic confusion, and 9) no answer. Responses from

[1] The categories were constructed based on responses provided by the participants from the norming sample.

categories 1-6 were scored as correct. Examples of responses per category are found in Appendix 1.

Controlled fluency

We counted the number of words generated on each of the sub-tests of the verbal fluency test. Correct responses for the phonological fluency task were words produced beginning with the specific sound /f/, /a/ or /s/, excluding proper nouns, numbers, repetitions and inflections of the same word. The total score is the sum of the words generated across the three trials. For semantic fluency, correct answers were types of animals produced, excluding repetitions and plurals.

Free word association

As mentioned previously, responses on FWA tasks have traditionally been coded as syntagmatic or paradigmatic (Gewirth et al., 1984) in relation to the cue word. In some cases, this two-way distinction is not always sufficient. For instance, '*life*' is often followed by '*death*' in normal speech, and can be judged as syntagmatic based on that criterion. However, the words belong to the same word class and could also be scored as paradigmatic. To account for such and other examples a more nuanced scoring system with a total of 14 categories has been proposed (Fitzpatrick, Playfoot, Wray, & Wright, 2015). Employing several categories allow for studying the finer details in responses between groups, as the traditional two-way scoring system is not sensitive enough to capture differences in responses, and does not properly account for responses that have several different relations to the cue word.

The 14 categories introduced by Fitzpatrick and colleagues (2015) have both strengths and weaknesses. Several categories allow for a more detailed analysis of the relationship between cues and responses; however, in some cases the number for each type of response is not large enough for meaningful analysis. Furthermore, their system recognizes that responses can be both *semantically* related to the cue and a *collocation*, but a category for responses that are both *meaning*- and *form*-based is lacking. Based on these shortcomings, Bøyum and colleagues proposed a scoring system that makes use of five main categories, of which three have sub-categories (Bøyum et al., in preparation). This system is employed in the current study. The categories and sub-categories are: 1) Meaning-based responses; 1a) synonyms, 1b) other semantic relations; 2) Collocations; 3) Form-based responses; 4) Mixed responses; 4a) meaning- and form-based, 4b) meaning-based and collocations and 4c) two-step associations – responses that are linked to the cue via a second word, and 5) No association; 5a) descriptions (whole phrases), 5b) no answer, 5c) seemingly unrelated responses, 5d) personal responses. Examples of each response type are found in Appendix 2.

All statistical calculations were performed using R version 3.6.0 (R core team, 2019) and RStudio version (1.2.1335) (RStudio Team, 2015). We use the Wilcoxon sum rank test to test the difference between participants and groups within each test. This test is less likely to indicate significance because of outliers, as it is based on the sums of ranks in the observed data rather than means. Non-parametric tests such as the Wilcoxon sum rank test are preferred for small sample sizes (Happ, Bathke, & Brunner, 2019).

Results

We first give a short description of the results from the cognitive tests as a background, before we go in depth on the linguistic tests. Participants with AD and PPA are always compared to the averages of the scores obtained by the two groups of control participants.

Cognitive tests

Most of the participants with AD and PPA scored relatively well on the Mini-Mental State Examination (MMSE), only 3 scoring below the cut-off (26/30). However, recall of word-lists and figures, as well as word-list recognition from the consortium to establish a register for

Alzheimer's disease (CERAD), was poor for most participants, except PPA02 whose scores on these tests were higher than for the other participants with dementia, but poorer than for controls. AD05 had among the lowest scores on MMSE and all sub-test from CERAD, yet scored relatively well on Trail-Making Test (TMT), Stroop and the digit span tasks. AD01 showed the opposite pattern, and scored high on MMSE and the sub-tests from CERAD, but poorer on TMT and Stroop – especially on the sub-tests that measure switching and inhibition. Note also AD06 who scored within normal limits on all cognitive tests except word-list recall and recognition, and figure recall. These observations show how different tasks tap into different strengths and that multiple tests are necessary to assess cognitive impairment.

Table 2 below shows the scores of the participants with AD and PPA on the neuropsychological test battery compared to healthy elderly controls (HCE) and healthy younger controls (HCY).

	AD01	AD02	AD03	AD04	AD05	AD06	AD07	PPA01	PPA02	HCE (Average)	HCY (Average)
Age/sex **(Span)**	72/f	78/f	76/f	76/m	79/f	66/m	77/f	75/m	81/m	73,14 (66-89)	24,8 (19-30)
Education **(years)**	14	16	13	18	14	15	10	15	14	16,06 (15-20)	16,3 (13-20)
MMSE	29	27	29	30	23	27	24	24	29	29,14	29,06
CERAD 10 words **(max. 30)**	15	11	10	17	10	19	18	15	18	22,43	24,6
CERAD 10 words Recall **(max. 10)**	0	2	1	4	1	3	1	4	4	6,58	8,67
CERAD 10 words Recognition **(max. 20)**	8	7	9	N/A	6	7	6	7	10	19,5	19,8
CERAD Figure copy **(max. 11)**	11	11	9	N/A	9	11	8	N/A	10	10,57	10,13
CERAD Figure recall **(max. 11)**	5	0	0	N/A	N/A	N/A	6	N/A	8	10,14	10
TMT-A	83 sec.	75 sec.	33 sec.	45 sec.	28 sec.	37 sec.	61 sec.	85 sec.	105 sec.	37,71 sec.	29,53 sec.
TMT-B	207 sec.	300 sec.	82 sec.	144 sec.	78 sec.	69 sec.	236 sec.	300 sec.	486 sec.	82,92 sec.	63,73 sec.
TMT cost **(B-A)**	124 sec.	225 sec.	49 sec.	99 sec.	50 sec.	32 sec.	175 sec.	215 sec.	381 sec.	45,21 sec.	34,2 sec.
Stroop **(word reading)**	58 sec.	N/A	51 sec.	41 sec.	60 sec.	53 sec.	95 sec.	90 sec.	N/A	41,28 sec.	52,8 sec.
Stroop **(color-word interference)**	225 sec.	N/A	115 sec.	143 sec.	128 sec.	98 sec.	225 sec.	200 sec.	N/A	98,78 sec.	106, 5 sec.
Stroop cost **(CWI-WR)**	167 sec.	N/A	64 sec.	102 sec.	68 sec.	45 sec.	130 sec.	110 sec.	N/A	57,5 sec.	53,73 sec.
Digit forward **(max. 16)**	10	7	8	7	8	10	9	8	6	9,71	10
Digit backward **(max. 14)**	5	5	5	5	9	7	4	2	3	7,64	7,13
Digit ordering **(max. 15)**	11	3	9	12	11	15	11	6	6	13,35	12,6

Table 2: Results on the neuropsychological test battery for participants with AD, PPA and neurologically healthy participants. Some scores on the CERAD sub-tests are unavailable for participants AD04, AD05, AD06 and PPA01. AD02 and PPA02 were unable to perform the Stroop task.

Linguistic test battery

The results from the linguistic test battery are summarised here before the results from each sub-test is discussed in more detail. Participants with AD and PPA are compared to the averages from the HCE and HCY groups. Table 3 gives an overview of the results from each test.

Overall, persons with AD and PPA provided fewer correct items on both the *Pyramids and Palm Trees test* (PPT) and the *Semantische Associatie Test* (SAT), the word-picture-matching task, and produced far fewer items in total on *verbal fluency* (VF) compared to controls. On the picture-naming tasks, participants with AD and PPA gave fewer target responses than controls, and they provided more multi-word responses to cues and less meaning-based responses than healthy participants on the FWA test.

	AD 01	AD 02	AD 03	AD 04	AD 05	AD 06	AD 07	PPA 01	PPA 02	HCE (Mean)	HCY (Mean)	Compared to normed data
Age/sex **(Span)**	72/f	78/f	76/f	76/m	79/f	66/m	77/f	75/m	81/m	73,14 (66-89)	24,8 (19-30)	
Pyramids and palm trees **(no. items 52)**	92,3	82,7	92,3	92,3	82,7	100	86,5	92,3	86,5	99,2	97,7	98,7
Semantic association test **(no. items 20)**	90	55	80	90	85	95	65	55	85	96,4	92,6	N/A
Word-picture matching **(no. items 60)**	100	98,3	100	100	98,3	100	96,6	100	95	99,9	100	100
Noun-naming **(no. items 30)**	80	96,66	93,33	83,33	90	100	86,66	83,33	73,33	96,9	97,1	98,66
Verb-naming **(no. items 30)**	93,33	86,66	93,33	100	90	100	90	93,1*	73,33	98	97,5	98,16
Total phonological fluency (3 min)	43	51	38	44	25	44	38	30	18	58,6	50,6	46,0
Verbal fluency animals	14	9	19	15	8	22	12	7	5	25,3	26,2	23,5

Table 3: Results from language tasks per participant with AD and PPA, and the neurologically healthy groups. Total number of items produced on the verbal fluency tasks, all other scores in percentages.
** One item is missing from PPA01's results on the verb-naming task due to technical difficulties with the audio recording.*

Bound tasks

PPT and SAT

For the bound semantic association tasks, the scores vary between the tests. On the PPT significant differences were found between AD and HCE (w = 9.5, p = 0.001), AD and HCY (w = 16.5, p = 0.008), and interestingly also between the HCE and HCY groups, with younger controls scoring lower than the older controls (w = 59, p = 0.028). PPA01's scores were marginally better than PPA02's scores. Both scored lower than the control groups, but on par with the participants with AD.

For the SAT significant differences were found between the AD and HCE ($w = 6$, $p < 0.001$), and between AD and HCY participants ($w = 17$, $p = 0.011$). The participants with PPA showed an opposite pattern than for the PPT, with PPA02 scoring higher than PPA01, and both yet again poorer than controls, but similarly to the participants with AD.

Participants with AD, as well as controls scored better on PPT (objects) than on SAT (actions) (AD: $w = 49$, $p = 0.002$; HCE: $w = 196$, $p < 0.001$; HCY: $w = 81$, $p < 0.001$), indicating a better performance with nouns than with verbs. There was no difference between the two tests for PPA02, but PPA01's scores were better for the PPT than for the SAT. The results from these tests are visualised in Figure 1. Note that the maximum score is different for the two tests.

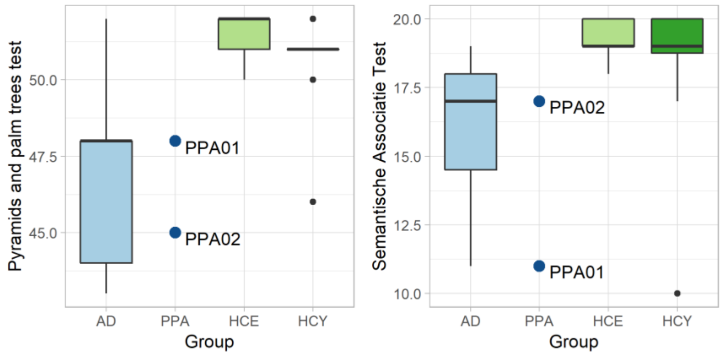


Figure 1: PPT (left) and SAT (right) responses per participant group (PPT max. 52, SAT max. 20). Participants with AD and PPA do not perform very differently from each other on the bound association tasks, but are impaired compared to HCE and HCY participants. Note the different maximum scores on the two tests.

Word-picture matching task

Most participants made no mistakes on the word-picture matching task, but a few errors were recorded for the AD and HCE groups and for one participant with PPA, enough to find a statistically significant difference between the AD and HCY participants ($w = 30$, $p = 0.009$).

PPA01 scored 100% on this task, whereas it was a more difficult task for PPA02 who made three mistakes. Figure 2 shows the responses on the word-picture-matching task.

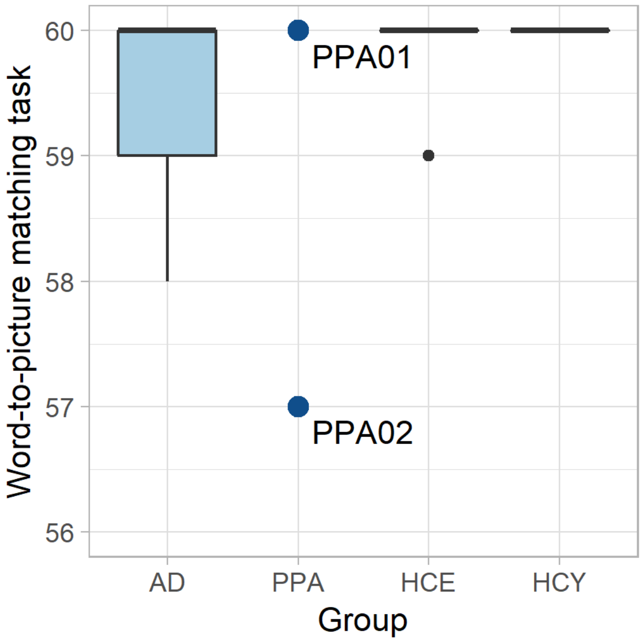


Figure 2: Word-picture-matching – participants with AD and PPA are slightly worse in connecting pictures to corresponding words than persons without dementia. However, the variation between the participants is great, and the number of mistakes is low for all participants.

Picture-naming

On the noun-naming task there were significant differences between participants with AD and both control groups (AD vs. HCE: $w = 20$, $p = 0.028$; AD vs. HCY: $w = 20.5$, $p = 0.02$). PPA02 made more noun-naming mistakes than PPA01, and both scored lower than the participants with AD and controls.

On the verb-naming task, there were no significant differences between any groups, suggesting that this sub-test was not sensitive enough to pick up on differences in naming performance. There were also no differences between the groups' performances on the verb- and the noun-naming tasks. PPA01's scores were again better than PPA02's scores, who scored equally on the noun- and verb-naming sub-tests. Both participants with PPA scored lower than the other participant groups.

In sum, AD and PPA participants had more difficulties than the age-matched controls on verb-naming as well as on noun-naming. However, as there was no difference between the scores on the noun– and verb– naming tests, it is not possible to determine a dissociation in noun– and verb– production. One of the participants with AD made mistakes on the same items in the matching and the naming task, indicating a difficulty with certain items from the test. The results from the naming tests are visualised in Figure 3 below.

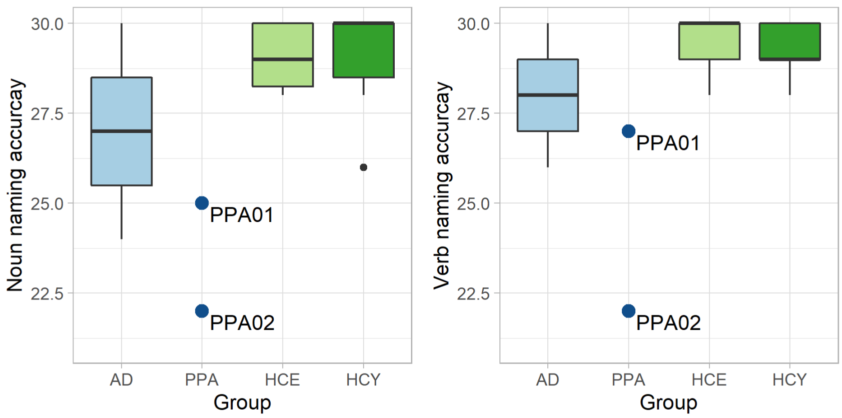


Figure 3: Scores per participant group on the picture-naming tasks. Participants with AD are marginally better on the verb-naming task than the noun-naming task, but this is not significant. The participants with PPA show a pattern where PPA01 is better with noun-naming than PPA02, who again is better on verb-naming than PPA01.

An error analysis of the responses from the picture-naming tasks reveals that AD and PPA participants provided fewer target responses than controls; instead, they provided responses that were semantically related to the target. For both control groups the most dominant response types after target responses were synonyms, whereas persons with AD provided either descriptions or other words entirely (semantic confusion). Hyper- and hyponyms were more common for AD and PPA participants than for the control groups. Participant AD06 provided more target responses than the averages for both HCE and HCY groups. For those with PPA, the most dominant response after target was description of the item, semantic confusion or no answer, but there was great variation between the two participants. Table 4 gives an overview of the percentage of responses for each response type per participant with AD and PPA, and averages per group for HCE and HCY.

	AD 01	**AD 02**	**AD 03**	**AD 04**	**AD 05**	**AD 06**	**AD 07**	**PPA 01***	**PPA 02**	**HCE (Mean)**	**HCY (Mean)**
Target	76,66	80	86,66	83,33	66,66	93,33	71,66	67,79	55	92,66	92,33
Synonym	5	3,33	0	1,66	3,33	1,66	1,66	1,69	1,66	2,61	2,91
Hyponym	1,66	0	3,33	1,66	3,33	5	5	1,69	0	1,53	1,25
Hypernym	1,66	1,66	1,66	0	1,66	0	1,66	3,38	5	0	0,55
Description	1,66	6,66	0	5	0	0	5	13,55	5	0,11	0,13
Code switch	0	0	0	0	0	0	0	0	0	0,23	0,13
Wrong word class	0	0	1,66	0	15	0	3,33	0	6,66	0,23	0
Visual confusion	1,66	1,66	1,66	0	5	0	1,66	3,38	8,33	0,58	1,1
Semantic confusion	11,66	6,66	5	3,33	3,33	0	10	6,77	10	1,9	1,51
No answer	0	0	0	5	1,66	0	0	1,69	8,33	0	0

*Table 4: Percentage of responses in each response category for both the noun and the verb-naming tasks combined. *PPA01's total score is out of 59, as one item is missing from the verb-naming task.*

Controlled test: verbal fluency

On the phonological fluency test, there was a difference in number of total items produced for AD and HCE ($w = 14$, $p = 0.009$), and AD and HCY participants (AD vs. HCY: $w = 18.5$, $p = 0.018$), as expected. There was no difference between the control groups. Again, there was a difference between PPA01 and PPA02, with PPA01 producing more words on the phonological fluency task than PPA02.

Similarly, differences between AD and HCE ($w = 7$, $p = 0.001$) and AD and HCY ($w = 5.5$, $p = 0.001$), and no difference between the control groups were also found for animal fluency. Both participants with PPA scored lower on animal fluency compared to persons with AD and controls. Figure 4 shows the distribution of items on each sub-test of both phonological and semantic fluency.

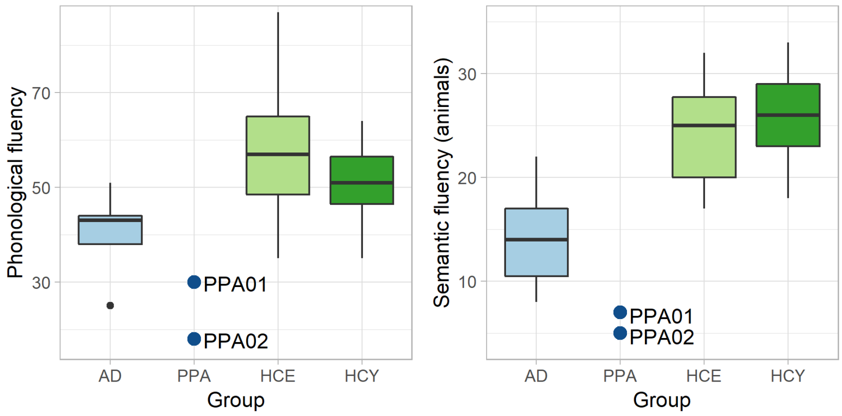


Figure 4: Distribution of responses on phonological (F-A-S) and semantic (animal) fluency for all participant groups. In all conditions participants with PPA named fewer words than all other groups; participants with AD were marginally better than PPAs on this task.

The results from these traditional tasks show that participants with AD and PPA scored lower on all tasks compared to both age- and education matched, as well as young, controls. This is in line with previous research on lexical retrieval in AD and PPA. Table 5 below sums up the results on these tests.

Test	**Summary of results**
PPT	HCE > HCY > AD = PPA
SAT	HCE = HCY > AD = PPA
Word-picture matching	HCE = HCY > AD = PPA
Noun-naming	HCE = HCY > AD = PPA
Verb-naming	HCE = HCY > AD = PPA
Verbal fluency - phonological	HCE = HCY > AD = PPA
Verbal fluency - semantic	HCE = HCY > AD > PPA

Table 5: Overview of results per test

Free word association

Results from the main response types from the FWA test show that there are some differences between participants with AD and the HCY group, and between the two control groups. Participants with AD provided more responses of the *No association* type than HCY participants (AD vs. HCY: $w = 81$, $p = 0.037$). This was also the main response type that was most dominant for PPA02. PPA01's responses were in general quite similar to the averages for the HCE group, however, he did provide more descriptions relative to controls.

Between the control groups there was a statistically significant difference for the *Mixed response* type, with HCY providing more of these responses than HCE (w = 57, p = 0.032). The lack of significant differences between other groups is possibly due to the small number of participants, and few responses in each category. However, as seen in Figure 5 below, the distribution of answers in each of the main response categories differs among the groups.

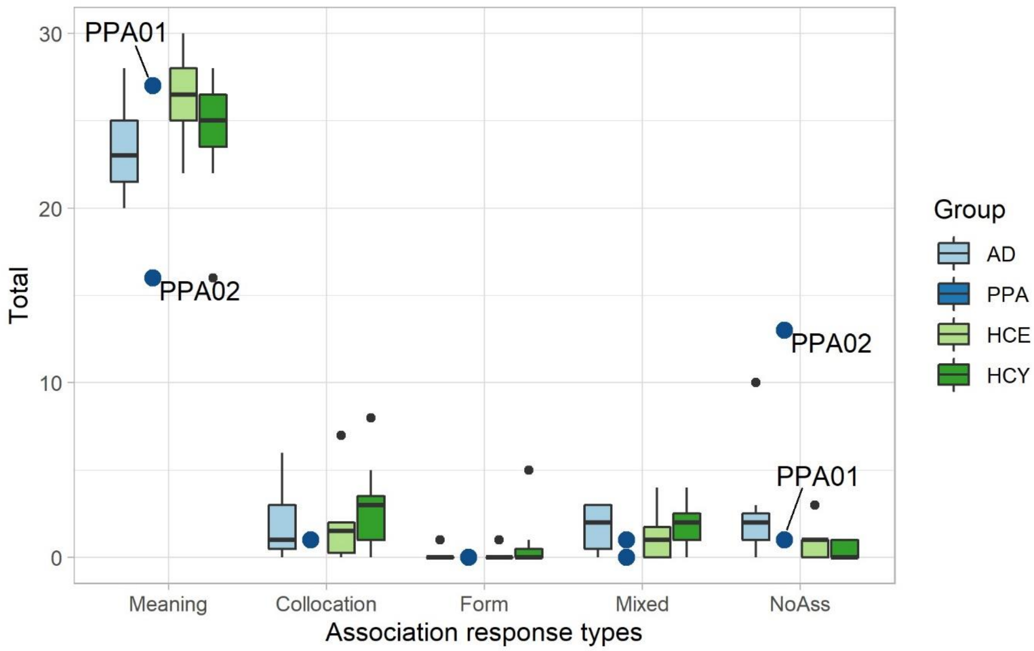


Figure 5: Distribution of responses within each main category on the free word association test. The main difference is found between the two participants with PPA. PPA01 gave more meaning related answers, on level with the older control groups, whereas PPA02 gave more answers that are classifyable as "No association". PPA01 provide one response that can be classified as Mixed, whereas PPA02 have none of these. Their collocation and form- based responses are similar, and therefore presented by only one dot for both participants in this figure. For more detail, see table 6

Participants with AD and PPA provided slightly fewer *Meaning-based* responses and more responses that are classified as *No association* than healthy participants. Participants with PPA mainly provided *no answer* or *seemingly unrelated* responses, and participants with AD gave more *descriptions* of the cue word. Table 6 gives an overview of the percentage of answers in each response category per participant with dementia, compared to averages for HCE and HCY groups.

Main response type	*Sub-type*	**AD 01**	**AD 02**	**AD 03**	**AD 04**	**AD 05**	**AD 06**	**AD 07**	**PPA 01**	**PPA 02**	**HCE (Mean)**	**HCY (Mean)**
Meaning - based	Synonym	50	36,66	33,33	26,66	36,66	10	36,66	33,33	3,33	33,6	18
	Other semantic	26,66	46,66	30	50	56,66	53,33	46,66	53,33	43,33	50	62,66
Collocation		3,33	3,33	0	16,66	0	20	3,33	3,33	3,33	5,23	8,88
Form - based		0	3,33	0	0	0	0	0	0	0	0,5	1,78
Mixed	Meaning + collocation	10	10	0	0	0	6,66	3,33	3,33	0	2,4	4,22
	Meaning + form	0	0	0	3,33	0	3,33	3,33	0	0	0,71	1,55
	Two-step	0	0	0	0	0	0	3,33	0	0	0,23	0,44
No association	Description	10	0	33,33	3,33	6,66	0	0	3,33	0	0,5	0
	No answer	0	0	0	0	0	0	0	0	30	0,5	0
	Unrelated	0	0	0	0	0	3,33	3,33	0	13,33	2	1,11
	Personal	0	0	3,33	0	0	3,33	0	3,33	6,66	4,3	1,33

Table 6: Percentage of responses within the different response categories for participants with AD, PPA and controls

An interesting observation is that participants AD04 and AD06, who scored relatively high on the neuropsychological and the other linguistic tests, almost at level with the healthy controls, have response patterns that stand out from the rest. These two participants provided fewer *synonyms* and more *collocations* than all other participants, and they have a higher number of responses that fall within the *mixed response* types *(meaning + form* and *meaning + collocation*).

There is also a small difference between the younger and the older control participants; the younger controls provide more *mixed* and fewer *meaning based* responses than the older controls. Although not significant, the younger controls give more *form based* responses than the older controls.

To sum up, participants with AD and PPA respond differently to healthy controls on free word associations, although the majority of their responses can be classified as *meaning-based* relative to the cue. Participants with AD are more similar to younger controls than to age-matched ones, indicating that not only cognitive capacity, but also age, influences responses on FWA tasks.

Discussion

We investigated how different tasks for lexical retrieval can broaden our knowledge about the semantic impairment for persons with Alzheimer's disease (AD) and Primary Progressive Aphasia (PPA). In this section we first recap the general findings on the traditional assessment tools, before answering each of the research questions regarding free word associations (FWA) introduced earlier.

In general, we found that participants with AD and PPA scored relatively well compared to healthy control participants on the comprehension tasks, suggesting that comprehension is better spared than production. In line with previous research, we also found that participants with AD and PPA provided responses that were semantically related to the target on the picture-naming tasks (Balthazar et al., 2008; Kathryn A Bayles & Tomoeda, 1983; Martin & Fedio, 1983). These findings indicate that their semantic knowledge is intact, but that the participants display an impairment in accessing the precise items for production.

Some studies report that persons with AD and PPA perform better on tasks that assess nouns relative to verbs, or vice versa (Druks et al., 2006; Hillis, Oh, & Ken, 2004; G. Robinson, Rossor, & Cipolotti, 1999; K. M. Robinson, Grossman, White-Devine, & D'Esposito, 1996; Rodríguez-Ferreiro, Davies, González-Nosti, Barbón, & Cuetos, 2009). We found no difference between naming of items from different word classes for most participants, as assessed with the picture-naming tasks, nor in comprehension, as measured by the word-picture matching tasks. However, there was a slight difference between nouns and verbs measured with the *Pyramids and Palm Trees* (PPT) and the *Semantische associatie test* (SAT), where most participants scored better on PPT than on SAT. One participant with PPA fared better on verb than on noun comprehension than the other, but there was no difference between the two on the naming tasks.

These two tests are not quite comparable, as the number of test items differ (52 in PPT vs. 20 in SAT), as does the number of distractors (participants choose between two images in PPT and three in SAT). The SAT is also not normed for Norwegian, and the images are visually more complex than in PPT.

There was substantial variation between participants on the controlled fluency tasks, and an interesting observation was the relatively high number of words produced by all participants in this study compared to norms for phonological fluency that have been published for Norwegian (Egeland, Landrø, Tjemsland, & Walbækken, 2006). One reason for this might be that the participants in this study were generally high-educated, whereas the norms from Egeland and colleagues are based on a population with a greater spread in education. Similar to other research contrasting AD and PPA (Kramer, 2003; Marczinski & Kertesz, 2006), we found that the persons with PPA in this study scored lower than persons with AD on semantic fluency.

How can results from FWA inform on semantic impairment in AD and PPA?

Our first research questions concerned how the results from the FWA test can broaden our knowledge of the semantic impairment in AD and PPA. We found that, similarly to the picture-naming task, most participants provided mainly *Meaning based* responses to cues. This is an indication of spared semantic knowledge, and that the participants associate mainly based on how words relate to each other in meaning.

However, we do see a reduction in semantically related responses for persons with AD and PPA compared to age-matched controls, and an increase in responses from other response categories. The response types *Collocations* and *Meaning + collocation* seem to be more common among the participants with AD. This indicates that in addition to meaning, the manner in which words co-occur in the language is important for lexical access for speakers with AD. This difference cannot be captured by means of bound tests.

The two participants with PPA showed a lot of variation on the FWA test; PPA01's responses were predominantly semantically related, whereas PPA02 provided many blank and unrelated responses to cues. In this respect, PPA01's responses were more similar to the profiles seen for the participants with AD than the one presented by PPA02. This might be seen in connection to his diagnosis, as he was initially diagnosed as having AD before it was later changed to PPA. There are many similarities between AD and lvPPA, and the two diseases are not always easy to tell apart (Henry & Gorno-Tempini, 2010; Sajjadi, Patterson, Arnold, Watson, & Nestor, 2012).

When used together with traditional tests for lexical retrieval, FWA tests reveal that semantic knowledge is still relatively intact for persons with AD and PPA, and that the observed impairment seems to lie in access to precise items. The FWA test can also show differences between persons with mild AD and healthy controls, indicating that there are changes in language behaviour early in AD that require more thorough investigation.

Theoretical implications

Our second research question concerned how responses from FWA can help inform on theories of the organisation of the mental lexicon. The results will be discussed in light of the two theories of lexical access introduced earlier.

The higher number of blank and/or multi-word responses provided by the participants with AD and PPA can be explained by the Transmission Deficit Hypothesis (TDH) (Burke & Shafto, 2008), which suggests that lexical access depends on sufficient activation between different nodes in the network. Difficulties in retrieving one specific word will lead the participants to provide multi-word responses to explain the word they are looking for. Prompts and context are known to aid retrieval, but as the cue words in the FWA task occur in isolation, there is less excitation between nodes, and this can lead to blank responses.

A high number of *Meaning based* responses, and responses from the sub-categories that combine *meaning* with *form* or *collocations*, indicate that the semantic knowledge about the words is still intact for the persons with AD and PPA, and that semantic information triggers other semantic responses.

The Inhibitory Deficit Hypothesis (IDH) (Zacks & Hasher, 1997) could also provide an explanation for the higher number of multi-word responses provided by persons with AD, PPA and older controls on the FWA task, as it assumes that older adults have trouble suppressing irrelevant information, resulting in personal observations and unrelated information in their speech. However, the results from this study show that personal associations are less common than pure descriptions for persons with AD. Four of the seven participants with AD provided no unrelated or personal associations at all. Descriptions were far more common, and for two participants (AD01 and AD03) they were among the most common responses. The persons with PPA showed divergent patterns, one favouring personal associations and one providing the same amount of pure descriptions as personal associations. For the older control participants, personal associations were more common than pure descriptions. The number of personal and unrelated responses were in total far fewer than other responses, which indicates that the majority of the participants in this study were able to suppress irrelevant and personal information during testing, providing less support to the IDH.

All of the participants with AD and PPA were poor at the inhibition sub-test in the *Stroop task*, both with regard to time to complete and number of mistakes. Despite showing impaired inhibition, these participants were able to steer clear of producing irrelevant responses on the FWA task.

There were also differences in the responses from the older and younger control participants, which can be explained by differences in the mental lexicon. Older participants provided more synonymous responses, suggesting that they have larger vocabularies than the younger participants. The younger participants provided more *Mixed* responses and *Collocations* than the older participants – not unlike the pattern observed for the participants with AD. These similarities may have different explanations; for persons with AD their naming strategies are based on an impairment in access to the intended lexical items, whereas the younger controls have smaller lexical networks, and therefore fewer synonyms to choose from.

Clinical implications

Our last research question related to whether response patterns from the FWA test can be diagnostically relevant for early identification of AD and PPA.

We see that our participants that do not show much impairment on conventional tests (MMSE, CERAD, picture-naming), do display differences on a test of free generation compared to healthy controls. Two of the participants, AD04 and AD06, had test scores more or less within the normal range on most tests, both cognitive and linguistic. On the FWA test, they provided more *collocation*-based responses than other participants, meaning that they associate based on words that frequently co-occur in the language. In addition to this, if we include responses that can be both *meaning* and *collocation*-based, we see a tendency that persons with AD produce more *collocation*-based associations than persons with PPA and controls. These differences cannot be captured by traditional bound and restricted tests, indicating that analysis of association responses may add valuable information about language change in early AD.

It is not possible to establish a pattern that distinguishes persons with AD from those with PPA based on only two participants. Furthermore, the response patterns from the participants with PPA were very divergent. However, compared to healthy controls and participants with AD, these two participants provided fewer collocations and mixed associations, which indicates that their associations were less dependent on how words co-occur in discourse. More research is needed to be able to say if this can distinguish between persons with AD and PPA.

Conclusion

The findings from this study reveal that persons with AD and PPA respond somewhat differently on a free word association task than neurologically healthy participants. The main differences are seen when examining the types of responses. Persons with AD provide markedly different associations than healthy controls, even in mild cases. Clinically, it is important to note that persons who score within normal ranges on cognitive tests respond differently on free word association tasks compared to controls.

All participants predominantly gave meaning-based responses; in most cases, AD participants provided semantically related responses that are not synonyms, HCE participants on the other hand gave predominantly synonym responses to cues. PPA participants' responses were more difficult to classify, and did not conform to one particular pattern. For participants with AD, responses that can belong to several categories (*Mixed*) were more common than for healthy controls. This indicates that persons with AD more easily retrieve words that have multiple connections in the mental lexicon network. Both form, meaning and usage patterns influence how words are processed, which can be explained by interactive processing models such as the *Transmission Deficit Hypothesis* (TDH).

Also in the picture-naming task, we saw more semantically related responses, indicating impairment in access and not an underlying semantic impairment, where the items are still available, but the phonological form is not. Following the TDH this means that the links between the semantic nodes in the network are stronger, and transmits enough activation based on semantics, but the phonological link is weakened.

Scoring FWA may be time consuming and requires proper knowledge of the scoring categories. Thus, it is a less convenient tool for quick analysis. However, a dichotomous scoring (syntagmatic vs. paradigmatic responses) does not fully capture the many possible relationships between cue and response, and a more nuanced scoring system may be better suited for analysis. Our study indicates that FWA may be a good addition to an assessment battery for dementia, being sensitive to deviant patterns that traditional tests, such as confrontation naming, verbal fluency and comprehension tests do not reveal, however, more research is needed to see if these patterns hold for larger participant groups. The differences between persons with AD and PPA needs to be explored further to investigate if responses on FWA tests can help differentiate between the two diseases.

Acknowledgements:

Thanks to Linda Vatne and Hanna Eidshagen Menne for assistance with data collection and coding, Pernille Hansen and Thomas Nygreen for help with statistics and visualisations in R. We also want to acknowledge the Norwegian register of persons assessed for cognitive symptoms (NorCog), for providing access to patient data, and to Anne-Brita Knapskog at the Oslo University Hospital for help with recruitment of participants.

Declaration of interest:

The authors report no conflict of interest.

Appendices

Appendix 1:

	Target	Alternative response	English translation
Synonym	“brekke”	“knekke”	Both: “to break, to snap”
Hyponym	“flue”	“insekt”	“insect” for target “a fly”
Hypernym	“sekk”	“ryggsekk”	“backpack” for target “bag”
Description	“pingvin”	“a bird from the south pole”	description of central property of the target “penguin”
Code switch	“ananas”	“pineapple” (Eng.)	Translation equivalent of Norwegian target word
Wrong word class	“danse”(v.)	“dans”(n.)	“a dance” for target “to dance”
Visual confusion	“palett”	“pensel”	“a paint brush” for target “palette” (both visible in the picture)
Semantic confusion	“hval”	“sel”	“seal” for target “whale”
No answer	“løk”	“I don’t have a word for this”	Onion

Appendix 1: Scoring categories for the different picture-naming tasks (verbs, nouns and BNT)

Appendix 2:

Main category	Sub-category	Cue	Response	English translation
1. Meaning-based	a) Synonym	"sykehus"	"hospital"	Hospital → infirmary
	b) Other semantic	"tolke"	"språk"	Interpret → language
2. Collocation		"konsert"	"hus"	Concert → venue
3. Form-based		"nyte"	"nype"	Enjoy → rosehip
4. Mixed associations	a) Meaning & collocation	"avis"	"papir"	Newspaper → paper
	b) Meaning & form	"håndtere"	"håndtak"	To handle → a handle
	c) Two-step	"tolke"	"Ringenes Herre"	To interpret → "Lord of the Rings" (via J. R. R. Tolkien – similar in form).
5. No association	a) Description	"hverdag"	"alle utenom lørdag og søndag"	Everyday → all but Saturday and Sunday
	b) No answer	"nyte"	"nei, det har jeg ingen assosiasjon til"	Enjoy → I do not have any association for this word
	c) Unrelated	"informere"	"hanske"	"inform" → "glove"
	d) Personal	"religion"	"bra"	Religion → good

Appendix 2: Scoring categories for the free word association task

References

Almor, A., Aronoff, J. M., MacDonald, M. C., Gonnerman, L. M., Kempler, D., Hintiryan, H., … Andersen, E. S. (2009). A common mechanism in verb and noun naming deficits in Alzheimer's patients. *Brain and Language*, *111*(1), 8–19. https://doi.org/10.1016/j.bandl.2009.07.009

Azuma, T., Sabbagh, M. N., & Connor, D. J. (2013). The effect of healthy aging and mild cognitive impairment on semantic ambiguity detection. *Journal of Neurolinguistics*, *26*(2), 271–282. https://doi.org/10.1016/j.jneuroling.2012.09.003

Balthazar, M. L. F., Cendes, F., & Damasceno, B. P. (2008). Semantic error patterns on the Boston Naming Test in normal aging, amnestic mild cognitive impairment, and mild Alzheimer's disease: Is there semantic disruption? *Neuropsychology*, *22*(6), 703–709. https://doi.org/10.1037/a0012919

Barresi, B. A., Nicholas, M., Connor, L. T., Obler, L. K., & Albert, M. L. (2000). Semantic Degradation and Lexical Access in Age-Related Naming Failures. *Aging, Neuropsychology, and Cognition*, *7*(3), 169–178. https://doi.org/10.1076/1382-5585(200009)7:3;1-Q;FT169

Bayles, K. A., & Tomoeda, C. K. (2007). *Cognitive-Communication Disorders of Dementia* (1 edition). San Diego: Plural Publishing Inc.

Bayles, Kathryn A, & Tomoeda, C. K. (1983). Confrontation naming impairment in dementia. *Brain and Language*, *19*(1), 98–114. https://doi.org/10.1016/0093-934X(83)90057-3

Bird, H., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2000). The Rise and Fall of Frequency and Imageability: Noun and Verb Production in Semantic Dementia. *Brain and Language*, *73*(1), 17–49. https://doi.org/10.1006/brln.2000.2293

Bowles, N. L., Obler, L. K., & Albert, M. L. (1987). Naming errors in healthy aging and dementia of the Alzheimer type. *Cortex*, *23*(3), 519–524. Retrieved from Scopus.

Bøyum, M., Hansen, P., & Ribu, I. S. (in preparation). *Assessing lexical processing with a free word association test: The roles of age, language proficiency and cognitive decline*.

Burke, D. M., & Shafto, M. A. (2004). Aging and Language Production. *Current Directions in Psychological Science*, *13*(1), 21–24. https://doi.org/10.1111/j.0963-7214.2004.01301006.x

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In *The handbook of aging and cognition, 3rd ed* (pp. 373–443). New York, NY, US: Psychology Press.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of the Alzheimer's Type: What do various measures measure? *Brain*, *113*(2), 397–417. https://doi.org/10.1093/brain/113.2.397

Chertkow, H., Whatmough, C., Saumier, D., & Duong, A. (2008). Chapter 25 Cognitive neuroscience studies of semantic memory in Alzheimer's disease. In J.-C. L. Wayne S. Sossin Vincent F. Castellucci and Sylvie Belleville (Ed.), *Progress in Brain Research*: *Vol. Volume 169* (pp. 393–407). https://doi.org/10.1016/S0079-6123(07)00025-8

Cuetos, F., Arce, N., Martínez, C., & Ellis, A. W. (2017). Word recognition in Alzheimer's disease: Effects of semantic degeneration. *Journal of Neuropsychology*, *11*(1), 26–39. https://doi.org/10.1111/jnp.12077

Cuetos, F., Rodríguez-Ferreiro, J., Sage, K., & Ellis, A. W. (2012). A fresh look at the predictors of naming accuracy and errors in Alzheimer's disease. *Journal of Neuropsychology*, *6*(2), 242–256. https://doi.org/10.1111/j.1748-6653.2011.02025.x

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS): Examiner's manual: Flexibility of thinking, concept formation, problem solving, planning, creativity, impluse control, inhibition*. San Antonio (Tex.): Pearson. /z-wcorg/.

Deyne, S. D., & Storms, G. (2015). Word Associations. *The Oxford Handbook of the Word*. https://doi.org/10.1093/oxfordhb/9780199641604.013.018

Druks, J., Masterson, J., Kopelman, M., Clare, L., Rose, A., & Rai, G. (2006). Is action naming better preserved (than object naming) in Alzheimer's disease and why should we ask? *Brain and Language*, *98*(3), 332–340. https://doi.org/10.1016/j.bandl.2006.06.003

Egeland, J., Landrø, N. I., Tjemsland, E., & Walbækken, K. (2006). Norwegian Norms and Factor-Structure of Phonemic and Semantic Word List Generation. *The Clinical Neuropsychologist*, *20*(4), 716–728. https://doi.org/10.1080/13854040500351008

Eustache, F., Cox, C., Brandt, J., Lechevalier, B., & Pons, L. (1990). Word-association responses and severity of dementia in Alzheimer disease. *Psychological Reports*, *66*(3 Pt 2), 1315–1322. https://doi.org/10.2466/pr0.1990.66.3c.1315

Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. J. (2015). Establishing the Reliability of Word Association Data for Investigating Individual and Group Differences. *Applied Linguistics*, *36*(1), 23–50. https://doi.org/10.1093/applin/amt020

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.

Gewirth, L. R., Shindler, A. G., & Hier, D. B. (1984). Altered patterns of word associations in dementia and aphasia. *Brain and Language*, *21*(2), 307–317.

Gollan, T. H., Salmon, D. P., & Paxton, J. L. (2006). Word association in early Alzheimer's disease. *Brain and Language*, *99*(3), 289–303. https://doi.org/10.1016/j.bandl.2005.07.001

Goral, M. (2004). First-language decline in healthy aging: Implications for attrition in bilingualism. *Journal of Neurolinguistics*, *17*(1), 31–52. https://doi.org/10.1016/S0911-6044(03)00052-6

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., … Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Happ, M., Bathke, A. C., & Brunner, E. (2019). Optimal sample size planning for the Wilcoxon-Mann-Whitney test. *Statistics in Medicine*, *38*(3), 363–375. https://doi.org/10.1002/sim.7983

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, *42*(9), 1212–1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Henry, M. L., & Gorno-Tempini, M. L. (2010). The logopenic variant of primary progressive aphasia. *Current Opinion in Neurology*, *23*(6), 633–637. https://doi.org/10.1097/WCO.0b013e32833fb93e

Hillis, Oh, S., & Ken, L. (2004). Deterioration of naming nouns versus verbs in primary progressive aphasia. *Annals of Neurology*, *55*(2), 268–275. https://doi.org/10.1002/ana.10812

Hodges, J. R., Patterson, K., Graham, N., & Dawson, K. (1996). Naming and knowing in dementia of Alzheimer's type. *Brain and Language*, *54*(2), 302–325. https://doi.org/10.1006/brln.1996.0077

Howard, D., Patterson, K. E., & Thames Valley Test Company. (1992). *The pyramids and palm trees test.* Thames Valley Test Co. /z-wcorg/.

Kaplan, E., Goodglass, Harold., Weintraub, Sandra., & Goodglass, Harold. (1983). *Boston naming test*. Philadelphia: Lea & Febiger. /z-wcorg/.

Kempler, D., & Goral, M. (2008). Language and Dementia: Neuropsychological Aspects. *Annual Review of Applied Linguistics*, *28*, 73–90. https://doi.org/10.1017/S0267190508080045

Kramer, J. H. (2003). Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cognitive and Behavioral Neurology : Official Journal of the Society for Behavioral and Cognitive Neurology*, *16*(4), 211–218.

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., & Andersen, E. S. (2001). Assessing Working Memory and Language Comprehension in Alzheimer's Disease. *Brain and Language*, *78*(1), 17–42. https://doi.org/10.1006/brln.2000.2436

MacKay, D. G., & Burke, D. M. (1990). Chapter Five Cognition and Aging: A Theory of New Learning and the Use of Old Connections. In F.G.A. Stone and Robert West (Ed.), *Advances in Psychology*: *Vol. Volume 71* (pp. 213–263). https://doi.org/10.1016/S0166-4115(08)60159-4

Marczinski, C. A., & Kertesz, A. (2006). Category and letter fluency in semantic dementia, primary progressive aphasia, and Alzheimer's disease. *Brain and Language*, *97*(3), 258–265. https://doi.org/10.1016/j.bandl.2005.11.001

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language*, *19*(1), 124–141. https://doi.org/10.1016/0093-934X(83)90059-7

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., … Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's &*

Dementia: The Journal of the Alzheimer's Association, *7*(3), 263–269. https://doi.org/10.1016/j.jalz.2011.03.005

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PLOS ONE*, *12*(2), e0171935. https://doi.org/10.1371/journal.pone.0171935

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., … Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165.

Nicholas, M., Obler, L., Albert, M., & Goodglass, H. (1985). Lexical Retrieval in Healthy Aging. *Cortex*, *21*(4), 595–606. https://doi.org/10.1016/S0010-9452(58)80007-6

Obler, L. K., & Gjerlow, K. (1999). *Language and the Brain*. Cambridge University Press.

R core team. (2019). R-Project (Version 3.6.0).

Ribu, I. S. (under revision). *First in, last out: A systematic review of the effects of (psycho-)linguistic variables on lexical retrieval in Alzheimer's disease and primary progressive aphasia*.

Robinson, G., Rossor, M., & Cipolotti, L. (1999). Selective sparing of verb naming in a case of severe Alzheimer's disease. *Cortex*, *35*(3), 443–450. Retrieved from Scopus.

Robinson, K. M., Grossman, M., White-Devine, T., & D'Esposito, M. (1996). Category-specific difficulty naming with verbs in Alzheimer's disease. *Neurology*, *47*(1), 178–182. Retrieved from Scopus.

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

RStudio Team. (2015). *RStudio: Integrated Development for R. RStudio, Inc.* Retrieved from http://support.rstudio.com/hc/en-us/articles/206212048-Citing-RStudio

Sajjadi, S. A., Patterson, K., Arnold, R. J., Watson, P. C., & Nestor, P. J. (2012). Primary progressive aphasia: A tale of two syndromes and the rest. *Neurology*, *78*(21), 1670–1677. https://doi.org/10.1212/WNL.0b013e3182574f79

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428.

Santo Pietro, M. J., & Goldfarb, R. (1985). Characteristic patterns of word association responses in institutionalized elderly with and without senile dementia. *Brain and Language*, *26*(2), 230–243.

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203–214. https://doi.org/10.1016/S0887-6177(03)00039-8

Visch-Brink, E., Stronks, D., & Denes, G. (2005). *Semantische Associatie test (SAT)*. Pearson.

Vitevitch, M. S., & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological Science*, *9*(4), 325–329.

Wechsler, D. (2008). *WAIS-IV : Wechsler adult intelligence scale*. San Antonio, Tex.: Pearson. /z-wcorg/.

Wingfield, A., & Stine-Morrow, E. A. L. (2000). Language and speech. In *The handbook of aging and cognition, 2nd ed* (pp. 359–416). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Zacks, R., & Hasher, L. (1997). Cognitive Gerontology and Attentional Inhibition: A Reply to Burke and McDowd. *The Journals of Gerontology: Series B*, *52B*(6), P274–P283. https://doi.org/10.1093/geronb/52B.6.P274

Ribu, I. S. & Kuzmina, E. (submitted). Tracking non-canonical sentence comprehension in Alzheimer's disease and Primary progressive aphasia

III

Tracking non-canonical sentence comprehension in Alzheimer's disease and Primary progressive aphasia

Abstract

Persons with Alzheimer's disease (AD) and Primary progressive aphasia (PPA) often experience sentence comprehension deficits, despite having intact single word comprehension. It has been suggested that these deficits are mediated by working memory impairments.

The goal of the study was twofold; (1) to investigate whether impairment in linguistic and cognitive abilities are related to the sentence comprehension deficits in AD and PPA; and (2) to investigate whether both online and offline processing are impaired in AD and PPA; and to explore the relationship between online and offline processing.

We measured both online (eye-tracking) and offline (accuracy) sentence comprehension in five participants with AD and two with a logopenic variant of PPA and a group of age- and education matched controls. Sentence comprehension performance was contrasted with results on linguistic and cognitive background tests.

Persons with AD and PPA scored lower on the sentence comprehension task relative to neurologically healthy controls. Accuracy was generally high for all participants, AD and PPA participants were slower at detecting the target image. The impairment was more pronounced on accuracy judgement than in the gaze data.

The offline performance (reflected in accuracy performance) of participants with AD and PPA seems to be worse than their online performance (reflected in gaze data) Both neurologically healthy and persons with AD and PPA rely on an agent-first parsing strategy for object-cleft sentences, which indicates that construction frequency plays a major role in online processing.

Key words: Alzheimer's disease, Primary Progressive Aphasia, Sentence comprehension, Eye tracking

Abbreviations:

A = Active

AD = Alzheimer's disease

AOI = Area of interest (image)

ET = Eye-tracking

HC = Healthy control

OC = Object cleft

PLT = Proportion of looks to target

PPA = Primary progressive aphasia

ROI = Region of interest (audio)

SC = Subject cleft

VWP = Visual world paradigm

WM = working memory

Introduction and background

Sentence comprehension requires processing of individual words, correct mapping between arguments and their thematic roles, relating pronouns to referents, establishing the time-frame of the sentence, among other things. Speakers are often able to map arguments correctly onto thematic roles based on world/semantic knowledge. For example, in the sentence “The boy kicks the ball” we know that only *the boy* can be an agent, because only animate entities can kick. For this reason, syntactically-based sentence comprehension can only be tested using semantically reversible sentences[1] such as, for example “The boy kicks the girl”.

Depending on the language, different linguistic devices may play a more prominent role than others. Word order, for instance, is more important for sentence parsing than inflectional morphology in English, whereas morphological cues, such as case marking, is more important than word order in languages such as German. In languages such as English and Norwegian, the agent of canonical sentences precedes the theme/patient of the sentence, e.g., “The boy paints the girl”. On the contrary, the theme/patient precedes the agent in non-canonical sentences; e.g., “The girl is painted by the boy”, where the theme/patient is the first element of the sentence. Non-canonical sentences are considered more complex, because they feature a mismatch between the word order and the order of thematic roles (Ferreira, 2003). Furthermore, since non-canonical sentences are less frequent than canonical sentences, they tend to be more difficult to process (Dąbrowska & Street, 2006; Street, 2017; Street & Dąbrowska, 2010).

Following usage-based theories of language, linguistic knowledge is shaped by usage factors such as frequency, and the language user’s perception of similarity between linguistic units (Bybee, 2010). Such theories of language predict frequency effects for shorter and longer constructions, even sentences, which can in part describe poorer performance on tasks including less frequent structures. According to this view participants’ performance on sentence processing tasks are shaped by the regularity of the constructions, and language experience in addition to frequency of individual linguistic units (Gahl & Menn, 2016; Maryellen C. MacDonald & Christiansen, 2002; Street, 2017). Constraint-based theories of sentence processing (i.e., Maryellen C. MacDonald & Seidenberg, 2006), fits well with this view, as such theories assume that the frequencies and distributions of events in the linguistic environment

[1] In semantically reversible sentences, the sentence is still meaningful if the main participants in the sentence change roles. The sentence in (2) becomes meaningless if *The boy* and *the ball* change positions (*The ball kicks the boy*).

can be used to infer the meaning of the sentence (M. C. MacDonald, Pearlmutter, & Seidenberg, 1994; Maryellen C. MacDonald & Seidenberg, 2006; Seidenberg & McClelland, 1989).

A different view is presented in the *Derived order problem hypothesis* (DOP-H) (Bastiaanse & van Zonneveld, 2005, 2006), which assumes that comprehension of non-canonical sentences, or sentences with non-base word order, are harder to process than sentences that follow base order. The former involve syntactic movement, which require more processing capacities (Bastiaanse & van Zonneveld, 2005, 2006; Jap, Martinez-Ferreiro, & Bastiaanse, 2016). For comprehension, the DOP-H predicts that sentences in which the thematic roles are not in the basic order (i.e., sentences where the theme precedes the agent), will result in chance level performance for persons with reduced processing capacities (Bastiaanse & van Zonneveld, 2006).

Persons with neurodegenerative diseases, such as Alzheimer's disease (AD) and Primary progressive aphasia (PPA), are typically impaired in sentence comprehension; In particular, non-canonical sentences have been found to be impaired in dementia (Waters & Caplan, 1997). The underlying causes for these deficits are not entirely clear. Deficits in lexical semantics, syntactic knowledge, and working memory may all play a role (Almor, Kempler, MacDonald, Andersen, & Tyler, 1999; Kempler, Almor, & MacDonald, 1998; Rochon, Waters, & Caplan, 1994). To shed light on factors underlying sentence comprehension deficit in AD and PPA, we addressed the following research questions:

1. Are there differences in auditory comprehension accuracy and processing strategies applied to canonical and non-canonical sentences between AD and PPA and neurologically healthy controls?
2. Is there a relationship between auditory sentence comprehension and cognitive functions in AD and PPA?
3. Can sentence comprehension tasks help distinguish between AD and PPA?

To answer these questions, we investigated offline performance and online processing patterns of canonical and non-canonical sentences for participants with AD and PPA in contrast to a group of healthy controls. Additionally cognitive abilities, naming and semantic knowledge of all participants were assessed.

Sentence comprehension deficits in PPA and AD

PPA is a collective term for three neurodegenerative conditions which mainly affects language abilities: (1) a semantic variant of PPA (svPPA); (2) a non-fluent/agrammatic variant of PPA (nfavPPA); and (3) a logopenic variant of PPA (lvPPA). The different variants are distinguished based on language manifestations and site(s) of neurological atrophy. Patients with svPPA have impaired confrontation naming, and word comprehension; persons with nfavPPA often have slow, haltering speech and show signs of agrammatism. lvPPA is the variant that is most similar to AD, and its core features are impaired single-word retrieval, in spontaneous speech and confrontation naming, and impaired repetition of sentences and phrases, but single-word comprehension and object knowledge are usually spared. In the remainder of this paper, the term PPA is used to refer to lvPPA unless otherwise specified, since the participants in this study were diagnosed with lvPPA.

Multiple studies on AD reported impaired sentence comprehension of complex structures, such as sentences with non-canonical/non-base word order (Caplan, 1996; Caplan, DeDe, Waters, Michaud, & Tripodis, 2011; Croot, Hodges, & Patterson, 1999; Kempler et al., 1998; Marková, Horváthová, Králová, & Cséfalvay, 2016; Price & Grossman, 2005; Rochon et al., 1994; Small, Kemper, & Lyons, 1997; Waters, Rochon, & Caplan, 1998); Comprehension of sentences that requires a lot of inferences and comprehension of multi-party conversations is also impaired (Bayles & Tomoeda, 2007; Müller, 2013). As the disease progresses, it becomes increasingly difficult for individuals with AD to comprehend spoken and written language (Müller, 2013).

It has also been found that persons with AD are impaired in comprehension of semantic roles associated with verbs. Grossman and White-Devine (1998) and Price and Grossman (2005) suggested that this impairment could be explained either by a decline in knowledge about verb thematic roles, or by a deficit in the processing of the thematic role component of verbs. They also argued that persons with AD have limitations in cognitive resources needed to support atypical semantic-thematic mapping, as well as an impoverished grasp of the thematic roles associated with verbs (ibid.). Working memory (WM) deficits are often observed for persons with AD, and have been cited as the source of sentence comprehension impairment (Small et al., 1997). A relationship between WM impairments and impaired sentence comprehension has been found in studies that investigate comprehension of sentences where the number of propositions (finite verbs) and sentence length are the main features under investigation. Croot, Hodges and Patterson (2009) found that participants with AD performed slightly worse than controls in understanding sentences with two propositions, than on sentences with one proposition. Additionally, participants with AD were more impaired in understanding two-proposition sentences with non-canonical order of thematic roles than two-proposition sentences with canonical order of thematic roles. The researchers ascribe this to difficulties with the post-interpretative stage of processing, where the propositional meaning of the sentence is matched to the appropriate picture in a sentence-to-picture matching task (Croot et al., 1999). Similar findings have been central in research on sentence comprehension in studies by Caplan and colleagues (e.g., Caplan et al., 2011; Rochon et al., 1994; Waters et al., 1998). According to these authors, performance on sentence comprehension tasks is not particularly affected by syntactic complexity, but that the difficulties arise at the post-interpretative processing, relying on non-linguistic capacities, such as WM are critically involved (Rochon et al., 1994; Waters et al., 1998).

Only few studies have investigated sentence comprehension in lvPPA: Wilson and colleagues (2012) compared sentence comprehension abilities between persons with nfavPPA and lvPPA, and found that sentence comprehension in persons with lvPPA was good for short sentences, even passives. However, the participants with lvPPA had more difficulties with longer sentences. The authors attributed this to the limited short-term memory capacity for individuals with lvPPA. In contrast, Charels and colleagues (2014) compared participants with lvPPA and svPPA to participants with nfavPPA. They found that persons with lvPPA showed a deficit for centre-embedded sentences (i.e., *the fox that followed the domesticated cat was fierce*). They did not relate their findings to short-term memory deficits, but rather to atrophy in fronto-temporal brain areas that involves a sentence-processing network (Charles et al. 2014).

Zimmerer, Dabrowska, Romanowski, Blank and Varley (2014) reported a case of a person with lvPPA who performed better on passive than active sentences in a sentence comprehension task. The participant was not able to determine agent and patient in transitive active structures, but he displayed comprehension of full passives (Zimmerer et al., 2014; Zimmerer & Varley, 2015).

To our knowledge, there are no eye-tracking (ET) studies published on sentence comprehension in AD and PPA. Data from ET studies can help shed light on the real time language processing, since a person's visual attention is associated with the point of fixation (Duchowski, 2007). The visual world paradigm (VWP), where visual and auditory stimuli are presented simultaneously, is a popular paradigm for the study of sentence comprehension. The VWP is based on the mind-eye hypothesis, which states that gaze is driven by attention (Sekerina, 2014). Healthy individuals shift their attention to the target object immediately, within 200 milliseconds (ms) after hearing the relevant words (Matin, Shao, & Boff, 1993; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). As eye-fixations are time-locked to the continuous auditory stimuli, ET data provides a window into real-time sentence processing (Tanenhaus et al., 1995).

The Norwegian cleft constructions

Norwegian declarative sentences follow a strict *Subject – Verb – Object* (SVO) word order, as shown in (1) below. Cleft sentences, like in English, are not relative clauses per se, although they share many similar properties. Cleft sentences consist of a main clause and a relative clause. The main clause includes a dummy subject *det* ('that'/'it'), which refers to the subject in the relative clause, followed by a copula *å være* ('to be') and then the relative clause (Faarlund, 1997, p. 1089). In subject clefts (SC), the canonical SVO pattern of Norwegian main declaratives is retained in the relative clause, whereas in object cleft (OC) sentences the word order in the relative clause is OSV. This is exemplified in (2) and (3) below.Active:

Subject	verb	Object
Gutten	maler	jenta
The boy	*paints*	*the girl*

(1) Subject cleft:

'Dummy subject'	COPULA	[Relative clause]
Det	er	gutten som maler jenta
		S V O
It	*is*	*the boy who is painting the girl*

(2) Object cleft:

'Dummy subject'	COPULA	[Relative clause]
Det	er	jenta (som) gutten maler
		O S V
It	*is*	*the girl (whom) the boy is painting*

The relative subjunction *som* is obligatory in SC sentences, and optional in OC sentences. In the material used in the current experiment, the subjunction is omitted in the OC condition, to further disambiguate the OC sentences from SC sentences.

There are no morphological cues in Norwegian (i.e., case marking) which aid mapping of thematic and syntactic roles to the lexical form, and listeners have to rely on word-order cues and semantic knowledge. In (2) the thematic roles follow canonical order (*agent* before *theme*), whereas there is a non-canonical order of thematic roles (*theme* before *agent*) in sentence (3).

A search in NorGramBank (Rosén, De Smedt, Meurer, & Dyvik, 2012), a syntactically annotated treebank for the two written variants of Norwegian (Bokmål and Nynorsk), returned a total of 781356 hits for the two cleft sentence types as well as declarative sentences in the active voice (A). Table 1 show the distribution between the constructions.

Table 1: Distribution of constructions in the NorGramBank tree bank.

Sentence type	Hits in NorGramBank	Percentage
Active	744 358	95,3%
Subject cleft	32 343	4,1%
Object cleft	4 656	0,6%

Methods

Participants:

Five persons with AD and two with lvPPA were recruited from the memory clinic at the Oslo University Hospital to participate in the current study. Nine neurologically healthy control participants (HC) were also included in the study.[2] Table 2 summarizes demographic details. The participants with AD and PPA were assessed and diagnosed by neurologists at the Oslo University Hospital. Diagnosis was based on a range of standardized tests, and supported by biomarkers in cerebrospinal fluid and through imaging (fMRI, SPECT, CT). Persons with other neurological diseases, premorbid language disorders or a history of alcohol and drug abuse were not included in the study.

[2] A total of seven persons with AD and 14 control participants were initially included in the study. Two participants were excluded because it was impossible to obtain good calibrations and validations from them due to scarring on the cornea from laser surgery. Further four participants (one with AD) were excluded after recording because of poor tracking ratio.

One participant with PPA (PPA1) was initially diagnosed with AD, but the diagnosis was later changed to lvPPA. In many cases, persons with a PPA diagnosis do not fit either of the three subtypes (Henry & Gorno-Tempini, 2010; Sajjadi, Patterson, Tomek, & Nestor, 2012), and some argue that lvPPA is a form of atypical AD (Ahmed, Jager, Haigh, & Garrard, 2012). Current trends in PPA research suggest that there might be two different types of lvPPA, one which is more similar to AD, and one which is more related to the originally proposed PPA criteria (Ahmed & Garrard, 2012; Matias-Guiu et al., 2018; Vandenberghe, 2016). The uncertainty surrounding PPA1's diagnosis is important to keep in mind for the discussion of his results.

All participants reported Norwegian as their main language of use. However, all participants reported exposure to English-language media on at least a weekly basis, and all participants had learned at least one foreign language in school or at university level. Two of the participants with AD, one with PPA and four of the older control participants reported having used another language than Norwegian professionally.

Wilcoxon sum rank tests showed no difference between the participants with AD or PPA and HC in age (AD vs. HC: $w = 33$, $p = 0.172$; PPA vs. HC: $w = 16.5$, $p = 0.095$). HC participants had higher education than AD ($w = 5.5$, $p < 0.05$), but there was no difference between HC and PPA ($w = 0.5$, $p = 0.055$).

The study was approved by the Regional committees for medical and health research ethics in South-East Norway (REK), the data protection officials at the University of Oslo, and the Oslo University Hospital. All participants gave their written informed consent. Participants with AD and PPA were considered by their primary physicians as competent to give consent.

Table 2: Demographics for participants with AD, PPA and neurologically healthy older (HC) control participants.

	AD (N = 5)		**PPA (N=2)**	**HC (N = 9)**	
Female:Male	3:2		0:2	5:4	
	Mean	SD		Mean	SD
Age	74	4,6	75; 81	70	4,34
Education (years)	13,8	2,71	15; 14	17,61	1,59
MMSE	27	2,44	24; 29	29,11	0,73

Materials:

Participants were tested with linguistic and cognitive tests, as well as a sentence-to-picture matching eye-tracking task.

Cognitive screening:

The cognitive test battery consisted of the *Mini-Mental State Examination* (MMSE) (Folstein, Folstein, & McHugh, 1975); a word-list learning and recall test from *The Consortium to Establish a Register for Alzheimer's Disease* (CERAD) (Morris et al., 1989); three digit span tasks (forward, backward and ordering) (MacDonald, Almor, Henderson, Kempler, & Andersen, 2001; Wechsler, 2008); figure copying and recall (from the CERAD battery); the Stroop task (Delis, Kaplan, & Kramer, 2001); and *Trail making test A and B* (TMT) (Tombaugh, 2004). Participants with AD and PPA were tested with MMSE, CERAD and TMT A and B at the memory clinic, and the additional tests at the University of Oslo.

Language screening: offline measures

The language test battery included object and action picture naming, and word-to-picture matching tasks,[3] phonological and semantic verbal fluency tasks (Egeland, Landrø, Tjemsland, & Walbækken, 2006), Pyramids and palm trees battery (PPT) (Howard, Patterson, & Thames Valley Test Company, 1992) and the Semantische Associatie test (SAT) (Visch-Brink, Stronks, & Denes, 2005).

Language screening: online measures:

The eye-tracking experiment consisted of 45 sentences (15 of each sentence type *Active* (A), *Subject cleft* (SC), and *Object cleft* (OC)), inspired by the sentences used in the Norwegian version of the sentence comprehension sub-test in the *Verb and Sentence Test* (VAST) (Bastiaanse, Lind, Moen, & Simonsen, 2006), with certain modifications outlined below:

1) Number of pictures were reduced from four to two per trial (i.e., target and distractor). This was done to ensure that it was sentence structure and not lexical comprehension that was in focus (i.e., by not introducing images which show different actions than the one in question).

[3] The picture naming and word-picture matching tasks have been specifically developed for this project. More information about the items selection for these tests can be found in Ribu, Norvik, Lehtonen, & Simonsen, (submitted).

2) A context sentence introducing the main participants was added before the stimulus sentence, so that e the stimulus sentences were not introduced in isolation, making them sound more naturalistic.[4]
3) 15 sentences were added based on the discarded distractor images, to create a larger set of test items, and to have enough sentences of each sentence type.

The sentences were constructed using 15 different verbs (see Appendix I), and each verb was used once for each sentence type (see examples (1) – (3)). All sentences (context and stimuli sentences) were recorded by a male professional voice actor, and manipulated post recording to ensure that all sentences in each sentence type test block were approximately of the same length. The complete list of stimuli with English translations can be found in Appendix II.

Test images were square 550x550 pixels, black-and-white line drawings, which either corresponded to the stimulus sentence (Target), or including the same actors and verb but with reversed roles (Distractor). The target picture alternated between the left and right sides of the screen. The pictures were separated from each other with a grey frame (see Figure 1).

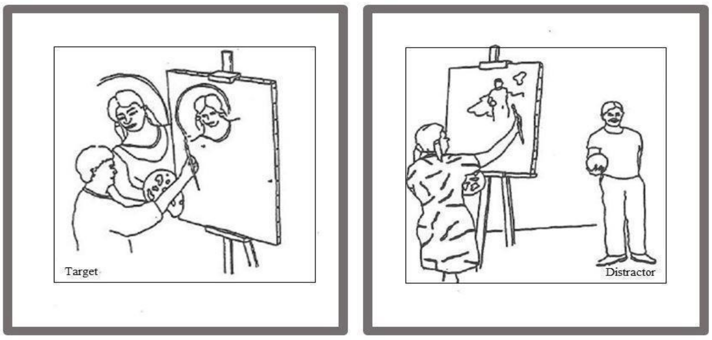


Figure 1 Example of the visual stimuli, with AOI Target (left) and Distractor (right). Everything which is not within the Target or Distractor AOI is considered White Space

Three *Areas of Interest* (AOI) were defined for each on-screen image stimulus; the *Target* picture, the *Distractor* and *White Space*. White Space included all parts of the screen that were not Target or Distractor. Furthermore, three *Regions of interest* (ROI) (i.e., Region 1, Region 2, Region 3) were assigned to the audio stimuli. For A and SC sentences, Region 2 was

[4] Example of a context sentence: *On these two pictures you see a boy and a girl.* The order of the main participants were varied across trial; in some sentences the boy would be introduced before the girl, and in other sentences the girl was introduced before the boy.

the critical region (where disambiguation took place), whereas the critical region for OC sentences was Region 3. In all cases, the critical region coincided with the main verb of the sentence. The verb serves important functions in the sentence, as it conveys information about the action being referred to, as well as about the number of, and roles of the participants or objects involved in this action. For an overview of the different ROIs, see Table 3.

Table 3: Regions of interest (ROI) for the auditory stimuli

	Region 1	**Region 2**	**Region 3**
ACTIVE (A)	Gutten	maler	jenta
Grammatical/thematic role	Subject/Agent	Verb	Object/Patient
	Sentence onset	Critical region	Sentence offset
SUBJECT CLEFT (SC)	Det er gutten som	maler	jenta
Grammatical/thematic role	Subject/Agent	Verb	Object/patient
	Sentence onset	Critical region	Sentence offset
OBJECT CLEFT (OC)	Det er jenta	gutten	maler
Grammatical/thematic role	Object/Patient	Subject/Agent	Verb
	Sentence onset	Sentence content	Critical region

Procedure:

The sentences were divided into three testing blocks with 15 sentences each. Within each block, each sentence type occurred five times, but each verb only once. The presentation of items was randomized for all participants, but never with more than three items of the same sentence type (A, SC or OC) after each other. After each testing block, participants themselves chose how much break-time they needed before continuing with the next block. In no case did the experiment last more than 20 minutes, including breaks. The first trial included three practice items (one of each sentence type) which were not included in the analysis. If participants showed poor understanding of the task during the practice trials, the experiment was stopped, instructions were repeated and the experiment was restarted with new calibration.

Each trial began with a fixation cross in the middle of the screen, to divert the participants' gaze towards the centre. After 1500 milliseconds (ms), the target and distractor image appeared, and after 1000 ms the context sentence introducing the actors on the images was played. This gave the participants in total 3500 ms to familiarize with the images before the stimulus sentence was spoken. The images were left on the screen for 7000 ms before the next trial started with a new fixation cross. Figure 2 gives a schematic representation of the test procedure.

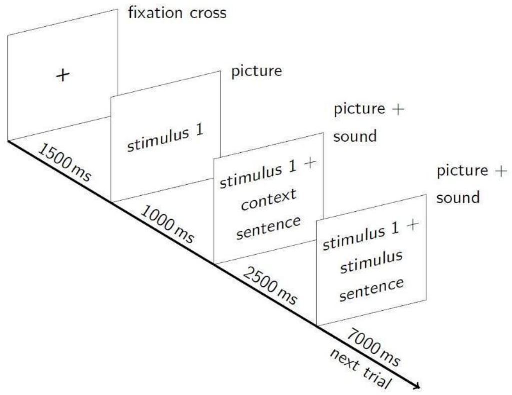


Figure 2 Test procedure for one trial in the eye-tracking experiment

Two control participants completed only two out of three sessions, due to technical difficulties with the apparatus. For these participants we only have data from 10/15 trials per sentence type.

Apparatus:

The experiment was programmed in SMI Experiment Center (SMI Experiment Suite TM3.6) on a laptop computer (Intel ® core ™ i7-4710MQ processor 2.500GHz 2,5 GHz; 8GB RAM) which was also used to control the experiment through the SMI eye-tracking software. Stimuli were presented on an external screen (34 cm diagonal size) with a SMI remote eye-tracker (SMI RED250Mobile) attached to the bottom of the screen. Sampling rate was set to 250 Hz.[5] At the beginning of each trial, a five-point calibration was executed.

Analysis:

Both accuracy and eye-tracking data were analysed. The statistical analysis was performed in in R (R core team, 2019).

Sentence comprehension accuracy was measured by having participants point to the left or the right picture, or to raise their left or right hand, indicating which image they believed

[5] For two participants, sampling rate was set to 60Hz by mistake, but this did not influence the analysis.

corresponded to the stimulus sentence. Because of the long offset time of each trial (700 ms), participants also had the time to change their mind regarding the judgment they had just made, in case they themselves wanted to do so. In these cases, the test administrator would ask which response the participant wanted to have recorded.

ET data from correctly performed trials only were used for creating the plots representing sentence processing patterns. To make these plots, we calculated mean proportion of time participants spent fixating on the target picture for each 50 ms timeslots using the following formula:

$$\frac{Target}{(Target + Distractor + White\ Space)}$$

Results

We first report the results from the cognitive and linguistic tasks used as background measures. The results from the sentence comprehension task is discussed separately; the behavioural data is reported before the eye-tracking data.

Cognitive and linguistic screening

Participants with AD and PPA are discussed individually and compared to averages from the neurologically healthy control (HC) participant group. On the cognitive tests, participants with AD and PPA performed poorer than control participants, as expected, this can be seen in Table 4.

Table 4: results from the cognitive screening for participants with AD and PPA, as well as averages for the HC group.

	AD1	AD2	AD3	AD4	AD5	PPA1	PPA2	HC (Mean)
Age/sex (Span)	72/f	76/m	79/f	66/m	77/f	75/m	81/m	70 (4m:5f) 66-78
Education (years) (Span)	12	18	14	15	10	15	14	17,61 (15-20)
MMSE	29	30	23	27	24	24	29	29,11
CERAD Word list	15	17	10	19	18	15	18	22,66

learning (max. 30)								
CERAD Word list recall (max. 10)	0	4	1	3	1	4	4	7
CERAD Word list recognition (max. 20)	8	N/A	6	7	6	7	10	19,33
CERAD Figure copy (max. 11)	11	N/A	9	11	8	N/A	10	10,55
CERAD Figure recall (max. 11)	5	N/A	N/A	N/A	6	N/A	8	10,44
TMT-A	83 sec.	45 sec.	28 sec.	37 sec.	61 sec.	85 sec.	105 sec.	35,77 sec.
TMT-B	207 sec.	144 sec.	78 sec.	69 sec.	236 sec.	300 sec.	486 sec.	78,22 sec.
TMT B-A	124 sec.	99 sec.	50 sec.	32 sec.	175 sec.	215 sec.	381 sec.	42,44 sec.
Stroop (word reading)	58 sec.	41 sec.	60 sec.	53 sec.	95 sec.	90 sec.	N/A	43,66 sec.
Stroop (color-word interference)	225 sec.	143 sec.	128 sec.	98 sec.	225 sec.	200 sec.	N/A	101,83 sec.
Stroop cost (CWI-WR)	167 sec.	102 sec.	68 sec.	45 sec.	130 sec.	110 sec.	N/A	56,85 sec.
Digit forward (max. 16)	10	7	8	10	9	8	6	9,66
Digit backward (max. 14)	5	5	9	7	4	2	3	8,11
Digit ordering (max. 15)	11	12	11	15	11	6	6	13,88
Total phonological fluency (/f/, /a/, /s/)	43	44	25	44	38	30	18	62,66

Semantic fluency (animals)	14	15	8	22	12	7	5	24,33

Raw scores for the cognitive tests (MMSE – Digit ordering) and verbal fluency.

NA = not all participants have completed all tests.

Table 5 summarizes the scores on the linguistic tests. The main differences between participants with AD/PPA and HC are found on tests for semantic knowledge (Pyramids and Palm Trees and Semantische Associatie Test), and picture naming. Word comprehension as measured by the word-to-picture matching task is not impaired.

Table 5: Results from the language battery for participants with AD and PPA, as well as averages for the HC group.

	AD1	AD2	AD3	AD4	AD5	PPA1	PPA2	HC (Mean)
Pyramids and palm trees (no. items 52)	92,3%	92,3%	82,7%	100%	86,5%	92,3%	86,5%	99,13%
Semantische associatie test (no. items 20)	90%	90%	85%	95%	65%	55%	85%	96,65%
Word-picture matching (no. items 60)	100%	100%	98,3%	100%	96,6%	100%	95%	100%
Noun-naming (no. items 30)	80%	83,3%	90%	100%	86,6%	83,3%	73,3%	97,4%
Verb-naming (no. items 30)	93,3%	100%	90%	100%	90%	93,1%*	73,3%	98,86%

*PPA1's scores on the verb naming test is based on a total of 29 items rather than 30, due to a problem with the recording.

Accuracy data

Most participants scored high on the sentence comprehension task, but there was some individual variation, mainly for the participants with AD and PPA. All results can be seen in Table 6, the scores for HC are in averages. All participants scored above chance level, indicating that they did not guess when selecting the target image, and that they are not impaired on the tested sentence structures.

Table 6: Offline judgments per participant with AD and PPA on each sentence type. Two control participants only performed 2/3 test sessions.

	AD1	AD2	AD3	AD4	AD5	PPA1	PPA2	HC (N=9)
Active	80% (12/15)	93,3% (14/15)	100%	93,3% (14/15)	100 %	86,6% (13/15)	93,3% (14/15)	100%
Subject cleft	86,6% (13/15)	93,3% (14/15)	100%	100%	100 %	86,6% (13/15)	100%	100%
Object cleft	93,3% (14/15)	93,3% (14/15)	100%	93,3% (14/15)	100 %	86,6% (13/15)	73,33% (11/15)	99,2%

Eye-tracking data

In the results from the eye-tracking data, we measure sentence comprehension by means of proportion of looks to target (PLT). Due to the small number of participants, each participant with AD and PPA will be discussed individually, and compared to the control group. Averaging scores for five, respectively two participants is futile, since the variation between individuals would provide misleading scores and results. The plots show the PLT on the y-axis, and time from sentence onset on the x-axis.

Healthy control (HC) participants:

For A and SC sentences, HC participants quickly detected the target image, indicated by a high proportion of looks to target (PLT) within, or right after, the region of interest (ROI). The PLT never dropped below 50% for these two sentence constructions. The two conditions are equally easy to comprehend, despite the difference in constructional frequency. This has also been found for English (Ferreira, 2003).

The OC sentence condition presents an interesting pattern. For the HC group there was a clear effect of an agent-first parsing strategy initially. Before the onset of the ROI the PLT indicated 75% looks to distractor: At this point in time, participants considered the image with reversed roles to correspond to the spoken sentence. Once the complete auditory stimulus was presented, gazes shifted towards the target image. The PLT was upwards of 90% by the ROI offset, indicating a fast transition during the few hundred milliseconds it took to hear the disambiguating cue word (the verb). This agent-first parsing strategy is likely mediated by the constructional frequency of sentences which follow this structure (cf. table 1 above).

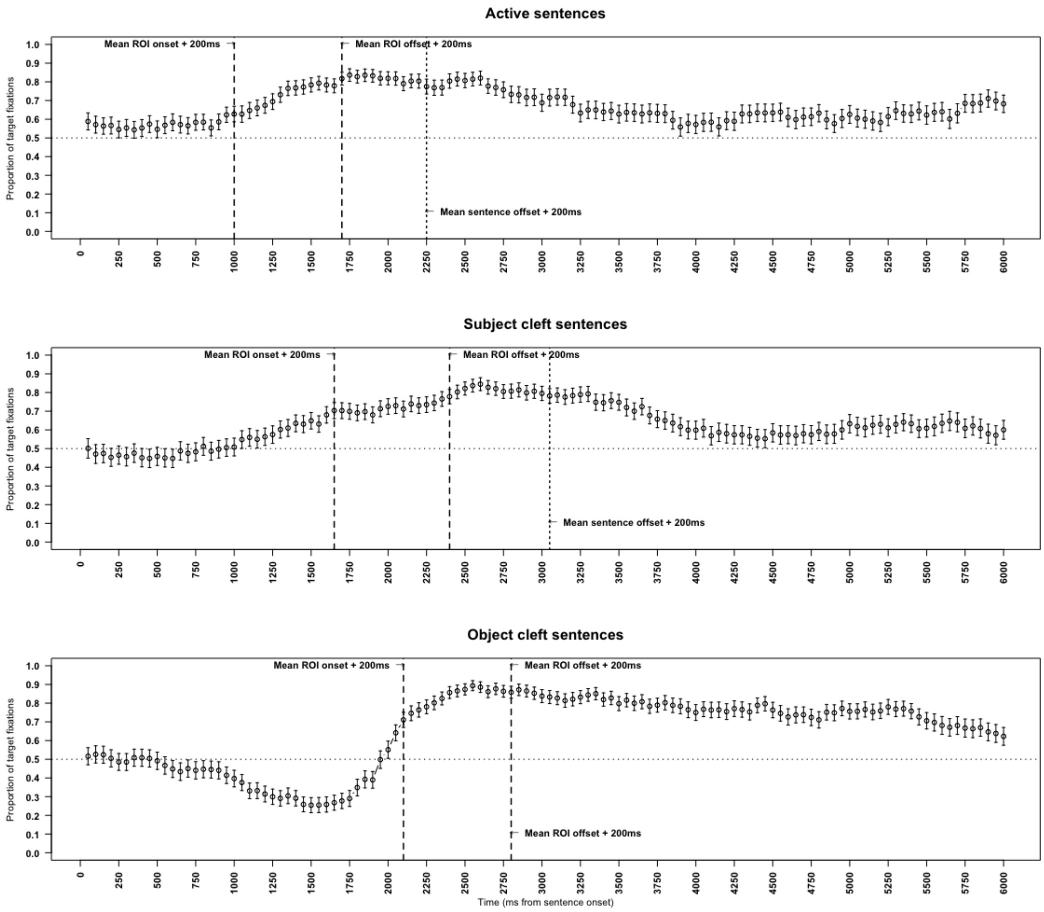


Figure 3 HC participants' gaze patterns on the ACTIVE (top), SUBJECT CLEFT (middle) and OBJECT CLEFT (bottom) sentence constructions.

Participant AD1:

Participant AD1 showed an interesting profile; she made three mistakes on the A sentence type, two on the SC and one on the OC condition in the offline task, indicating that she had some difficulties pointing to the correct image, for the simpler sentence structures, A and SC compared to the more difficult OC. However, if we look at her gaze data, we see that she was slower at reaching a high PLT on the A and OC sentence types compared to controls. On the SC sentence type, she followed the HC participants both in terms of time to detect target and PLT.

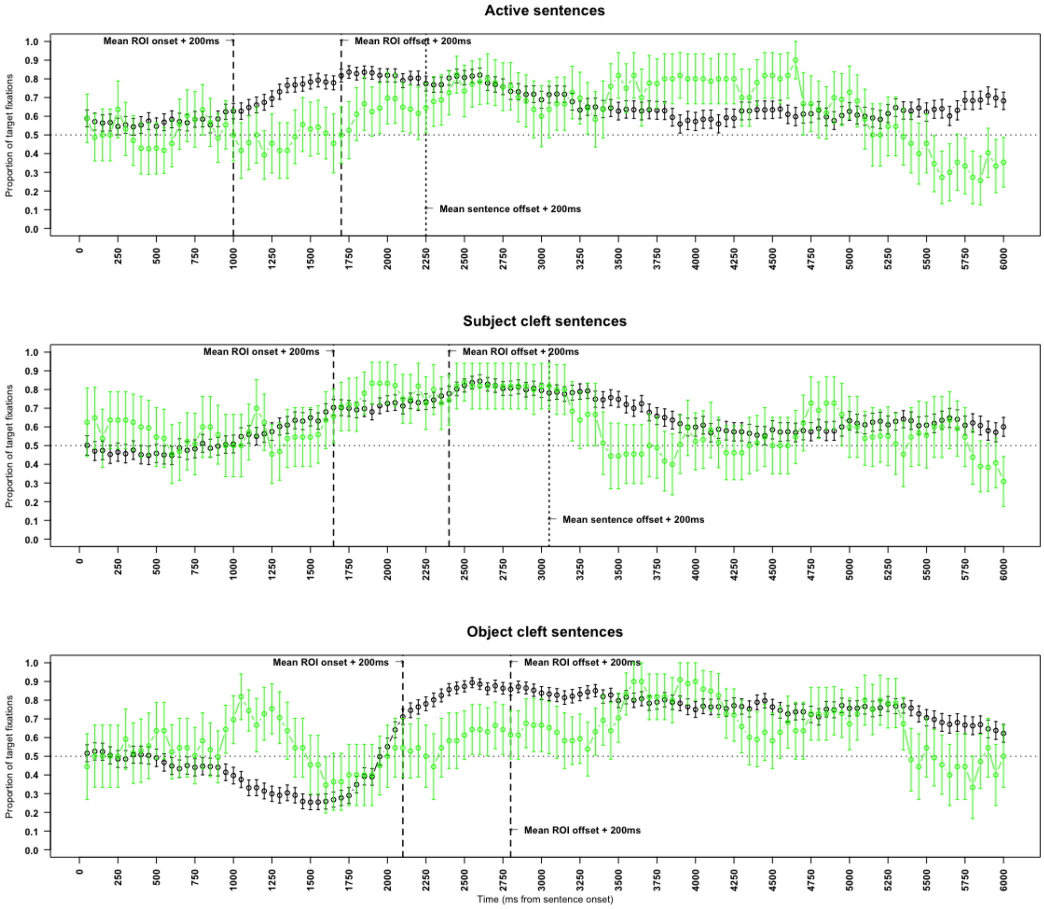


Figure 4: AD1's gaze patterns (green) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom.

Participant AD2:

On the active sentence condition, AD2's gaze data indicates that he was slightly slower than HC participants, but his PLT is high. On the SC condition, AD2 was very similar to the HC participants; at the offset of the ROI, his PLT was as high as 90%. On the OC condition, he was slower than the HC participants, but he followed the same general pattern with an initial preference for the distractor image, indicating an agent-first parsing strategy. On the offline data, AD2 fared quite well, pointing to the wrong image only once for each sentence condition.

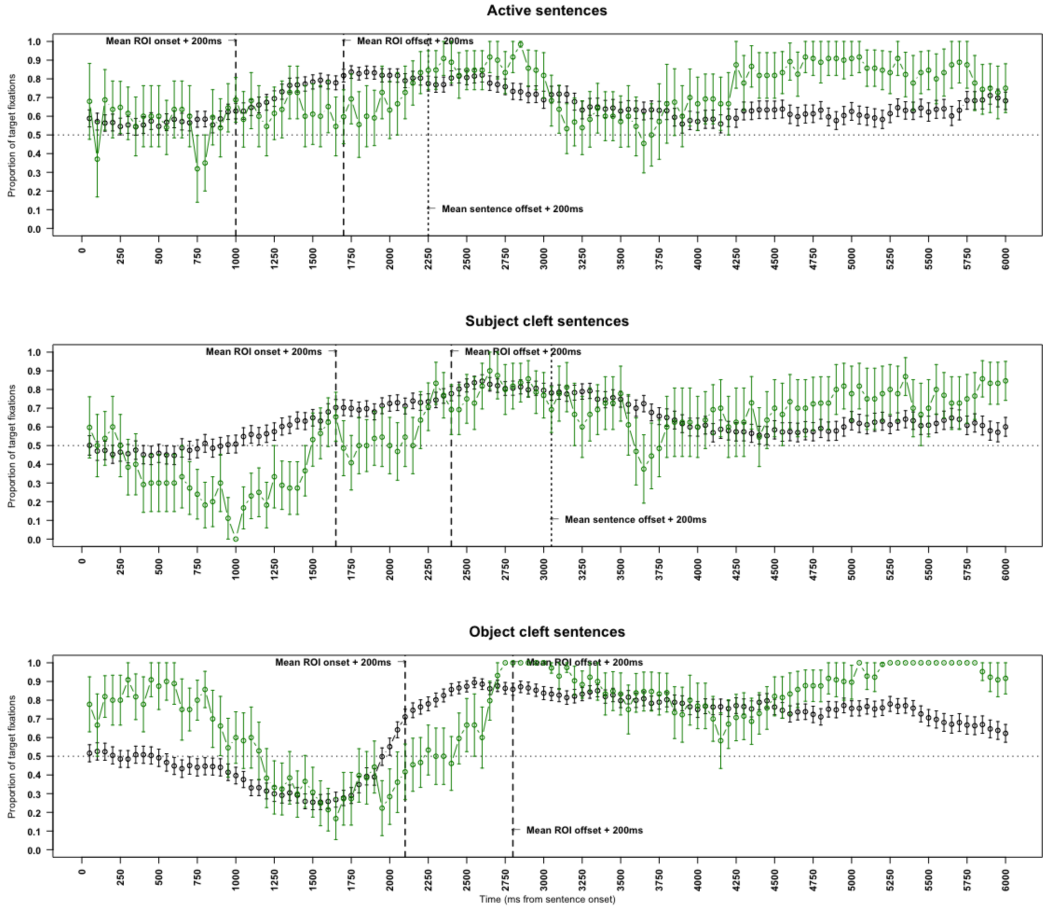


Figure 5 AD2 's gaze patterns (green) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom

Participant AD3:

Participant AD3 made no mistakes on the offline judgment task. However, the eye tracking data show some uncertainty for the A sentence type, but with an initial preference for the target image. PLT reached about 60% during the ROI, before it dropped to a low 10% after sentence offset. However, not long after, there was another spike in looks to target, and PLT rises to almost 70%. This could indicate the need for consciously examining both pictures before making a selection. We call this pattern *Conscious inhibition*. On the SC and OC sentences, AD3 was slower to detect target than HC participants were, but the general gaze patterns were comparable.

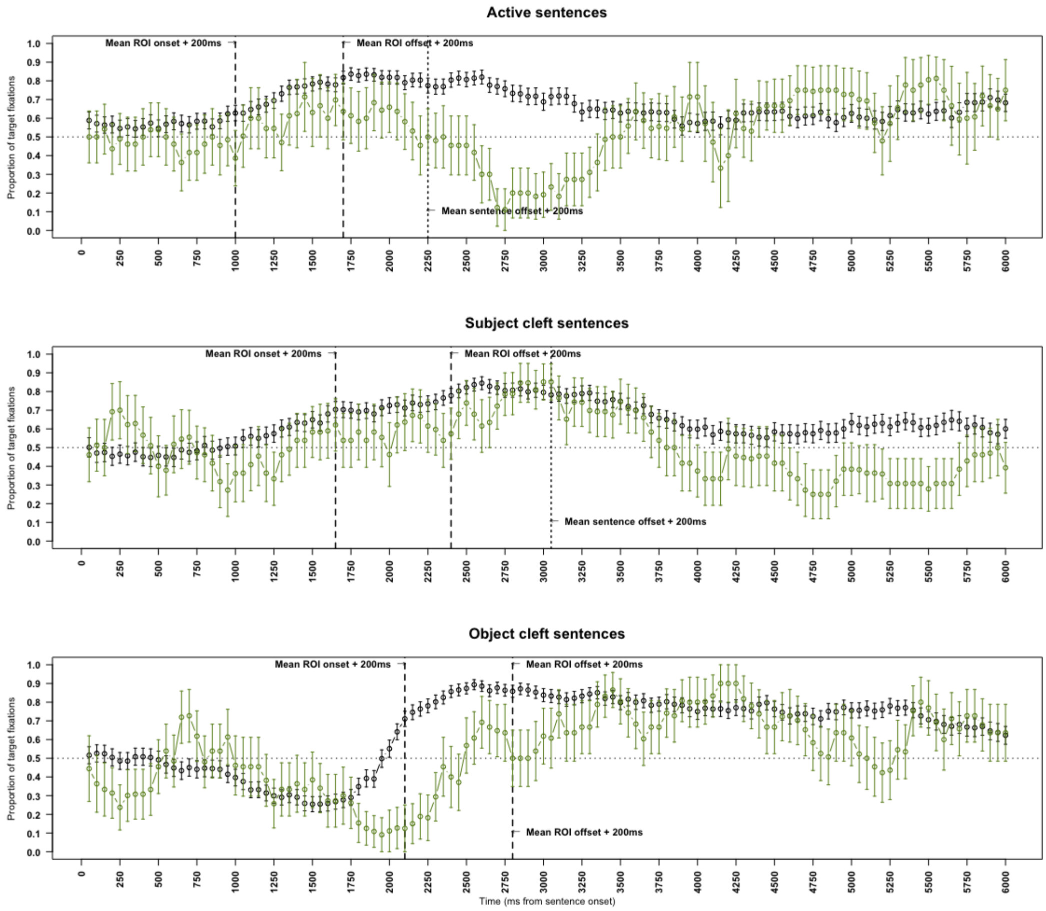


Figure 6: AD3's gaze patterns (green) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom

Participant AD4:

Participant AD4 showed a pattern that was quite similar to control participants, albeit somewhat slower for A and SC sentences. On the OC condition, AD4's gaze pattern did not show the initial preference for the target, but PLT stayed above chance level. Accuracy on the offline task was good overall, with one mistake each on the A and the OC conditions.

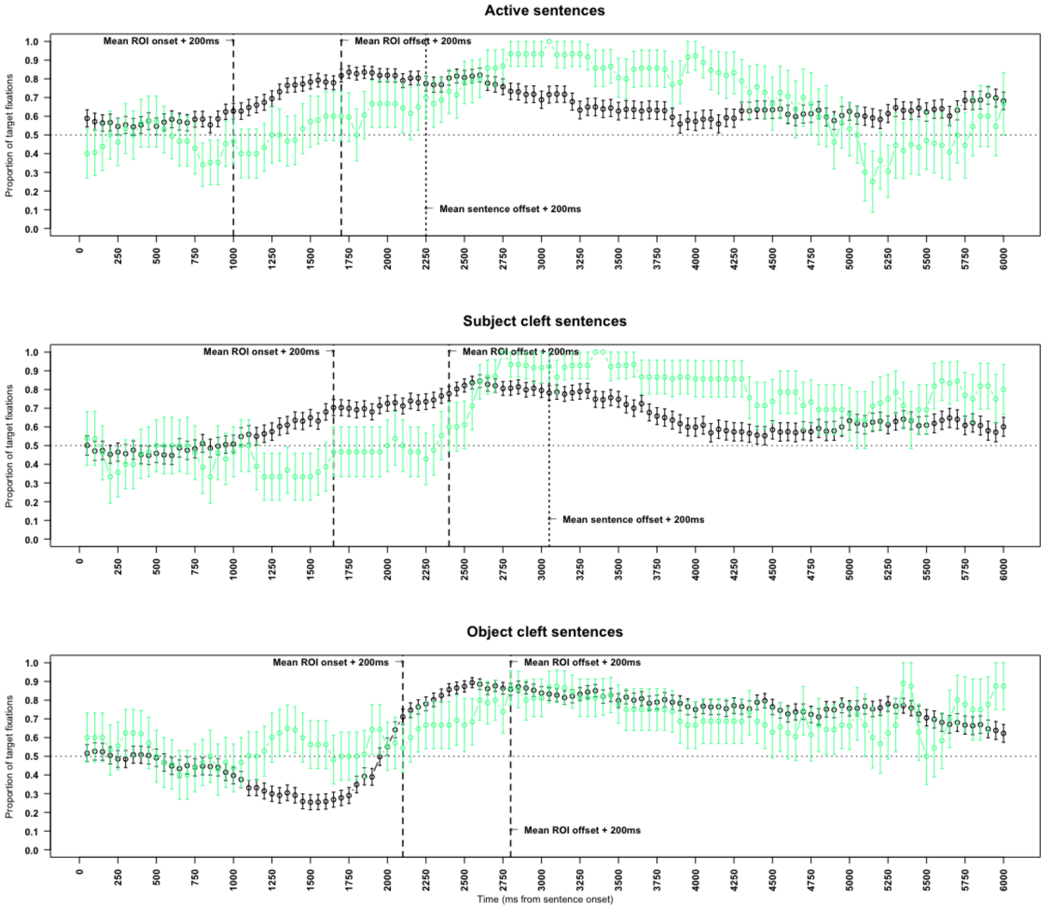


Figure 7: AD4's gaze patterns (green) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom

Participant AD5:

AD5 was the slowest of the participants with AD in all conditions. However, her accuracy scores on the offline task were high (100%), which indicated that comprehension is slow but correct. On the OC condition, she spent much time going back and forth between the target and distractor images initially, before showing a preference for the target only after sentence offset. This can look like the conscious inhibition pattern mentioned for AD3, but it is more likely a general slowing of almost 3 seconds compared to the HC group, based on her slowed processing also for A sentences.

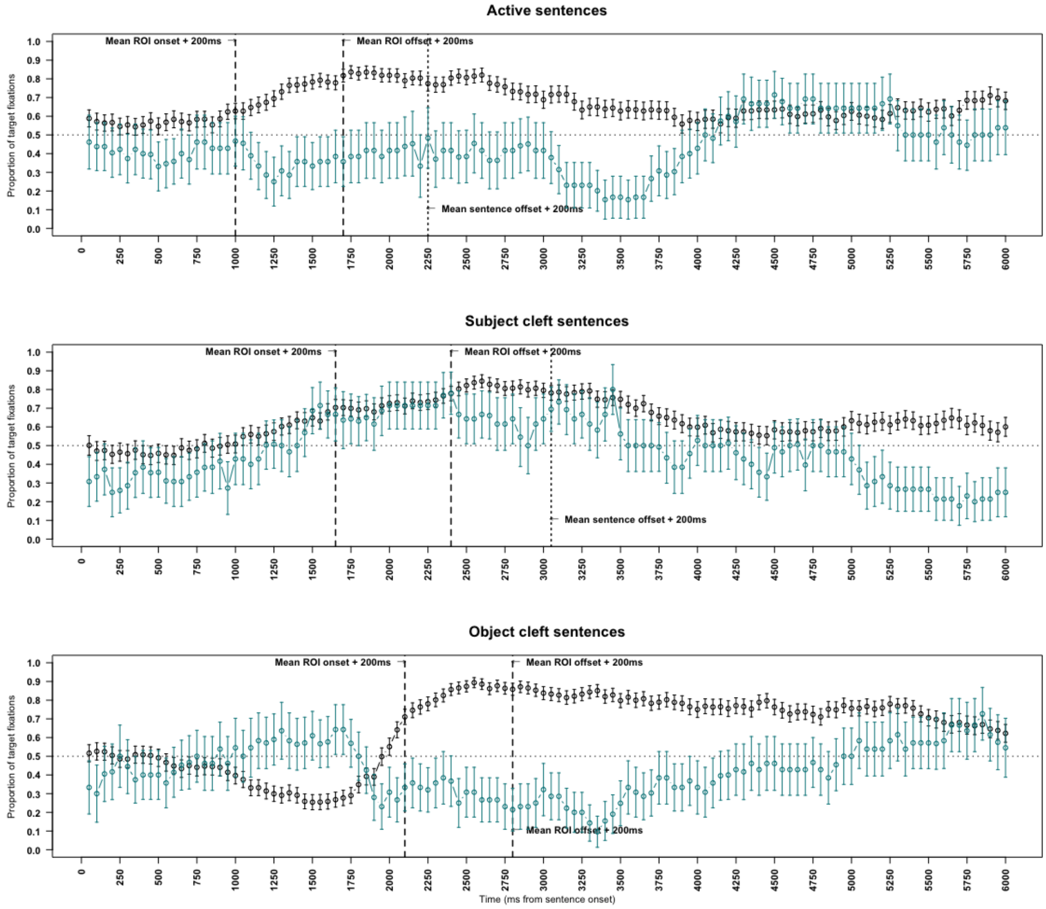


Figure 8: AD5's gaze patterns (green) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom

Participant PPA1:

On all three sentence structures, PPA1 showed uncertainties about which picture to focus on, showing conscious inhibition in all conditions. However, in all three conditions, there was a preference for the target initially. After an initial high PLT there was a drop, and another spike in PLT. His accuracy scores also reflect this uncertainty, as he made a few mistakes on all conditions on the offline task.

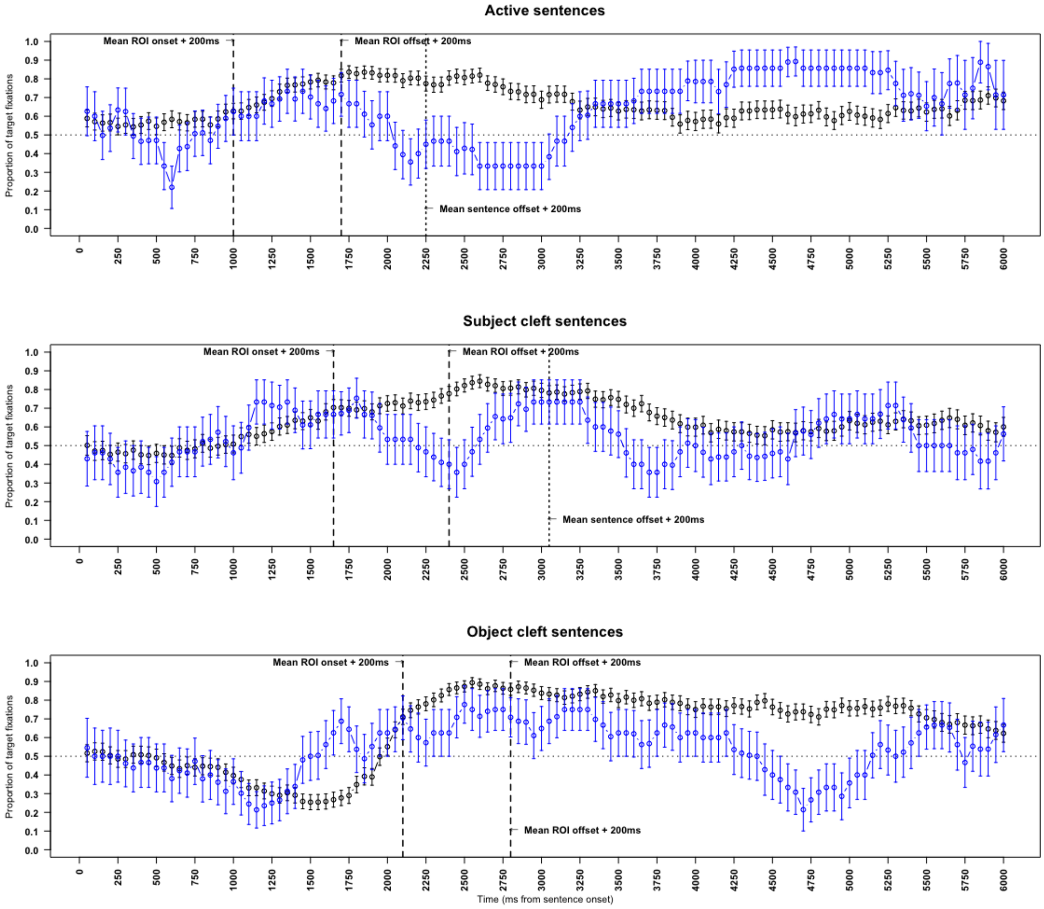


Figure 9: PPA1's gaze patterns (blue) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom

Participant PPA2:

The participant who showed a pattern most similar to the HC group overall was PPA2. He was fast at recognizing the target (at ROI offset), and the PLT reached a high peak of almost 100% for both A and SC conditions. It is interesting to note that PPA2 scored high on these two sentence conditions, both online and offline although his scores on the linguistic and cognitive tests were among the lowest for all participants. On the OC condition, however, PPA2’s gaze data show more uncertainty. Similar to the HC group he showed an initial preference for the distractor, before shifting gaze towards the target during the ROI. However, he never reached more than 65-70% PLT, and the proportion stayed low throughout. His accuracy on this condition is also poor (4 mistakes), proving that OC sentences were more difficult for PPA2 than the other conditions.

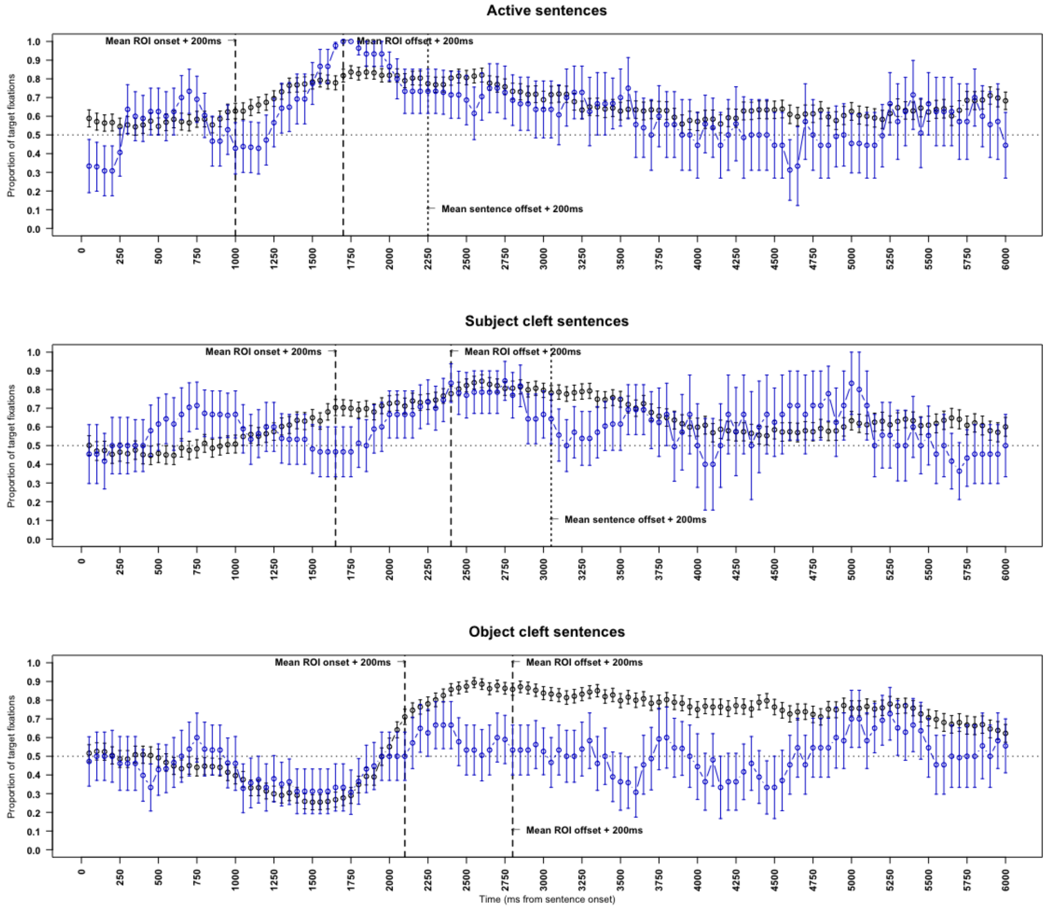


Figure 10: PPA2's gaze patterns (blue) compared to HC (black) participants. ACTIVE sentences on top, SUBJECT CLEFT in the middle and OBJECT CLEFT at the bottom

Summary of online results

For participants with dementia, we found three general processing patterns: 1) Sentence processing was *similar* for participants with AD/PPA and healthy controls in terms of PLT and time to detect target, 2) Sentence comprehension for persons with AD/PPA was *slower* than for healthy controls, but the PLT was comparable, and 3) Participants with AD/PPA showed uncertainty about the target picture, and gazes fluctuated between the two images before gazes were locked at target. We call this pattern *conscious inhibition* since the participants clearly need to properly inspect both images before making a selection, even though they had 3500ms to familiarize with images before the onset of the stimulus sentences.

For the A sentences, participant PPA2 was the only one who was similar to HC participants, whereas AD1, AD2, AD4 and AD5 were all slower than the HC group. AD3 and PPA1 showed the hesitant pattern.

In the SC condition, all participants with AD, as well as PPA2 exhibited a gaze pattern that was similar to the one for HC participants. PPA1 followed the same hesitant pattern as for A sentences, moving between target and distractor. The subjunction *som* is a strong cue for agency, since it is obligatory in subject cleft contexts, furthermore the amount of information in the first clause may provide enough context information for all participants to be more confident about their decision. This may explain why this structure is easier to comprehend than the two others.

On the OC condition, PPA2 and AD2 were similar to HC participants, AD3, AD4 and AD5 were slower than HC, and PPA1 and AD1 showed a pattern of uncertainty. Table 7 summarises the gaze data per participant on the different conditions.

Table 7: Gaze data per participant with dementia per sentence condition

	AD1	AD2	AD3	AD4	AD5	PPA1	PPA2
Active	Slower than HC	Slower than HC	conscious inhibition	slower than HC	slower than HC	conscious inhibition	similar to HC
Subject cleft	similar to HC	similar to HC	slower than HC	slower than HC	slower than HC	conscious inhibition	similar to HC
Object cleft	slower than HC	slower than HC	slower than HC	slower than HC	slower than HC	conscious inhibition	similar to HC

Discussion

We tested participants with AD and PPA, and neurologically healthy participants on sentences with canonical and non-canonical word order, to answer the following questions: (1) Are there differences in auditory comprehension accuracy and processing strategies applied to canonical and non-canonical sentences between AD and PPA and neurologically healthy controls?; (2) Is there a relationship between auditory sentence comprehension and cognitive functions in AD and PPA?; and (3) Can sentence comprehension tasks help distinguish between AD and PPA? We use the next few paragraphs to answer each of these questions separately.

Are there differences in auditory comprehension accuracy and processing strategies applied to canonical and non-canonical sentences between AD and PPA and neurologically healthy controls?

The control group, however small, showed a uniform pattern with little individual variation for each of the three sentence types on the online data. In the A and SC conditions we saw a pattern where the proportion of looks to target (PLT) spikes early during, or slightly after, the region of interest (ROI). This indicates that by the time the verb is presented, the participants' gazes are mainly focused on the target image. On the OC condition, the gaze pattern showed an initial preference for the distractor image, until the onset of the ROI. All participants show this preferred agent-first parsing strategy. Once it becomes clear that the first role introduced is not the agent but the theme, gazes shift quickly to the target image. There main difference between HC and participants with dementia was mainly found in the time it took to detect target.

The *Derived order processing hypothesis* (DOP-H) suggest that the sentence types which include movement are more difficult to process than sentences without overt movement. The DOP-H predicts chance level performance for sentences where the theme precedes the agent, such as in the OC sentences in the current experiment (Bastiaanse & van Zonneveld, 2006; Jap et al., 2016). The participants in this current study all performed above chance level, in terms of accuracy, on the OC (non-canonical) condition.

Contrary to models that assume that language users rely on syntactic movement in comprehension, the constraint-based model (Maryellen C. MacDonald & Seidenberg, 2006; Wells, Christiansen, Race, Acheson, & MacDonald, 2009) suggests that people make use of a vast amount of probabilistic information available in the linguistic signal. Language users arrive at a particular interpretation over another during sentence processing by rapidly integrating probabilistic constraints. For instance, in the OC condition, the initial fixations are directed to the distractor, indicating that the participants parse the sentence based on the most frequent pattern in Norwegian, which is one where the agent precedes the theme. Once the second noun, and later the verb of the relative clause are introduced, participants understand that the initial agent-first pattern is broken in this particular construction, and the looks to distractor are suppressed. This is similar for all participants in the current study, both controls and persons with AD and PPA.

On the offline task, the sentence comprehension difficulties for both AD and PPA seem minimal. No participant makes more than six mistakes in total (out of 45), and the mistakes were in most cases similar across the different conditions. Based on the offline task alone, there is no evidence to claim that these participants have a sentence impairment.

Most participants with AD showed better offline judgment than online processing. The participants with PPA show similar performance in online and offline processing, however their performance is different from each other: For PPA1 this can be seen in the uncertainty in his gaze data on all conditions, and the similar number of mistakes on all sentence conditions on the offline task. PPA2 also performs similarly on the online and offline measures, with more mistakes and longer times to detect target on the OC condition. This indicates that there are differences both between individuals and groups which cannot be captured by online or offline methods alone. Although some of our participants (at least in one of two conditions) showed slower than "normal" or deviant online processing, all participants showed above-chance offline (accuracy) comprehension performance, which was comparable to the offline performance exhibited by the control group. It appears, thus, that offline performance is determined by online processing routines.

Contrary to what can be expected, the A sentence condition was not the simplest structure for all participants. In fact, the SC condition elicited least mistakes on the offline task for all participants with dementia. Furthermore, all participants with dementia, except PPA1, exhibited similar gaze patterns as the HC group on this condition. Some participants with dementia were slower to detect the target than the HC participants, but since their gaze profiles overlapped in temporality, this difference was not significant. Better performance on the SC compared to the A condition in both measures may be due to the extra amount of context available in the matrix clause (*It is the [agent] who...*). There might be two reasons for this; 1) the matrix clause adds more time before the onset of the ROI, and 2) the subjunction *som* helps to condition who is the agent of the sentence. A combination of these two factors may facilitate comprehension. In contrast, more of the participants with dementia are slower on the A sentence condition on the online measure, and also make mistakes on the offline judgement task, which might be due to the short interval before the onset of the sentence and the onset of the ROI. However, the number of mistakes are marginal, and not indicative of a sentence comprehension impairment.

Is there a relationship between auditory sentence comprehension and cognitive functions in AD and PPA?

The participants with the lowest MMSE scores (AD3, AD4, AD5 and PPA1) also have gaze profiles which deviates most from the HC group. This may indicate that global cognitive functioning affects speed, and to some degree automatic processing for the offline judgement. These participants had good offline judgment, but their gaze pattern showed slower processing, and clear conscious gaze shifts between target and distractor, especially for participants PPA1 and AD3. Participants with lower MMSE scores need to assess both images before making a conscious decision about which image is the target and which is the distractor. This uncertainty would not have been visible by means of only offline methods.

Participants AD3, AD5 and PPA1 also show poor verb knowledge, as measured with the Semantische Associatie Test (Visch-Brink, Stronks, & Denes, 2005), and all three perform poorer on the online than on the offline measures. This may indicate that their automatic processing is influenced by poor verb comprehension.

Furthermore, participants AD1 and PPA1 both have slightly impaired executive functioning (as measured by TMT and Stroop), and working memory (digits backwards and ordering), but have good short-term memory (digits forward). These two participants make more mistakes on the “simpler” sentence structures A and SC compared to the other participants. However, since there are only two participants who respond this way, this may be coincidental and not indicate a pattern. The difference between conditions is more pronounced for AD1 than for PPA1.

AD1’s responses, which were poorest for active sentences, and best for the object cleft sentences, is similar to a finding by Zimmerer, Dabrowska et al. (2014) for one participant with lvPPA who was more impaired on active sentences than on passives. The authors argue that this cannot be explained by theories that assume overt syntactic movement, or even by working memory (WM) decline, as both movement-models and WM impairments would assume that a participant who can comprehend structures with movement, which places a heavy load on WM, would also be able to comprehend simple active structures. One possible explanation can be found in construction grammar accounts of how one, or a set of constructions can be selectively impaired, since constructions can be partially or completely lexicalized. Depending on usage-related factors such as frequency, listeners can find some constructions easier to process than others (Gahl et al., 2003; Zimmerer, Dąbrowska, Romanowski, Blank, & Varley, 2014).

Five of the seven participants with dementia rely on the agent-first parsing strategy, albeit to different degrees, indicating that this strategy is retained even when cognitive capacities are impaired. This may imply that linear processing is automatic and necessary for interpretation of constructions with little or no morphological cues.

How can results from sentence comprehension studies including participants with AD and PPA help distinguish between the two diseases?

There was not one pattern that was similar for both participants with PPA, and one that is similar for all participants with AD. However, we identify some similarities; the participants with AD were in general slower than the HC group, except on the SC condition. Furthermore, the participants with PPA show two different patterns on both online and offline tasks. PPA2 was consistently similar to the HC group on all three conditions on the ET data, and PPA1, on the other hand, showed confusion on all sentence types. Thus, it is not possible to distinguish AD and PPA based on these sentence comprehension data.

Both on the cognitive and the linguistic tests, PPA1 was more similar to participants with AD than he was to PPA2. This can be seen in relation to their diagnoses: PPA1 was initially diagnosed with AD, before it was changed to PPA. PPA2's diagnosis seems to be a more clear instance of lvPPA, with phonological short-term memory impairments and poor confrontation naming skills (Gorno-Tempini et al., 2011).

However, an alternative view is that PPA1 and PPA2 show profiles that can be related to two different types of lvPPA; one variant which is resembles atypical AD, and one which follows the originally proposed criteria for lvPPA (Marshall et al., 2018; Matias-Guiu et al., 2018; Vandenberghe, 2016).

Limitations and future research

An obvious methodological shortcoming of the current study is the low number of participants. Few participants were recruited initially, and several were excluded due to difficulties with calibration of the eye tracking apparatus. Two participants were impossible to track (the eye tracker could not find their eyes), and three participants had very poor tracking ratio, probably due to glasses, eye make-up and heavy eye-lids which throw shadow on the cornea, making it difficult to obtain good eye tracking data.

The two participants in this study who are diagnosed with PPA show different profiles on several of the linguistic and cognitive tests, which may be related to different underlying neurological change for the two participants. It might also mirror recent research on lvPPA, which suggests that lvPPA is not in fact one, but rather two different sub-types of PPA (Matias-

Guiu 2019, Vanderberghe, 2016). More research is needed on persons with PPA to corroborate these findings.

An interesting observation is that AD1 seems to have more difficulties with the canonical sentence conditions than the non-canonical one. This has been found for at least one other person with dementia (Zimmerer et al., 2014; Zimmerer & Varley, 2015), but it is uncertain how common this pattern is globally due to the few studies of sentence comprehension in dementia. More research is needed to identify the theoretical and clinical reasons behind this atypical pattern.

Conclusion

Eye movements are closely linked to focus of attention (Tanenhaus et al., 1995). In all conditions in the current experiment, we see that participants do focus on the target image quite soon after the onset of the ROI. Even the participants who show signs of conscious inhibition have an initial preference for the target image before they move to the distractor. Listeners display an anticipatory assignment of thematic roles while completing sentence resolution tasks (Kamide, Altmann, & Haywood, 2003; Knoeferle, Crocker, Scheepers, & Pickering, 2005). This pattern is clearly seen on the OC sentences; due to the expectation of a sentence structure where the agent precedes the theme, participants look to the image where the person who performs the action of the sentence corresponds to the first-mentioned participant in the sentence. This expectation of agent-first is based on the listener's experience with this type of structures, and experience is a powerful factor in sentence comprehension expertise (Wells et al., 2009). This can explain why the A and SC sentences are equally easy to comprehend for most participants, despite the vast difference in frequency between the constructions.

We find that sentence comprehension in AD and PPA is less impaired than reported in the literature (Croot et al., 1999; Kempler et al., 1998; Rochon et al., 1994; Waters et al., 1998). The novelty of this study is that we combine online and offline measures within the same task to identify sentence comprehension difficulties for persons with dementia. We see that most of the participants with AD perform better on the offline than on the online measures, indicating that their automatic processing is slower than for neurologically healthy participants. Participants with PPA perform similarly on the online and the offline measures for each sentence condition. However, their profiles are different from each other. This can either be due to individual variation, or it might reflect different types of lvPPA.

Sentence comprehension needs to be assessed during dementia screening to make sure that persons with objective or subjective memory impairments actually understand what is being asked of them on other tasks. We argue that sentence comprehension tasks can be a useful addition for assessment of functional behaviour for persons with AD and PPA. Since all standardized testing in the specialist health care system requires a certain level of language comprehension, it is important to know that the patients do in fact have the ability to understand what is being asked of them. Furthermore, for research purposes, both online and offline methods should be implemented to get a more comprehensive view of possible impairments.

Acknowledgments:
We wish to extend our gratitude to Linda Vatne and Hanna Eidshagen Menne for assistance with data collection and coding; Björn Lundquist for help with building a script for the eye tracking data; Valantis Fyndanis, Minna Lehtonen, Monica Norvik and Hanne Gram Simonsen for feedback on a previous version of the manuscript; and The Center of Multilingualism in Society Across the Lifespan (MultiLing) at the University of Oslo for generous access to their lab facilities. We also want to acknowledge the Norwegian register of persons assessed for cognitive symptoms (NorCog), for providing access to patient data, and to Anne-Brita Knapskog at the Oslo University Hospital for help with recruitment of participants. Helge Dyvik at the University of Bergen helped us searching in the IESS corpus.

Declaration of interest:

The authors report no conflict of interest.

Appendix

Verbs used in the sentence construction:

Target verb Norwegian	**English translation**
Bite	Bite
Bære	Cary
Dytte	Push
Filme	Film
Fotografere	Photograph
Klore	Scratch
Klype	Pinch
Kvele	Strangle
Kysse	Kiss
Male	Paint
Redde	Save
Slå	Hit/punch
Sparke	Kick
True	Threaten
Vaske	Wash/clean

Full list of tested sentences in each session:

Set 1:

:e type	Norwegian sentence	English translation
A	Jenta bærer mannen	The girl carries the man
A	Jenta dytter gutten	The girl pushes the boy
A	Mannen klyper damen	The man pinches the woman
A	Jenta kysser mannen	The girl kisses the man
A	Mannen slår damen	The man hits the woman
SC	Det er damen som vasker barnet	It is the woman who washes the child
SC	Det er hesten som sparker kua	It is the horse who kicks the cow
SC	Det er jenta som filmer gutten	It is the girl who films the boy
SC	Det er jenta som klorer gutten	It is the girl who scratches the boy
SC	Det er damen som truer mannen	It is the woman who threatens the man
OC	Det er gutten jenta kveler	It is the boy the girl strangles
OC	Det er kua hesten biter	It is the cow the horse bites
OC	Det er mannen damen redder	It is the man the woman saves
OC	Det er gutten jenta fotograferer	It is the boy the girl photographs
OC	Det er gutten jenta maler	It is the boy the girl paints

Set 2:

Sentence type	Norwegian sentence	English translation
A	Mannen redder damen	The man saves the woman
A	Mannen truer damen	The man threatens the woman
A	Gutten klorer jenta	The boy scratches the girl
A	Hesten biter kua	The horse bites the cow
A	Gutten maler jenta	The boy paints the girl
SC	Det er mannen som bærer damen	It is the man who carries the woman
SC	Det er gutten som kveler jenta	It is the boy who strangles the girl
SC	Det er gutten som fotograferer jenta	It is the boy who photographs the girl
SC	Det er jenta som kysser mannen	It is the girl who kisses the man
SC	Det er jenta som klyper gutten	It is the girl who pinches the boy
OC	Det er barnet damen dytter	It is the child the woman pushes
OC	Det er barnet damen vasker	It is the child the woman washes
OC	Det er hesten kua sparker	It is the horse the cow kicks

OC	Det er barnet damen slår	It is the child the woman hits
OC	Det er damen mannen filmer	It is the woman the man films

Set 3:

Sentence type	**Norwegian sentence**	**English translation**
A	Damen vasker barnet	The woman washes the child
A	Damen kveler mannen	The woman strangles the man
A	Jenta fotograferer gutten	The girl photographs the boy
A	Kua sparker hasten	The cow kicks the horse
A	Mannen filmer damen	The man films the woman
SC	Det er mannen som redder damen	It is the man who saves the woman
SC	Det er gutten som dytter jenta	It is the boy who pushes the girl
SC	Det er gutten som maler jenta	It is the boy who paints the girl
SC	Det er damen som slår mannen	It is the woman who hits the man
SC	Det er jenta som biter gutten	It is the girl who bites the boy
OC	Det er jenta gutten klyper	It is the girl the boy pinches
OC	Det er mannen barnet kysser	It is the man the child kisses
OC	Det er damen mannen truer	It is the woman the man threatens
OC	Det er barnet mannen klorer	It is the child the man scratches
OC	Det er jenta mannen bærer	It is the girl the man carries

References

Ahmed, S., Jager, C. A. de, Haigh, A.-M. F., & Garrard, P. (2012). Logopenic aphasia in Alzheimer's disease: Clinical variant or clinical feature? *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(11), 1056–1062. https://doi.org/10.1136/jnnp-2012-

Ahmed, S., Jager, C. A. de, Haigh, A.-M. F., & Garrard, P. (2012). Logopenic aphasia in Alzheimer's disease: Clinical variant or clinical feature? *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(11), 1056–1062. https://doi.org/10.1136/jnnp-2012-302798

Almor, A., Kempler, D., MacDonald, M. C., Andersen, E. S., & Tyler, L. K. (1999). Why Do Alzheimer Patients Have Difficulty with Pronouns? Working Memory, Semantics, and Reference in Comprehension and Production in Alzheimer's Disease. *Brain and Language*, *67*(3), 202–227. https://doi.org/10.1006/brln.1999.2055

Bastiaanse, R., Lind, M., Moen, I., & Simonsen, H. G. (2006). *Verb- og Setningstesten*. Novus. /z-wcorg/.

Bastiaanse, R., & van Zonneveld, R. (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia. *Journal of Neurolinguistics*, *18*(1), 57–66. https://doi.org/10.1016/j.jneuroling.2004.11.006

Bastiaanse, R., & van Zonneveld, R. (2006). Comprehension of passives in Broca's aphasia. *Brain and Language*, *96*(2), 135–142; discussion 157-170. https://doi.org/10.1016/j.bandl.2005.06.012

Bayles, K. A., & Tomoeda, C. K. (2007). *Cognitive-Communication Disorders of Dementia* (1 edition). San Diego: Plural Publishing Inc.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Caplan, D. (1996). Syntactic Processing in Sentence Comprehension Under Dual-task Conditions in Aphasic Patients. *Language and Cognitive Processes*, *11*(5), 525–551. https://doi.org/10.1080/016909696387141

Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging*, *26*(2), 439–450. https://doi.org/10.1037/a0021837

Charles, D., Olm, C., Powers, J., Ash, S., Irwin, D. J., McMillan, C. T., … Grossman, M. (2014). Grammatical comprehension deficits in non-fluent/agrammatic primary progressive aphasia. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(3), 249–256. https://doi.org/10.1136/jnnp-2013-305749

Croot, K., Hodges, J. R., & Patterson, K. (1999). Evidence for impaired sentence comprehension in early Alzheimer's disease. *Journal of the International Neuropsychological Society: JINS*, *5*(5), 393–404.

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, *28*(6), 604–615. https://doi.org/10.1016/j.langsci.2005.11.014

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS): Examiner's manual: Flexibility of thinking, concept formation, problem solving, planning, creativity, impluse control, inhibition*. San Antonio (Tex.): Pearson. /z-wcorg/.

Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice* (2nd ed.). Retrieved from https://www.springer.com/gp/book/9781846286087

Egeland, J., Landrø, N. I., Tjemsland, E., & Walbækken, K. (2006). Norwegian Norms and Factor-Structure of Phonemic and Semantic Word List Generation. *The Clinical Neuropsychologist*, *20*(4), 716–728. https://doi.org/10.1080/13854040500351008

Faarlund, J. T. (1997). *Norsk referansegrammatikk*. Oslo: Universitetsforl.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*(2), 164–203. https://doi.org/10.1016/S0010-0285(03)00005-7

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.

Gahl, S., & Menn, L. (2016). Usage-based approaches to aphasia. *Aphasiology*, *30*(11), 1361–1377. https://doi.org/10.1080/02687038.2016.1140120

Gahl, S., Menn, L., Ramsberger, G., Jurafsky, D. S., Elder, E., Rewega, M., & Audrey, L. H. (2003). Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. *Brain and Cognition*, *53*(2), 223–228.

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., … Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Grossman, M., & White-Devine, T. (1998). Sentence Comprehension in Alzheimer’s Disease. *Brain and Language*, *62*(2), 186–201. https://doi.org/10.1006/brln.1997.1898

Henry, M. L., & Gorno-Tempini, M. L. (2010). The logopenic variant of primary progressive aphasia. *Current Opinion in Neurology*, *23*(6), 633–637. https://doi.org/10.1097/WCO.0b013e32833fb93e

Howard, D., Patterson, K. E., & Thames Valley Test Company. (1992). *The pyramids and palm trees test.* Thames Valley Test Co. /z-wcorg/.

Jap, B. A., Martinez-Ferreiro, S., & Bastiaanse, R. (2016). The effect of syntactic frequency on sentence comprehension in standard Indonesian Broca’s aphasia. *Aphasiology*, *30*(11), 1325–1340. https://doi.org/10.1080/02687038.2016.1148902

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156. https://doi.org/10.1016/S0749-596X(03)00023-8

Kavé, G., & Levy, Y. (2004). Preserved Morphological Decomposition in Persons With Alzheimer's Disease. *Journal of Speech, Language, and Hearing Research*, *47*(4), 835–847. https://doi.org/10.1044/1092-4388(2004/062)

Kempler, D., Almor, A., & MacDonald, M. C. (1998). Teasing Apart the Contribution of Memory and Language Impairments in Alzheimer's DiseaseAn Online Study of Sentence Comprehension. *American Journal of Speech-Language Pathology*, *7*(1), 61–67. https://doi.org/10.1044/1058-0360.0701.61

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127. https://doi.org/10.1016/j.cognition.2004.03.002

MacDonald, M. C., Almor, A., Henderson, V. W., Kempler, D., & Andersen, E. S. (2001). Assessing Working Memory and Language Comprehension in Alzheimer's Disease. *Brain and Language*, *78*(1), 17–42. https://doi.org/10.1006/brln.2000.2436

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54; discussion 55-74.

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 581–611). Academic Press.

Marková, J., Horváthová, Ľ., Králová, M., & Cséfalvay, Z. (2016). Sentence comprehension in Slovak-speaking patients with Alzheimer's disease. *International Journal of Language & Communication Disorders*, n/a-n/a. https://doi.org/10.1111/1460-6984.12284

Marshall, C. R., Hardy, C. J. D., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D., … Warren, J. D. (2018). Primary progressive aphasia: A clinical approach. *Journal of Neurology*, *265*(6), 1474–1490. https://doi.org/10.1007/s00415-018-8762-6

Matias-Guiu, J. A., Díaz-Álvarez, J., Ayala, J. L., Risco-Martín, J. L., Moreno-Ramos, T., Pytel, V., … Cabrera-Martín, M. N. (2018). Clustering Analysis of FDG-PET Imaging in Primary Progressive Aphasia. *Frontiers in Aging Neuroscience*, *10*. https://doi.org/10.3389/fnagi.2018.00230

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics*, *53*(4), 372–380. https://doi.org/10.3758/BF03206780

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., … Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165.

Müller, N. (2013). Dementia. In *The Handbook of Speech and Language Disorders*. Chichester, West Sussex: Wiley-Blackwell Publishing Ltd.

Price, C. C., & Grossman, M. (2005). Verb agreements during on-line sentence processing in Alzheimer's disease and frontotemporal dementia. *Brain and Language*, *94*(2), 217–232. https://doi.org/10.1016/j.bandl.2004.12.009

R core team. (2019). R-Project (Version 3.6.0).

Ribu, I. S., Norvik, M. I., Lehtonen, M., & Simonsen, H. G. (submitted). *Free word associations in Alzheimer's disease and Primary Progressive Aphasia*.

Rochon, E., Waters, G. S., & Caplan, D. (1994). Sentence Comprehension in Patients with Alzheimer's Disease. *Brain and Language*, *46*(2), 329–349. https://doi.org/10.1006/brln.1994.1018

Rosén, V., De Smedt, K., Meurer, P., & Dyvik, H. (2012). An open infrastructure for advanced treebanking. In J. Hajič, K. De Smedt, M. Tadić, & A. Branco (Eds.), *Meta-research Workshop on Advanced Treebanking at LREC2012*.

Sajjadi, S. A., Patterson, K., Tomek, M., & Nestor, P. J. (2012). Abnormalities of connected speech in the non-semantic variants of primary progressive aphasia. *Aphasiology*, *26*(10), 1219–1237. https://doi.org/10.1080/02687038.2012.710318

Sekerina, I. A. (2014). Visual World Eye-Tracking Paradigm. In *Encyclopedia of Language Development* (pp. 658–658). https://doi.org/10.4135/9781483346441

Small, J. A., Kemper, S., & Lyons, K. (1997). Sentence comprehension in Alzheimer's disease: Effects of grammatical complexity, speech rate, and repetition. *Psychology and Aging*, *12*(1), 3–11.

Street, J. A. (2017). This is the native speaker that the non-native speaker outperformed: Individual, education-related differences in the processing and interpretation of Object Relative Clauses by native and non-native speakers of English. *Language Sciences*, *59*, 192–203. https://doi.org/10.1016/j.langsci.2016.10.004

Street, J. A., & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua*, *120*(8), 2080–2094. https://doi.org/10.1016/j.lingua.2010.01.004

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science (New York, N.Y.)*, *268*(5217), 1632–1634.

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203–214. https://doi.org/10.1016/S0887-6177(03)00039-8

Vandenberghe, R. (2016). Classification of the primary progressive aphasias: Principles and review of progress since 2011. *Alzheimer's Research & Therapy*. Retrieved from https://alzres.biomedcentral.com/articles/10.1186/s13195-016-0185-y

Visch-Brink, E., Stronks, D., & Denes, G. (2005). *Semantische Associatie test (SAT)*. Pearson.

Waters, G. S., & Caplan, D. (1997). Working Memory and On-Line Sentence Comprehension in Patients with Alzheimer's Disease. *Journal of Psycholinguistic Research*, *26*(4), 377–400. https://doi.org/10.1023/A:1025073902269

Waters, G. S., Rochon, E., & Caplan, D. (1998). Task Demands and Sentence Comprehension in Patients with Dementia of the Alzheimer's Type. *Brain and Language*, *62*(3), 361–397. https://doi.org/10.1006/brln.1997.1880

Wechsler, D. (2008). *WAIS-IV : Wechsler adult intelligence scale*. San Antonio, Tex.: Pearson. /z-wcorg/.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–271. https://doi.org/10.1016/j.cogpsych.2008.08.002

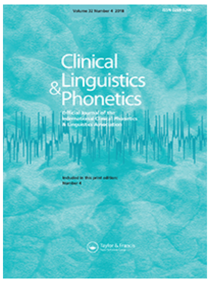
Wilson, S. M., Galantucci, S., Tartaglia, M. C., & Gorno-Tempini, M. L. (2012). The neural basis of syntactic deficits in primary progressive aphasia. *Brain and Language*, *122*(3), 190–198. https://doi.org/10.1016/j.bandl.2012.04.005

Zimmerer, V. C., Dąbrowska, E., Romanowski, C. A. J., Blank, C., & Varley, R. A. (2014). Preservation of passive constructions in a patient with primary progressive aphasia. *Cortex*, *50*, 7–18. https://doi.org/10.1016/j.cortex.2013.09.007

Zimmerer, V. C., & Varley, R. A. (2015). A case of "order insensitivity"? Natural and artificial language processing in a man with primary progressive aphasia. *Cortex*, *69*, 212–219. https://doi.org/10.1016/j.cortex.2015.05.006

Lind, M., Simonsen, H. G., Ribu, I. S., Svendsen, B. A., Svennevig, J., & de Bot, K. (2018). Lexical access in a bilingual speaker with dementia: Changes over time. *Clinical Linguistics & Phonetics*, *32*(4), 353–377. doi:10.1080/02699206.2017.1381168

IV



Clinical Linguistics & Phonetics

ISSN: 0269-9206 (Print) 1464-5076 (Online) Journal homepage: https://www.tandfonline.com/loi/iclp20

Lexical access in a bilingual speaker with dementia: Changes over time

Marianne Lind, Hanne Gram Simonsen, Ingeborg Sophie Bjønness Ribu, Bente Ailin Svendsen, Jan Svennevig & Kees de Bot

CLINICAL LINGUISTICS & PHONETICS
2018, VOL. 32, NO. 4, 353–377
https://doi.org/10.1080/02699206.2017.1381168

Check for updates

Lexical access in a bilingual speaker with dementia: Changes over time

Marianne Lind[a,b], Hanne Gram Simonsen[a], Ingeborg Sophie Bjønness Ribu[c], Bente Ailin Svendsen[a], Jan Svennevig[a], and Kees de Bot[d,e]

[a]Center for Multilingualism in Society across the Lifespan, University of Oslo, Norway; [b]Department of Speech and Language Disorders, Statped, Oslo, Norway; [c]Department of Linguistics and Scandinavian Studies, University of Oslo, Norway; [d]Department of Applied Linguistics, University of Groningen, The Netherlands; [e]Department of Applied Linguistics, University of Pannonia, Veszprém, Hungary

ABSTRACT
In this article, we explore the naming skills of a bilingual English-Norwegian speaker diagnosed with Primary Progressive Aphasia, in each of his languages across three different speech contexts: confrontation naming, semi-spontaneous narrative (picture description), and conversation, and at two points in time: 12 and 30 months post diagnosis, respectively. The results are discussed in light of two main theories of lexical retrieval in healthy, elderly speakers: the Transmission Deficit Hypothesis and the Inhibitory Deficit Theory. Our data show that, consistent with the participant's premorbid use of and proficiency in the two languages, his performance in his L2 is lower than in his L1, but this difference diminishes as the disease progresses. This is the case across the three speech contexts; however, the difference is smaller in the narrative task, where his performance is very low in both languages already at the first measurement point. Despite his word finding problems, he is able to take active part in conversation, particularly in his L1 and more so at the first measurement point. In addition to the task effect, we find effects of word class, frequency, and cognateness on his naming skills. His performance seems to support the Transmission Deficit Hypothesis. By combining different tools and methods of analysis, we get a more comprehensive picture of the impact of the dementia on the speaker's languages from an intra-individual as well as an inter-individual perspective, which may be useful in research as well as in clinical practice.

KEYWORDS
Bilingualism; dementia; lexical access; longitudinal study; primary progressive aphasia

Introduction

While decline in cognitive functions is a central feature in dementia, language problems are also recognised as a core clinical criterion in certain types of dementia, among them Alzheimer's disease (AD) (McKhann et al., 2011) and Primary Progressive Aphasia (PPA). In the latter case, language difficulty is the most prominent deficit at symptom onset (Gorno-Tempini et al., 2011). Word-finding problems, while also found in the healthy aging population (Burke & Shafto, 2004, 2008; Mortensen, Meyer, & Humphreys, 2006; Vogel-Eyny, Galletta, Gitterman, & Obler, 2016), seem to be among the earliest and most pervasive symptoms in these types of dementia, in AD

CONTACT Marianne Lind marianne.lind@statped.no Center for Multilingualism in Society across the Lifespan, University of Oslo, P.O. Box 1102 Blindern, 0317 Oslo, Norway.
This article was originally published with errors. This version has been corrected/amended. Please see Erratum (http://dx.doi.org/10.1080/02699206.2017.1402503).

both in earlier (Chen et al., 2001; Mickes et al., 2007; Nicholas, Obler, Au, & Albert, 1996) and later stages (Locascio, Growdon, & Corkin, 1995; Salmon, Heindel, & Lange, 1999), and in all subtypes of PPA, but most prominently in the semantic and logopenic subtypes (Gorno-Tempini et al., 2011; Grossman & Ash, 2004; Hilger, Ramsberger, Gilley, Menn, & Kong, 2014; Kempler & Goral, 2008; Wilson et al., 2010). Two recent review articles, one on AD (Kavé & Goral, 2017) and one on several neurodegenerative disorders including AD and PPA (Boschi et al., 2017), show that word retrieval problems are evident both in single word production and in connected speech in these groups, and point to the importance of using different cognitive and linguistic tasks in the assessment of persons with dementia. In the present longitudinal study, we investigate lexical access and word finding difficulties in a bilingual speaker with PPA across three different contexts of language use: confrontation naming, semi-structured narrative production, and conversation.

Models of word production, whether serial (Levelt, 1999, 2001) or interactive (Dell, 1986; Dell & O'Seaghdha, 1992), agree that word retrieval implies two different processing levels: semantic processing (word meaning) and phonological processing (word form). In confrontation naming, the picture or object first activates the concept, then the lemma with its semantic and grammatical information, and the phonological form of the word with its articulatory encoding. The two models are based on different theoretical frameworks and thus differ in their assumptions about semantic representations as well as spreading of activation – whether it is unidirectional from meaning to form, or interactive between the levels.

Concerning lexical retrieval in healthy, elderly speakers, two main theories have been proposed. The Transmission Deficit Hypothesis (TDH) (Burke, MacKay, Worthley, & Wade, 1991) assumes that the retrieval problems reflect a weakening of the connections in the lexical network. Semantic processing is less affected by weakening since there are many connections between the semantic nodes, but phonological processing is vulnerable since there is only one connection between the semantic representation and the phonological form of the actual word. The Inhibitory Deficit Theory (IDT) (Zacks & Hasher, 1997), on the other hand, suggests that the word retrieval problems reflect a weakening in the inhibitory processes of working memory, resulting in a reduced ability to suppress competing lexical alternatives.

In dementia, difficulties in lexical retrieval have mainly been attributed to problems at the semantic level. In AD patients, errors in confrontation naming are typically semantically related to the target item (Balthazar, Cendes, & Pereira Damasceno, 2008; Bayles & Tomoeda, 1983; Martin & Fedio, 1983; Moreaud, David, Charnallet, & Pellat, 2001; Obler & Albert, 1981), but whether the errors are the result of an underlying semantic impairment, or rather a result of impaired access to semantic information is still under discussion (Balthazar et al., 2008; Kavé & Goral, 2017; Kempler & Goral, 2008; Nicholas et al., 1996). In speakers with PPA, the different subtypes may reflect different underlying problems: While the semantic variant of PPA seems to reflect an underlying deficit in semantic memory, retrieval problems in the logopenic and nonfluent/agrammatic variants may rather be attributed to a limitation in phonological access in word production (Grossman & Ash, 2004; Kempler & Goral, 2008; Rogers, Ivanoiu, Patterson, & Hodges, 2006).

The study of language processing in bilingual[1] speakers is challenging. The list of confounding factors which may affect their linguistic performance is long and includes i.a. the age at which the

[1]We use the term bilingual in the broad sense, referring to speakers of two or more languages.

second (third, fourth, etc.) language was acquired (simultaneous vs. sequential bilingualism), the manner in which it was acquired (e.g. by immersion or instruction), the usage patterns of the different languages (e.g. in which contexts the various languages are used), the proficiency in different modalities (speaking, reading, writing, etc.) in the two or more languages, structural similarities or differences between the languages, and societal attitudes towards the different languages. Adding a pathological condition affecting language and communication to the picture makes it even more complex. Finding homogeneous groups of bilingual speakers for research is thus at best challenging and often impossible. It may even not be desirable since bilingualism is governed by the complementarity principle, which means that the languages of a bilingual speaker are usually acquired and used for different purposes, with different people, and in different domains of life (Grosjean, 1998).

Despite the lack of a general, unified definition of bilingualism, it is acknowledged that a bilingual speaker is not the same as two monolingual speakers in one mind (Grosjean, 1998). Thus, using monolingual norms as a comparison may not give a correct picture. Instead, it has been argued that the use of 'composite scoring' may be more sensible, at least when investigating language processing in balanced bilinguals (Goral, 2013:192). 'Composite scoring' of a naming test means that semantically correct responses are scored as correct irrespective of the language in which they are produced. Such a scoring method has resulted in higher scores on naming tests in studies of younger bilingual adults (Kohnert, Hernandez, & Bates, 1998) as well as of older bilingual speakers (Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007), than if the tests were scored for each language separately.

The few studies that have been conducted on naming in bilingual speakers with dementia so far have not yielded conclusive results. Several have found an earlier decline of naming performance in the L2 than in the L1 (e.g. Machado, Rodrigues, Simões, Santana, & Soares-Fernandes, 2010; Mendez, Saghafi, & Clark, 2004). However, there are also studies reporting comparable impairments across the languages (e.g. Filley et al., 2006; Hernández, Costa, Sebastián-Gallés, Juncadella, & Ramón, 2007; Veenstra, Huisman, & Miller, 2014) or greater impairment in the dominant language (not necessarily the L1 of the speakers) (Gollan, Salmon, Montoya, & Da Pena, 2010).

Among the many factors that may influence naming performance, are grammatical class, frequency, cognate status, and communicative task. The effect of grammatical class on the speed and accuracy of language processing has been studied extensively, not least across child and adult populations with language impairments. Noun-verb dissociations are evident in a range of types of disorders (Kambanaros & Grohmann, 2015). Dissociations in both directions are found for many of the populations including speakers with dementia (cf. Druks et al. (2006) for an overview). Frequency effects in naming in adults with language disorders are disputed, but when found, they occur more in nouns than in verbs (Bastiaanse, Wieling, & Wolthuis, 2016; Bird, Ralph, Patterson, & Hodges, 2000). In studies of bilinguals with and without language impairment, cognates (words that are similar in form and meaning in the two languages) are often found to be easier to retrieve than non-cognates (Costa, Santesteban, & Caño, 2005; Kohnert, 2004).

The majority of the studies on word retrieval in healthy aging are based on single word production tasks, which have certain limitations. On the one hand, single word production does not require the speaker to retrieve and produce more than one lexical unit; no integration of the unit in a larger syntactic or discursive frame is required. On the other hand, the speaker is usually required to retrieve one very specific item without the support of an interlocutor. A recent review by Kavé and Goral (2016a) shows that word retrieval difficulties found on single word production tasks do not generally extend to connected speech in healthy aging subjects. In speakers with aphasia, on the other hand, word finding difficulties are clearly present in connected speech. Also in speakers with dementia, connected language production (both orally and in writing) is affected compared to normal controls (e.g. Ahmed, Haigh, De Jager, & Garrard, 2013; Kavé & Goral, 2016b, 2017; Pekkala et al., 2013; Wilson et al., 2010), evident for instance in fewer words overall and fewer content words in particular, and increased proportions of closed-class words, pronouns and verbs.[2] Most of the studies reported on here have investigated semi-spontaneous forms of connected language production in the form of picture description tasks. Such tasks are closer to ordinary language use than single word production tasks, and the speaker is somewhat freer to choose which lexical item to produce. However, the picture to be described also restricts the speaker's choices, and usually, the speaker cannot rely on the interlocutor to scaffold the production.

The most common context of language use is conversation. Spontaneous talk in conversation sets high demands on language processing by requiring extensive planning and production of sentences and longer coherent stretches of talk, such as narratives. On the other hand, word finding is facilitated by the fact that the lexical items occur in a natural context. Furthermore, the speaker can appeal to the interlocutor for help in word search sequences.

Aims

In this study, to address the abovementioned gap in previous research on lexical retrieval, we aim to investigate the naming skills and the strategies used to cope with naming problems in a bilingual (US-Norwegian) male with PPA across three different speech contexts (confrontation naming, semi-structured narrative, and conversation) which put different demands on the speaker in the language production process. In particular, we focus on naming of nouns and verbs across these contexts. We investigate the progression of his naming problems over time. We also aim to discuss the results in the light of the two different theories described above, TDH and IDT.

Methods

We used a test battery consisting of a naming test and various cognitive tests as well as a background questionnaire on communicative functions. Moreover, we elicited oral narratives and collected conversational data. The cognitive tests and the questionnaire function as background information.

Data were collected at two time points: at approximately 12 months post diagnosis (T1) and at approximately 30 months post diagnosis (T2). At both time points, data were collected in two sessions, once in Norwegian and once in English, with one week apart. For each session, we aimed for a monolingual environment, in the sense that the main test administrator used only one language in each session, and in the English session a researcher with no knowledge of Norwegian was present to assist with data

[2]None of these studies include bilingual participants.

collection.[3] Ethical approval was obtained from the Norwegian Centre for Research Data (NSD).

The participant

The data were collected from a male bilingual speaker of US-English and Norwegian, here referred to as JJ. He was referred from his GP to a neurological examination at the age of 67 because of language problems, and after a brain scan (SPECT) 'showing reduced perfusion in large areas of the brain on both sides', he received a diagnosis of 'a variant of Alzheimer's disease, frontal lobe dementia, Primary Progressive Aphasia'. This is the wording (translated from Norwegian) from his medical journal, and when we first met JJ (at T1), he and his wife communicated to us that he had Alzheimer's disease. Only later, at the second data collection point when we got access to his medical records, did we realise that his diagnosis was actually Primary Progressive Aphasia. It is unclear to us why the diagnosis was formulated like this. Apart from the brain scan results we do not have knowledge about any data on the underlying pathology. Possibly, the neurologist himself did not distinguish clearly between these diseases, or he did not think that he had evidence to choose one over the other. There was no attempt in the medical journal to classify JJ's condition into any subtype of PPA. However, it is clear that the sole initial symptom was a language deficit, in accordance with the basic clinical diagnostic criteria of PPA (Gorno-Tempini et al., 2011: 1008), and the description of the brain scan results indicates a diffuse rather than a focal brain damage, ruling out a diagnosis of aphasia.

From biographical interviews with JJ and his wife, we learned that JJ has a higher academic education from USA, where he met his Swedish wife. They moved to Norway when he was 31 and got a job in a Norwegian subsidiary of an international company. At work, he used both Norwegian and English; he had some formal training in Norwegian, but learned most of it through immersion. At home, he spoke English with his wife and children, but there was much Norwegian and code mixing in the family. He spoke Norwegian with friends and neighbours, read Norwegian newspapers and watched Norwegian news. His wife rated him as very proficient and fluent in Norwegian, but with an US accent, prior to the onset of dementia. JJ reported that since the onset of dementia he had not been using Norwegian a lot. However, at the times of data collection, JJ was living at home with his wife and regularly attended Norwegian conversation groups at a day care center for elderly people with dementia.

Background information on cognitive functioning

JJ was tested with a battery of cognitive tests consisting of *The Rowland Universal Dementia Assessment Scale* (RUDAS) (Storey, Rowland, Basic, Conforti, & Dickson, 2004) and a

[3]For practical purposes, we had to use the same test administrator for both sessions, and this is of course unfortunate with regard to creating a monolingual condition. However, it is well-known that a great majority of the adult population in Norway, particularly those with a higher, academic education, understands and speaks English reasonably well, so completely monolingual settings are generally less likely when the languages in question are Norwegian and English.

Flanker task (Eriksen & Eriksen, 1974). These cognitive tests were carried out twice, in his L1, at T1 and at T2, 12 and 30 months post diagnosis, respectively.

RUDAS is a screening test which can be administered in less than ten minutes. It is very similar to the *Mini-Mental State Examination* (MMSE) (Folstein, Folstein, & McHugh, 1975), but RUDAS is specifically designed to assess cognitive impairment in culturally and linguistically diverse populations. It covers the following domains: memory, spatial orientation, praxis, visuo-constructional drawing, judgment, and language. The maximum score is 30, and the cut-off score, indicating possible cognitive impairment, is <23 (Storey et al., 2004).[4]

The Flanker task (Eriksen & Eriksen, 1974) measures both attention and inhibitory control. In the version we used, the subject sees five arrows on the screen either pointing in the same direction (congruent condition) or with all but the middle one pointing in the same direction (incongruent condition). The participant is asked to focus on the middle arrow, ignoring the arrows flanking it, and indicate if the middle arrow is pointing left or right. JJ was tested on 40 sequences in randomised order, half in the congruent condition and half in the incongruent condition. The task was presented electronically on a laptop computer using the E-Prime 2.0 software.[5]

As part of his medical examinations, JJ was also tested with MMSE four times: 1) at the time of diagnosis, 2) 10 months, 3) 22 months, and 4) 32 months post diagnosis. The first two times he was also tested with *Trail Making Test A* (TMT-A) and *B* (TMT-B), a test for assessing focused and joint attention and psycho-motoric speed (Reitan & Wolfson, 1985). Finally, at the first two and the last medical examination sessions, he was also administered the *Clock Drawing Test* (Strobel, Johansen, Wetterberg, & Engedal, 2012) assessing visuospatial abilities and semantic memory.

Table 1 gives an overview of the results of the cognitive testing at different points in time.

Table 1. Results of cognitive testing.

Time of testing	MMSE	RUDAS	Clock drawing	TMT-A (seconds)	TMT-A (z-score)[6]	TMT-B (seconds)	TMT-B (z-score)	Flanker
At diagnosis	30/30		5/5	48	1.17	301[7]	8.05	
10 mths p.d.	28/30		5/5	69	2.91	197	2.15	
12 mths p.d. (=T1)		22/30						23/40
22 mths p.d.	17/30		1/5					
30 mths p.d. (=T2)		16/30						26/40
32 mths p.d.	13/30							

[4]The full test is available from https://www.health.qld.gov.au/tpch/html/rudas. In a systematic review and meta-analysis with data from 1236 participants, Naqvi, Haider, Tomlinson, and Alibhai (2015) found a pooled sensitivity of 77.2% and a pooled specificity of 85.9% for RUDAS. A pooled estimate of the correlation between RUDAS and MMSE was 0.77, and scores on RUDAS were less affected by language and education level than scores on MMSE. They concluded that RUDAS is an assessment tool that 'has shown strong psychometric properties in several countries [and] shows particular advantage in culturally and linguistically diverse populations' (Naqvi et al., 2015: E175). Our reason for using RUDAS was that this case study is a pilot for a larger study of bilinguals with dementia, including participants from different countries and with low levels of education.

[5]Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

[6]Z-scores are calculated based on estimates of averages published in the MOANS age-corrected scaled scores (Ivnik et al., 1996).

[7]JJ was unable to complete TMT-B at time of diagnosis, and following scoring criteria his time was recorded as 301 seconds.

The MMSE scores do not give any indication of cognitive decline at the time of diagnosis, nor 10 months post diagnosis. However, the RUDAS results show that at 12 months post diagnosis the score is below cut-off. The MMSE and the RUDAS seem to correlate well, and indicate a steady decline over the next 20 months. The Clock Drawing Test indicates a similar trajectory, with a full score at the first the time of diagnosis as well as 10 months later, but a steep decline at 22 months post diagnosis. The TMT-A shows a performance below the 15th percentile (according to MOANS age-corrected scale scores (Ivnik, Malec, Smith, Tangalos, & Petersen, 1996)) at the time of diagnosis, and below the 5th percentile 10 months later; on the TMT-B he scores below the 10th percentile at the second test point. At the time of diagnosis, he was unable to complete the TMT-B, and was interrupted after making several errors. His time was then recorded as 301 seconds. We calculated z-scores for TMT-A and -B based on mean time ranges reported in Ivnik et al. (1996). These scores reflect a good estimate of JJ's performance relative to the normal population. As the norms in Ivnik et al. (1996) are reported as ranges rather than exact numbers, the z-scores in Table 1 are influenced by this. The low scores on the TMT tests indicate deterioration of attention and psychomotoric speed already at the time of diagnosis. On the Flanker task, even if the results at T1 (12 months post diagnosis) indicate reduced inhibitory control, this has not deteriorated to T2 (30 months post diagnosis). That JJ made a correct judgment to just over half of the trials at both T1 and T2, could be just by chance. A more thorough analysis of his responses shows that he answered all the congruent sequences and three of the incongruent sequences correctly at T1, and 19 of the congruent sequences and seven of the incongruent sequences correctly at T2, which indicates that he has understood the task, but has reduced inhibitory control.

Background information on functional communication

To get an impression of JJ's communicative abilities before and after the illness, his wife was asked to fill in the Norwegian version of the *Communicative Effectiveness Index* (CETI) (Lomas, Pickard, Bester, Erlbard, Finlayson, & Zoghaib, 2006).[8] This questionnaire was developed to measure the functional communication skills of a person with aphasia, but it has also been used with speakers with dementia (Burgeois & Hickey, 2009). It consists of 16 descriptions of situations, referring to different aspects of functional communication. A significant other, for instance the spouse of the person with the speech or language impairment, rates to what extent the person with the impairment is able to do whatever the relevant description refers to, by marking on a ten cm long visual-analogue scale ranging from 'not able at all' (at zero cm) to 'as able as before the stroke', i.e. the onset of the disease (at ten cm). The rating for each situation is converted into a score by measuring where along the ten cm long scale the mark is made, and a total score is calculated by dividing the sum of the individual situation ratings by the total number of

[8]JJ's wife was not instructed to evaluate his communication skills in just one of his languages exclusively, as the CETI is not linked to any specific language, but rather to functional, verbal communication in general. It is possible that the scores would have been different if she had been asked to evaluate his functional communication in his L1 separately from his L2 and vice versa.

situations (maximum: 100). It is important to remember that with the CETI, the person is compared to him-/herself rather than to a given norm. The focus is not primarily on the absolute score, but on the change in score over time (Lomas et al., 1989). A change in the total score of at least 12 points is regarded as clinically important (Lomas et al., 1989).

When JJ's wife completed the CETI at T1 (12 months after the diagnosis), the total score was 79.2, and at T2 (30 months post diagnosis) it was 63.1. This decline by 16.1 points indicates a clinically important deterioration in his functional communication skills as the disease progresses. Since a maximum score of 100 would have meant that there had been no change since the onset of dementia, we may also interpret the score of 79.2 at T1 as indicating a clinically important change (a change by >12 points) in his functional communication skills already at the first measurement point. The items that received particularly low scores (55 or below) at T1, were items number 2 ('Getting involved in group conversations that are about him/her'), 4 ('Communicating his/her emotions'), and 14 ('Being part of a conversation when it is fast and there are a number of people involved'). Also, items number 6 ('Having coffee-time visits and conversations with friends and neighbours (around the bed or at home)'), 12 ('Starting a conversation with people who are not close family'), and 16 ('Describing or discussing something in depth') were rated lower than the rest of the items. At T2 the pattern is about the same, but in addition items number 3 ('Giving yes and no answers appropriately'), 13 ('Understanding writing') and 15 ('Participating in a conversation with strangers') are rated markedly lower than at T1. The results of the CETI thus indicate that cognitively and interactionally demanding speech contexts have become more problematic for JJ since the onset of dementia, and as the disease progresses comprehension also seems to have become more problematic.

Data on naming abilities and strategies – collection and analysis

Data on naming abilities were collected in three different settings: a) a confrontation naming task, b) a cartoon description task, and c) interview and small talk between the participant and the test administrator. The confrontation naming task consists of 60 items selected from the Norwegian versions of two test batteries developed primarily for the assessment of aphasia: The *Psycholinguistic Assessments of Language Processing in Aphasia* (PALPA) (Kay, Lesser, & Coltheart, 2009) and the *Verb and Sentence Test* (VAST) (Bastiaanse, Lind, Moen, & Simonsen, 2006). 30 drawings depicting objects were selected from subtest 53 of PALPA (oral picture naming), and 30 drawings depicting actions were selected from subtest 4 of VAST (naming of actions).[9] The aim of the task was to elicit single nouns and verbs. The same set of pictures was used for this task in English and Norwegian. The words were not matched for any underlying variables, but post hoc analyses show that 16/30 of the nouns and 15/30 of the verbs are cognates, for example Norwegian *katt, buss, (å) danse, (å) filme* vs. English *cat, bus, (to) dance, (to) film*. A Wilcoxon rank sum test show that there is no significant difference in the frequencies of

[9]Both of these tests originally comprise 40 items, and norms based on healthy, monolingual speakers for the full versions of the tests show high average scores (98% on both tests).

the words across the languages ($W = 2122.5$, $p = 0.091$). We should note, however, that the selection of words is slightly skewed, with few very high frequency items. The responses in the confrontation naming task were coded as correct or incorrect in the target language. Incorrect responses were further analysed as code switches (CS) and/or as related to the target item (semantically, phonologically) or not. A chi-square test was used to investigate the difference in scores between the two languages at both time points. Frequency scores were logarithmically transformed and differences were checked with a Wilcoxon rank sum test for nouns and verbs in both languages at both time points. Finally, a Fisher's exact test was used to investigate if there was a cognate effect between word classes in both languages.

The cartoon description task (a sequence of six pictures) is taken from the *Bilingual Aphasia Test* (Paradis & Libben, 1987). The descriptions were audiotaped and subsequently transcribed orthographically, including all words and attempts at words (neologisms, false starts, and so on) and dysfluencies, such as repetitions. For the analysis, we counted the number of words in each text. Only words in the target language were included, as were repetitions, truncated words, and lexical paraphasias. Incomprehensible words (transcribed with an X or X=) and hesitations (*eh*) were excluded. We also calculated the lexical density by dividing the number of lexical word tokens (nouns, adjectives, and verbs (excluding auxiliary verbs and semantically light verbs such as *be, have, come, go, give, take, make, do, get, put* (Gordon, 2008))) by the total number of words.[10] Furthermore, we defined ten key content components of the cartoon's plot for the purpose of coding and analysis. As there is no previously established set of key components for performing a concept analysis of the cartoon in the BAT, pre-analysis, we decided on ten key content components (Table 2). We analysed the extent to which the ten components are accurately and completely presented (Nicholas & Brookshire, 1995), which in essence means that they should be presented in such a way that the gist of the story can be understood by a listener who does not necessarily have access to the pictures. For this, a certain number of content words has to be present; for instance, a bare pronoun (*he*) does not suffice as an accurate and complete presentation of the first key component.

A well-produced narrative does not necessarily have to include each and every one of the components, and they do not necessarily have to appear in the order listed. However, for the complete story to be conveyed, most of these components will have to be represented. Since the narrative is based on pictures that could be seen both by the

Table 2. Key conceptual components of the BAT cartoon.

1	Introduction of the man (main participant in the story)
2	Introduction of the woman (main participant in the story)
3	Introduction of the birds in the nest (main participants in the story)
4	The man climbs the tree
5	The branch breaks
6	The man falls down
7	The man's leg is broken
8	The ambulance arrives
9	The man is in hospital
10	The bird is crying over her dead chicks outside the hospital

[10]Auxiliaries and semantically light verbs were not included in the calculation of lexical density since they are best regarded as function words and not content words (Malvern, Richards, Chipere, & Durán, 2004).

participant and the test administrator, the explicitness of the narratives may vary; implicitly, they could co-construct the narratives. However, in this task it was made clear to JJ that he was supposed to be the narrator.

The conversational data were gathered from an autobiographical interview with JJ in his L1 (English) at T1 and from small talk sequences before and in between the test sessions in both English and Norwegian at both test times. The informal autobiographical interview dealt with his life history and especially his use of different languages in various situations. All the test sessions were audiotaped, and the conversational data were subsequently transcribed in a detailed manner, using conversation analytic conventions (Jefferson, 2004) and including such features as pauses, speech overlap, prosody, and intonation (for transcription key, see Appendix A). Instances of word search sequences were excerpted and analysed qualitatively, using a conversation analytic methodology (Sidnell, 2011). Word search sequences are defined in this tradition as sequences where the speaker a) interrupts a turn-in-progress, b) initiates a searching activity, indexed by pauses, vocal features (filled pauses, sound stretches), meta comments, and/or non-verbal behaviour[11] (gestures, averted eye gaze, 'thinking face' etc.), and c) finds a solution to the problem (finds the word, finds another way of saying the same thing, or abandons the utterance) (Goodwin & Goodwin, 1986). Word searches may also become interactive by the speaker appealing to the interlocutor for help (seeking eye contact, asking questions, producing inviting gestures), and the latter may provide suggestions for the speaker to accept or decline. During the conversations, no specific guidelines were given to the conversation partners on how to handle word-finding problems; the most important thing was to keep the conversation going as naturally as possible.

Results

The results of the three tasks are presented consecutively below. For all three tasks, it is interesting to note that JJ showed no phonological distortions or paraphasias in either language (apart from the fact that he had an US accent in his Norwegian, a feature noted by his wife from before the onset of the disease).

The confrontation naming tasks

Figures 1 and 2 summarise the results of the confrontation naming tasks. Figure 1 shows that at T1, JJ performs relatively well on the nouns in his L1 (English), but he has more problems with the verbs. It is clear, however, that in most cases he understands the depicted verbal concept, but that he is unable to retrieve the correct lexical item from his mental lexicon. Instead, he responds with a lexical verb semantically related to the target verb, usually a synonym, such as *sweeping the floor* for the target *vacuum* or a description of the target action. Interestingly, while there is a frequency effect on the L1 nouns – he performs better on high frequency nouns ($W = 12$, $p = 0.002$) – there is no such effect on his L1 verbs ($W = 66$, $p = 0.064$). He has no instances of code switching to Norwegian when naming objects and actions in his L1 at the first measurement point.

At T1, his scores in the L2 (Norwegian) (Figure 2) are significantly lower ($\chi^2 = 14.72$, df = 1, $p = 0.0001$). For both objects and actions, he manages to produce the correct noun

[11]Since we only had audiotaped data, non-verbal behaviour was obviously not taken into account.

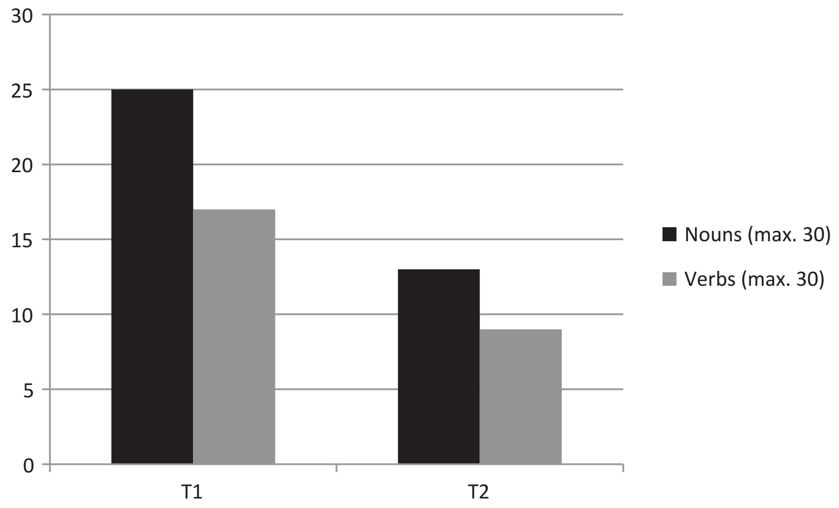


Figure 1. Results of the confrontation naming tasks in English (L1) at T1 and T2.

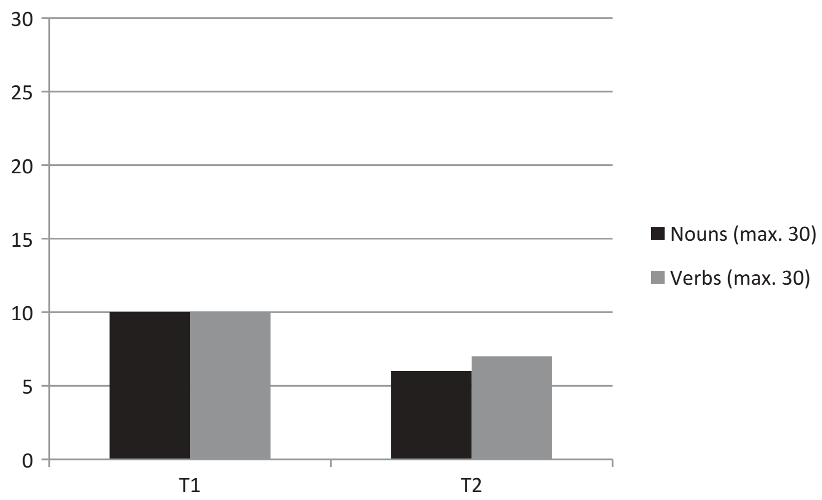


Figure 2. Results of the confrontation naming tasks in Norwegian (L2) at T1 and T2.

or verb in the target language in only one third of the test items. In the object naming task we see a clear cognate effect in the correct responses; 9 of the 10 correct responses are cognates, such as *bok/book, fisk/fish* and *buss/bus* (p = 0.006). In the action naming task, the cognate effect is not significant (p = 1), although half of the correct verb responses are cognates. Unsurprisingly, as English is his strongest language, there is no cognate effect in the L1 at T1. We may also note that among the correct responses (both for objects and actions) in his L2 there are almost no low frequency items. For both nouns and verbs, there is a significant frequency effect: nouns (W = 53, p = 0.039); verbs (W = 35, p = 0.003).

The error patterns in the object and action naming tasks in Norwegian at T1 differ somewhat. When failing to come up with the correct target verb in the action naming task, JJ's main strategy is to produce a semantically bleached, all-purpose verb such as *bruke* (use), *ta* (take) or *gå* (go), often in combination with a noun. Some examples are *bruke en saw* (CS) (use a saw) instead of *sage* ((to) saw), and *ta opp* (take up) instead of *plukke* (pick (flowers)). In the object naming task JJ has another strategy. When failing to come up with the correct target noun, he often code switches to English. For ten items, he produces the correct English equivalent of the target noun, e.g. *pig* for *gris, train* for *tog* and *dart* for *pil.*

As in the L1 action naming task, his production of correct English equivalents for the Norwegian target nouns indicates that he understands the depicted concept, but that he is unable to retrieve the correct lexical item in the target language from his mental lexicon. If we add the ten 'correct' English nouns to the ten correct Norwegian (target) nouns, using a composite scoring method, we see that JJ has a better grasp of nouns than of verbs – at least at a conceptual level – not only in his L1, but also in his L2 at T1. For Norwegian, his composite score for nouns is 20/30, while the composite score for verbs remains low at 12/30. His very low score and the dominant error pattern in the Norwegian action naming task (an overuse of highly frequent, semantically light verbs) may indicate more severe difficulties with verbs as concepts in his L2 (Norwegian) than in his L1 (English) at T1.

Over time, we see a clear deterioration on all scores in both languages. At T2, there is no longer any difference in the performance between L1 and L2 ($\chi 2$ = 2.58, df = 1, p = 0.108). His score on nouns in the L1 has fallen to 13/30 (Figure 1), 9 of which are cognates; however, there is no effect of either cognates (p = 0.07) or frequency (W = 66, p = 0.064). A closer look at his error pattern still reveals a better retained understanding of the depicted nominal concepts as 9 of the 17 errors are descriptions or synonyms. His L1 score on verbs is even lower (9/30), and only three cognates are correct. For half of the items he is unable to provide any response, and for the remaining six he gives a description of or a synonym for the target action. There is no effect of frequency (W = 57, p = 0.094). There are no instances of code switching to Norwegian on object and action naming in the L1 at T2.

In the L2 (Figure 2), the scores on nouns and verbs are equally low, and the cognate effect is only found for nouns (p = 0.01), where all the six correct answers are cognates; only 3/7 correct verbs are cognates, which is not statistically significant (p = 1). For the nouns, but not for the verbs, there is also a statistically significant frequency effect (W = 26, p = 0.015). In the object naming test, the composite score reaches 12/30; in addition, there are seven descriptions, indicating a still reasonably well-preserved grasp of the nouns in his L2 at a conceptual level. In the action naming test, a composite score only adds one item to the original 7/30 correct responses. For 18 verbs, he is unable to provide a response.

In sum, his results on the confrontation naming tasks indicate greater impairment of his L2 than his L1 at T1, a difference that decreases as his dementia progresses (T2); more severe difficulties with verbs than nouns; and a better retained understanding of concepts, in particular nominal ones, than words in both languages as indicated by his use of synonyms, descriptions, and code switching when he cannot access the target word.

Naming in semi-spontaneous elicited narratives

JJ's descriptions of the cartoon from the BAT are longer at T2 than at T1, both in English and in Norwegian, but the lexical density diminishes in both languages as time passes and the disease progresses (cf. Table 3). Given the low and decreasing lexical density in the

Table 3. Results of the narrative tasks.

Test time	Language	Number of words (target language)	Lexical density	Number of key component units
T1	English	61	0.30	3/10
	Norwegian	92	0.23	3/10
T2	English	204	0.20	3/10
	Norwegian	138	0.14	1/10

narratives, we may expect them to convey little information. An analysis of the number of key component units that are accurately and completely presented in the texts, indeed shows that very few of the ten pre-defined components are present in each of the narratives in a sufficiently accurate and complete way (cf. Table 3 and Appendix B).

Naming in interaction

Despite the fact that JJ has challenges of naming in the experimental task and of conveying substantial information in the narrative task, the interactional data show that he manages to communicate rather successfully in conversation, at least in his L1, English, and most clearly at the first data collection point. However, the difference in performance in his L1 and L2 as observed in the tests is also observable in his conversational contributions. We will here describe some of the strategies he uses when facing a word retrieval problem.

In the English conversation at T1, the main naming problem seems to be retrieving low frequency nouns. A very common strategy used by JJ is to replace a word he does not find with an explanation of the meaning of that word. This is what we find in excerpt (1),[12] where he is talking about changing jobs, and seems to be searching for the word *application*:

Excerpt (1)

```
1  JJ   [And] then (0.5) and then at that time I sent eh: (1.0) (tsk) eh:
2       (1.5) sent out er (0.5) ((thump)) er you know (.) t- trying to
3       get a job (0.5) eh: (1.0) in (.) in places in Sweden.
4  HGS  Mhm.
```

The utterance initiated in line 1 projects a complement in the form of a noun phrase. However, instead of a noun phrase there is a series of filled and silent pauses, displaying the activity of searching for a word. There is also a recycling of the finite verb (*sent* – *sent out*) and a potential appeal to the interlocutor to search for the intended referent (*you know*). When the search is still unsuccessful and the interlocutor does not provide any sign of understanding, he produces an explanation of the meaning of the word instead of the word itself. After this he continues the turn, and the interlocutor provides an acknowledgement token, claiming understanding of the turn so far. This strategy of circumlocution (Tarone, 1980) or paraphrase (Kurhila, 2006) is well documented in second language speakers lacking a word in the target language. Here we can relate it to JJs problems with retrieving low frequency words, and more specifically, with phonological access, since he is able to provide a semantically equivalent description.

[12]In the excerpts, HGS refers to the test administrator/main interlocutor. A research assistant (AS) was also present, but she contributed mainly with minimal responses.

Sometimes the interlocutor may help in the search for a word by providing suggestions. In excerpt (2), we have an instance of this.

Excerpt (2)

1 JJ And (1.5) and then from (.) from there I went to eh Johnston
2 University, (0.5) [in] eh Wyoming,
3 HGS [mm]
4 AS Yeah,
5 JJ and that's (.) where I (0.5) eh: got my (2.0) s- whatever I got
6 [he he he]
7 HGS [The degree?]
8 AS [mm] Mhm
9 JJ A degree yeah,
10 HGS Yeah, [what] eh what eh: in what field?

A word search occurs in line 5, where JJ suspends an utterance after a possessive pronoun (*my*), projecting an NP head, and leaves a long silence of two seconds. This time he seems to abandon the search and give up finding the referring expression (Tarone, 1980). The laughter accompanying this abandonment may seem oriented to compensating for the embarrassment created by this communicative failure (Lindholm, 2008; Wilkinson, 2007). However, at this point the interviewer comes in with a candidate solution to the retrieval problem by suggesting *the degree* with rising intonation. This suggestion is accepted by JJ in that he produces an echo answer – a repeat with an affirmative response token (Svennevig, 2003).

In the Norwegian conversation at T1, the problems seem to be of a more fundamental character and occur at points when it is not just a single noun that is projectable, but a main verb, and thus the complete sentence structure. An example can be seen in the next excerpt. JJ has excused himself for the low test scores, and the interlocutor (HGS) responds by saying that he has been using English much more during his life time and in addition that he was 'born with it'. After a minimal response, JJ continues:

Excerpt (3)

1 JJ Jeg har altså vært ((KREMT)) (1.3) at (1.2) e:h (3.0) når jeg s-
I have (PART) been ((COUGH)) (1.3) that (1.2) e:h (3.0) when I s-
2 (sa) det var (.) problem,
(said) it was (.) problem,
3 AS [Mhm]
4 HGS [Mhm] =
5 JJ =e:h (0.7) eller (sånn) (0.5) eh gikk (.) som f- f- funnet hva det
=e:h (0.7) or (like) (0.5) eh went (.) that f- f- found what it
6 JJ var [og så] videre .hh eh ((KREMT)) (1.0) (tsk) (1.0) o:g (2.0)
was and so on .hh eh ((COUGH)) (1.0) (tsk) (1.0) a:nd (2.0)
7 HGS [ja .hhja]
Yeah yeah
8 HGS [mm]
9 JJ det er ehm (0.7) (det er eh) (0.9) (hårdt) nok
It's ehm (0.7) (it's eh) (0.9) (hard) enough
10 HGS Mhm=
11 JJ =til å eh (.) bruke: en.
=to eh (.) use one.

12 AS Mm=
13 JJ =heheheh
14 HGS hehhehhhh ja (nemlig) [ja]
hehhehhhh yeah (exactly) yeah

In this extract, there are several instances of dysfluencies and pauses at points where there is no specific noun projected. In line 1 the search activity starts after a semantically light verb, *å være* (to be), which does not provide many clues about what sort of item he is searching for. The search continues after a subordinate clause in lines 1–2 (‘when I said it was problem’), before the main verb of the main clause is presented. The same seems to be the case in the next round of searching in line 5. Here he produces what may be a self-repair of the previous subordinate clause, as indicated by the self-repair marker *eller* (or). Finally, in line 9 he pauses and restarts after a semantically empty introductory construction, *det er* (it is).

What is common to these instances of apparent production problems in his L2 is that they occur before a clear structure may be assigned to the utterance as a whole. Thus, they are not just related to retrieving a specific lexical item, but rather to constructing a whole propositional utterance. A consequence of this is that in the Norwegian interaction the interlocutors do not have any basis for guessing what JJ is searching for and thus for assisting him by providing suggestions. As can be seen, they do not provide any candidate solutions in the long pauses that occur while JJ is searching.[13]

The picture that emerges from analysing the spontaneous conversations at T1 is that in the English data, the problem concerns retrieval of low frequency nouns. JJ and his interlocutors manage to overcome JJ’s lexical limitations to a large extent. JJ himself uses semantic strategies such as circumlocution and approximation (for instance using a hyponym) as alternative paths to the identification of the referent. The interlocutors participate actively by providing candidate suggestions in cases where JJ does not find a solution himself. In the Norwegian data, by contrast, the problem seems to concern the planning of the sentence structure as a whole. The problems often occur at an early point before the main verb has been produced. This severely limits the opportunities for the interlocutors to scaffold the speaker in his speech production.

At T2, JJ’s conversational performance had deteriorated in both his L1 and his L2. While testing him in Norwegian, the interviewers tried to initiate some small talk in between the tests. However, JJ either answered in English or had such large problems in trying to formulate an answer in Norwegian that the interviewers rather quickly abandoned the attempt. In his English conversation, however, JJ managed to keep a conversation going, but had more word finding problems, and more importantly, more difficulties in using compensatory strategies to remedy the problems. The next excerpt is an example of this. JJ is telling about his daughter, but encounters a word finding problem when he tries to tell what her profession is:

Excerpt (4)

1 JJ: [Yeah] and eh: (1.0) so she’s been (1.2) been working with
2 ehm (.) ts °e:::h° ne:w (1.1) un- (.) (when) (.) people get the

[13]For a more thorough analysis of this extract, see Svennevig & Lind (2016).

```
3          (.)°eh (.) e:hm e:h e:h with the (.) if they get a ca-
4          eh you know° hh hh [hh]
5  HGS:                       [he] he
6  JJ:     a- how how how you go get (.) e:h you know the the first first
7          when they first they get- and especially .hh a::h if they're
8          (.) if they're from outside of of Norway.
9  HGS:    Oh i:mmigrants?
10 JJ:     [Yeah it's] it's it is [is sort] of [yeah] so:me of them are.
11 AS:     [Mhm yeah]             [yeah]
12 HGS:                                        [wha-]
13 HGS:    [Yeah] oka:y
14 AS:     [Mhm]
15 JJ:     .hh (.) but but e:h (.) but to any of the: (.) the: (0.8)
16         young people when they=
17 AS:     =Mhm
18 JJ:     the the f- first get eh: eh: (.) a chi:ld.
19 HGS:    Ok↑a::y.
```

JJ here initiates a search for a word to specify what his daughter had *been working with* (line 1). After some time searching for the word without success, he initiates a new clause (*when people get the*) that is not syntactically fitted to the point of suspension of his previous clause. Rather, it seems to be an attempt to explain the phenomenon in a paraphrase or a circumlocution. However, also this attempt fails, since he stops mid-course and starts to search for a new word to complete this clause (line 3). He tries to appeal to his interlocutors for help with the discourse marker *you know* and displays his embarrassment of not finding words by starting to laugh (line 4). HGS joins in with the laughter, but does not contribute any suggestion, being without any contextual clues to guess what he might be searching for.

In lines 6–8, the search continues and some contextual clues are added. His specification of the numeral *first* neither helps him or his interlocutors, but the second specification of a typical condition ('being from abroad') leads to a candidate suggestion from HGS (line 9). The suggestion is only partially accepted by JJ, who instead gives another contextual clue by adding that the people are *young* (line 16). Finally, in line 18, he retrieves the word he had been searching for since line 2, the object of the verb *get*, namely *a child*. HGS responds by an emphatic claim of understanding.

Although this word search is very extended and complicated, it only solves a part of the problem, namely the formulation of an explanation about the daughter's profession. What this profession is, still remains to be established. The excerpt shows the deterioration in JJ's skills in compensating for his word finding problems by making use of alternative means such as explanations of meaning and circumlocution. It also shows how his linguistic deterioration limits the opportunities for the interlocutors to assist in solving his word finding problems, much like in his Norwegian conversations at T1.

Discussion

In this article, we have presented analyses of the naming skills of a bilingual speaker with PPA in each of his languages across three different speech contexts: confrontation naming, semi-spontaneous picture description, and conversation, at two points in time.

Our data show firstly that the dementia has clearly affected his language skills and his conversational performance, and thus corroborate the report by his wife in the CETI questionnaire that his abilities to communicate functionally have deteriorated since the onset of dementia and progressively so. Consistent with his premorbid use of and proficiency in the two languages, his performance in his L2 Norwegian is lower than in his L1 English, but this difference diminishes as the disease progresses. This is the case across the three speech contexts; however, the difference is smaller in the narrative task, where his performance is very low in both languages already at T1.

Secondly, in our data in the confrontation naming task, we find effects of word class, frequency, and cognateness. The test scores show a better performance on nouns than on verbs in the L1 only, at both T1 and T2. However, also in L2, the error patterns reveal that the two word classes are treated differently. If we use a composite score including correct L1 answers in the L2, he has a lower performance on verbs than on nouns also in his L2, at both time points. High frequency words are clearly easier to retrieve at T1 in both languages (except for verbs in L1), but this effect has nearly disappeared at T2 (except for nouns in L2). In L2, his less dominant language, there is a clear cognate effect for nouns, but not for verbs, at both T1 and T2.

In conversation, JJ performs better than he does in the semi-spontaneous narrative. At T1, especially in his L1, English, he communicates quite well, and manages to convey a rich and detailed life story. His retrieval problems mainly concern specific low frequency nouns, and they are generally remedied by compensatory strategies such as circumlocution or approximation. In his Norwegian (L2) conversation, by contrast, the naming problems are more frequent and severe, and concern more fundamental aspects of sentence construction. Consequently, the opportunities for scaffolding are much more limited. At T2, the compensatory strategies for word retrieval become more limited even in English, resulting in less opportunities for the conversation partner to bring the conversation forward.

Despite some variation across the tasks, the overall impression is that JJ performs best in his first language, in accordance with his premorbid proficiency and use. However, across both languages nouns are easier to retrieve than verbs, and in semi-spontaneous narratives his performance is very low in both L1 and L2. Thus, although it may look as if JJ's L1 is less affected by the dementia than his L2, in line with prior research by e.g. Machado et al. (2010) and Mendez et al. (2004), both his narrative performance and the dissociation between nouns and verbs indicate that there are parallel impairments across his two languages. As mentioned, in the literature the findings so far are mixed regarding the question of parallel or non-parallel impairment of the languages of bilingual speakers with dementia. Further research is obviously needed in this field.

In confrontation naming, his error strategies when word retrieval fails – descriptions, use of synonyms and code switching – all indicate that his concepts are better retained than his words. This suggests that his impairment is not related to semantic memory, but rather is a problem of phonological processing, of accessing the phonological form of the word. His extensive use of code switching when naming nouns in his L2 might be seen as more of a compensatory strategy than as lack of inhibition. As mentioned above, JJ could safely assume that the test administrator knew both his languages, so code switching was a feasible strategy. Thus, his performance seems to support the Transmission Deficit Hypothesis (Burke et al., 1991). However, his poorer performance on verbs than on nouns across all contexts – and worsening with the progression of the disease – indicates a deficit in semantic and syntactic processing in particular related to action naming. This finding supports the research pointing to a dissociation between nouns and verbs found in different types of disorders (Druks et al., 2006; Kambanaros & Grohmann, 2015).

The diagnosis based on JJ's medical examination and cognitive tests was Primary Progressive Aphasia, but also described as a variant of Alzheimer's disease. It seems clear to us that since his disease so clearly started with a language deficit with cognitive functioning intact, followed by a gradual, progressive impairment of both cognitive and linguistic functioning, this is a case of PPA (Gorno-Tempini et al., 2011: 1008). The discussion in the literature concerning the sub-classification of PPA (e.g. Gorno-Tempini et al., 2011; Wilson et al., 2010), shows that it is not always easy to distinguish clearly between the different subtypes in each case. Although we are not in a position to classify JJ's PPA definitely, our linguistic investigation can bring us closer to an answer. His performance across all speech contexts shows the following pattern: His speech is generally fluent with an acceptable speech rate, not agrammatic, and he has no phonological distortions or paraphasias; this should exclude the non-fluent variant of PPA. On the other hand, his word finding problems result in many hesitations, pauses and repairs in connected speech. He has clear problems with word retrieval in confrontation naming as well as in connected speech, but spared object knowledge; this seems to exclude the semantic variant of PPA, and rather indicates that he suffers from a logopenic subtype of PPA.[14] The fact that in a large proportion of cases of this subtype of PPA, the underlying pathology is similar to that of AD (Gorno-Tempini et al., 2011), may support the 'double' diagnosis given by his neurologist.

In this study, we used several tools to explore the naming skills and strategies of JJ, allowing us to examine the phenomenon from different perspectives, in our case, contexts with different demands on language processing. By combining different tools and methods of analysis we get a 'thicker' description and thus a more comprehensive picture of the impact of the dementia on JJ's languages from an intra-individual as well as an inter-individual perspective. For instance, the greater difficulties in the Norwegian interaction at T1, which seem to be linked to fundamental aspects of sentence construction, are more readily understandable in view of JJ's particularly low scores on Norwegian verbs in the confrontation naming task. On a more general level, this corroborates the crucial role of verbs in sentence construction and communication (e.g. De Diego Balaguer et al., 2006). Finally, our data show that it is necessary to assess a bilingual person in both his languages, to get a comprehensive view of strengths and weaknesses in language functioning.

A combination of different tools and methods is also useful in clinical practice as it allows both a systematic exploration of possible effects of confounding factors such as grammatical class, cognate status and frequency, and an exploration of the phenomenon in speech contexts with different demands for language processing. By examining talk-in-interaction one also gets the opportunity to study the role of the interlocutor in managing the linguistic and communicative challenges of the person with dementia. The knowledge gained from studies of conversations, preferably combined with studies of language

[14]Unfortunately, sentence repetition was not tested, so this important diagnostic feature remains unattested.

processing in more restricted contexts, provides a good basis for guiding significant others and professional caregivers. These interlocutors may need advice on how best to interact with the person with dementia in such a way that the person is able to make use of the linguistic and communicative resources she or he still has despite the impairments.

Limitations

Our study has several limitations. First, the medical diagnosis given at the outset was not entirely clear, and since we do not have access to the actual brain scan results, we cannot confidently evaluate the neurological basis for the disease. Second, this is a case study, and we have no control group to compare with. Although control groups for bilinguals are always problematic since it is near impossible to find groups of individuals with comparable language histories, a proper baseline control against which we could have evaluated JJ's performance, would have been preferable.

Acknowledgements

We thank JJ and his wife for participating in this study, Malene Bøyum and Gözde Mercan for assistance with the data collection, and Pernille Hansen for help with the statistical analyses.

Declaration of interest

The authors report no conflicts of interest.

Funding

This work was supported by The Research Council of Norway [223265 and 250093].

ORCID

Kees de Bot http://orcid.org/0000-0001-7442-2693

References

Ahmed, S., Haigh, A.-M. F., De Jager, C. A., & Garrard, P. (2013). Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain*, *136*, 3727–3737. http://doi.org/10.1093/brain/awt269

Balthazar, M. L. F., Cendes, F., & Pereira Damasceno, B. (2008). Semantic error patterns on the Boston Naming Test in normal aging, amnestic mild cognitive impairment, and mild Alzheimer's disease: Is there semantic disruption?. *Neuropsychology*, *22*, 703–709. doi:10.1037/a0012919

Bastiaanse, R., Lind, M., Moen, I., & Simonsen, H. G. (2006). *Verb- og setningstesten* (VOST). Oslo, Norway: Novus.

Bastiaanse, R., Wieling, M., & Wolthuis, N. (2016). The role of frequency in the retrieval of nouns and verbs in aphasia. *Aphasiology*, *30*(11), 1221–1239. doi:10.1080/02687038.2015.1100709

Bayles, K. A., & Tomoeda, C. K. (1983). Confrontation naming impairment in dementia. *Brain and Language*, *19*, 98–114. doi:10.1016/0093-934X(83)90057-3

Bird, H., Ralph, M. A. L., Patterson, K., & Hodges, J. R. (2000). The rise and fall of frequency and imageability: Noun and verb production in semantic dementia. *Brain and language*, *73*(1), 17–49. doi:10.1006/brln.2000.2293

Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S. F. (2017). Connected speech in neurodegenerative language disorders: a review. *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.00269

Burgeois, M. S., & Hickey, E. M. (2009). *Dementia. From diagnosis to management – a functional approach*. New York, NY: Psychology Press.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults?. *Journal of memory and language*, *30*(5), 542–579. doi:10.1016/0749-596X(91)90026-G

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science*, *13*, 21–24. doi:10.1111/j.0963-7214.2004.01301006.x

Burke, D. M., & Shafto, M. A. (2008). Language and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 373–443). New York, NY: Psychology Press.

Chen, P., Ratcliff, G., Belle, S. H., Cauley, J. A., DeKosky, S. T., & Ganguli, M. (2001). Patterns of cognitive decline in presymptomatic Alzheimer disease: A prospective community study. *Archives of General Psychiatry*, *58*, 853–858. doi:10.1001/archpsyc.58.9.853

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and language*, *94*(1), 94–103. doi:10.1016/j.bandl.2004.12.002

De Diego Balaguer, R., Rodríguez-Fornells, A., Rotte, M., Bahlmann, J., Heinze, H. J., & Münte, T. F. (2006). Neural circuits subserving the retrieval of stems and grammatical features in regular and irregular verbs. *Human Brain Mapping*, *27*, 874–888. doi:10.1002/hbm.20228

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, *93*(3), 283. doi:10.1037/0033-295X.93.3.283

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*(1), 287–314. doi:10.1016/0010-0277(92)90046-K

Druks, J., Masterson, J., Kopelman, M., Clare, L., Rose, A., & Rai, G. (2006). Is action naming better preserved (than object naming) in Alzheimer's disease and why should we ask?. *Brain and Language*, *98*, 332–340. doi:10.1016/j.bandl.2006.06.003

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. doi:10.3758/BF03203267

Filley, C. M., Ramsberger, G., Menn, L., Wu, J., Reid, B. Y., & Reid, A. L. (2006). Primary progressive aphasia in a bilingual woman. *Neurocase*, *12*(5), 296–299. doi:10.1080/13554790601126047

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

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21*, 354–375. doi:10.1093/applin/21.3.354

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *The Journal of the International Neuropsychological Society*, *13*, 197–208.

Gollan, T. H., Salmon, D. P., Montoya, R. I., & Da Pena, E. (2010). Accessibility of the nondominant language in picture naming: A counterintuitive effect of dementia on bilingual language production. *Neuropsychologia*, *48*, 1356–1366. doi:10.1016/j.neuropsychologia.2009.12.038

Goodwin, M. H., & Goodwin, C. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, *62*, 51–75.

Goral, M. (2013). Bilingualism, language, and ageing. In J. Altarriba & L. Isurin (Eds.), *Memory, language, and bilingualism. Theoretical and applied approaches* (pp. 188–210). Cambridge, NY: Cambridge University Press.

Gordon, J. K. (2008). Measuring the lexical semantics of picture description in aphasia. *Aphasiology*, *22*(7–8), 839–852. doi:10.1080/02687030701820063
Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. E. E. A., … Manes, F. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. doi:10.1212/WNL.0b013e31821103e6
Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingualism: Language and cognition*, *1*, 131–149. doi:10.1017/S136672899800025X
Grossman, M., & Ash, S. (2004). Primary progressive aphasia: a review. *Neurocase*, *10*(1), 3–18. doi:10.1080/13554790490960440
Hernández, M., Costa, A., Sebastián-Gallés, N., Juncadella, M., & Ramón, R. (2007). The organisation of nouns and verbs in bilingual speakers: A case of bilingual grammatical category-specific deficit. *Journal of neurolinguistics*, *20*, 285–305. doi:10.1016/j.jneuroling.2006.10.002
Hilger, A., Ramsberger, G., Gilley, P., Menn, L., & Kong, A. (2014). Analysing speech problems in a longitudinal case study of logopenic variant PPA. *Aphasiology*, *28*(7), 840–861. doi:10.1080/02687038.2014.895974
Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, STROOP, TMT, and JLO. *The Clinical Neuropsychologist*, *10*(3), 262–278. doi:10.1080/13854049608406689
Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the first generation* (pp. 13–31). Amsterdam, The Netherlands: John Benjamins.
Kambanaros, M., & Grohmann, K. K. (2015). Grammatical class effects across impaired child and adult populations. *Frontiers in Psychology: Language Sciences*, *6*, 1670. doi:10.3389/fpsyg.2015.01670
Kavé, G., & Goral, M. (2016a). Do age-related word retrieval difficulties appear (or disappear) in connected speech?. *Aging, Neuropsychology, and Cognition*. doi:10.1080/13825585.2016.1226249
Kavé, G., & Goral, M. (2016b). Word retrieval in picture descriptions produced by individuals with Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *38*, 958–966. doi:10.1080/13803395.2016.1179266
Kavé, G., & Goral, M. (2017). Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses. *Aphasiology*, 1–23. doi:10.1080/02687038.2017.1338663
Kay, J., Lesser, R., & Coltheart, M. (2009). *Psykolingvistisk kartlegging av språkprosessering hos afasirammede* (PALPA) (Bredtvet kompetansesenter, Logopedtjenesten – Helse Bergen, Statped Vest & Øverby kompetansesenter, adaptation). Oslo, Norway: Novus. (Original published 1992).
Kempler, D., & Goral, M. (2008). Language and dementia: Neuropsychological aspects. *Annual review of applied linguistics*, *28*, 73–90. doi:10.1017/S0267190508080045
Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and language*, *91*(3), 294–302. doi:10.1016/j.bandl.2004.04.001
Kohnert, K. J., Hernandez, A. J., & Bates, E. (1998). Bilingual performance on the Boston Naming Test: preliminary norms in Spanish and English. *Brain and Language*, *65*, 422–440. doi:10.1006/brln.1998.2001
Kurhila, S. (2006). *Second language interaction*. Amsterdam, The Netherlands: John Benjamins.
Levelt, W. J. (1999). Models of word production. *Trends in cognitive sciences*, *3*(6), 223–232. doi:10.1016/S1364-6613(99)01319-4
Levelt, W. J. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*(23), 13464–13471. doi:10.1073/pnas.231459498
Lindholm, C. (2008). Laughter, communication problems and dementia. *Communication & medicine*, *5*(1), 3–14. doi:10.1558/cam.v5i1.3
Locascio, J. J., Growdon, J. H., & Corkin, S. (1995). Cognitive test performance in detecting, staging and tracking Alzheimer's disease. *Archives of Neurology*, *52*, 1087–1099. doi:10.1001/archneur.1995.00540350081020
Lomas, J., Pickard, L., Bester, S., Erlbard, H., Finlayson, A., & Zoghaib, C. (1989). Communicative Effectiveness Index: Development and psychometric evaluation of a functional communication

measure for adult aphasia. *Journal of Speech and Hearing Disorders*, *54*, 113–124. doi:10.1044/jshd.5401.113
Lomas, J., Pickard, L., Bester, S., Erlbard, H., Finlayson, A., & Zoghaib, C. (2006). *Spørreskjema for nærpersoner. Communicative Effectiveness Index (CETI) på norsk* (L. Haaland-Johansen, M. Lind, & L. Hammersvik, translation). Oslo, Norway: Bredtvet kompetansesenter. (Original published 1989).
Machado, Á., Rodrigues, M., Simões, S., Santana, I., & Soares-Fernandes, J. (2010). The Portuguese who could no longer speak French: Primary progressive aphasia in a bilingual man. *The Journal of neuropsychiatry and clinical neurosciences*, *22*(1), E31–E32. doi:10.1176/jnp.2010.22.1.123.e31
Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). Lexical diversity and language development. *Quantification and Assessment*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.
Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language*, *19*, 124–141. doi:10.1016/0093-934X(83)90059-7
McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., … Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute of Aging and the Alzheimer's Association workgroup. *Alzheimer's and Dementia*, *7*, 263–269. doi:10.1016/j.jalz.2011.03.005
Mendez, M. F., Saghafi, S., & Clark, D. G. (2004). Semantic dementia in multilingual patients. *Journal of Neuropsychiatry and Clinical Neuroscience*, *16*, 381. doi:10.1176/jnp.16.3.381
Mickes, L., Wixted, J. T., Fennema-Notestine, C., Galasko, D., Bondi, M. W., Thal, L. J., & Salmon, D. P. (2007). Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *Neuropsychology*, *21*, 696–705. doi:10.1037/0894-4105.21.6.696
Moreaud, O., David, D., Charnallet, A., & Pellat, J. (2001). Are semantic errors actually semantic? Evidence from Alzheimer's disease. *Brain and Language*, *77*, 176–186. doi:10.1006/brln.2000.2427
Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes*, *21*, 238–290. doi:10.1080/01690960444000278
Naqvi, R. M., Haider, S., Tomlinson, G., & Alibhai, S. (2015). Cognitive assessments in multicultural populations using the Rowland Universal Dementia Assessment Scale: a systematic review and meta-analysis. *Canadian Medical Association Journal*, *187*(5), E169–E175. doi:10.1503/cmaj.140802
Nicholas, L. E., & Brookshire, R. H. (1995). Presence, completeness and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech and Hearing Research*, *38*, 145–156. doi:10.1044/jshr.3801.145
Nicholas, M., Obler, L. K., Au, R., & Albert, M. L. (1996). On the nature of naming errors in aging and dementia: A study of semantic relatedness. *Brain and Language*, *54*(2), 184–195. doi:10.1006/brln.1996.0070
Obler, L. K., & Albert, M. L. (1981). Language in the elderly aphasic and in the demented patient. In M. T. Sarno (Ed.), *Acquired aphasia* (pp. 385–398). New York, NY: Academic Press.
Paradis, M., & Libben, G. (1987). *The assessment of bilingual aphasia*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Pekkala, S., Wiener, D., Himali, J. J., Beiser, A. S., Obler, L. K., Liu, Y., … Au, R. (2013). Lexical retrieval in discourse: An early indicator of Alzheimer's dementia. *Clinical Linguistics & Phonetics*, *27*, 905–921. doi:10.3109/02699206.2013.815278
Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (Vol. 4). Reitan Neuropsychology. Tucson, AZ: Neuropsychological Press.
Rogers, T. T., Ivanoiu, A., Patterson, K., & Hodges, J. R. (2006). Semantic memory in Alzheimer's disease and the frontotemporal dementias: a longitudinal study of 236 patients. *Neuropsychology*, *20*(3), 319. doi:10.1037/0894-4105.20.3.319
Salmon, D. P., Heindel, W. C., & Lange, K. L. (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the

integrity of semantic memory. *Journal of the International Neuropsychological Society*, 5, 692–703. doi:10.1017/S1355617799577126

Sidnell, J. (2011). *Conversation Analysis. An introduction*. Oxford, UK: Wiley-Blackwell.

Storey, J. E., Rowland, J. T., Basic, D., Conforti, D. A., & Dickson, H. G. (2004). The Rowland Universal Dementia Assessment Scale (RUDAS): A multicultural cognitive assessment scale. *International Psychogeriatrics*, *16*, 13–31. doi:10.1017/S1041610204000043

Strobel, C., Johansen, H., Wetterberg, P., & Engedal, K. (2012). KT-NR. Norsk Revidert Klokketest. Oslo, Norway: Nasjonalt Kompetansesenter for Aldring og Helse.

Svennevig, J. (2003). Echo answers in native/non-native interaction. *Pragmatics*, *13*, 285–309.

Svennevig, J., & Lind, M. (2016). Dementia, interaction, and bilingualism: An exploratory case study. In H.-O. Enger, M. I. N. Knoph, K. E. Kristoffersen, & M. Lind (Eds.), *Helt fabelaktig! Festskrift til Hanne Gram Simonsen på 70-årsdagen* (pp. 199–218). Oslo, Norway: Novus.

Tarone, E. (1980). Communication strategies, foreigner talk and repair in interlanguage. *Language Learning*, *30*, 417–431. doi:10.1111/j.1467-1770.1980.tb00326.x

Veenstra, W. S., Huisman, M., & Miller, N. (2014). Age of acquisition and naming performance in Frisian-Dutch bilingual speakers with dementia. *Dementia & Neuropsychologia*, *8*, 249–259. doi:10.1590/S1980-57642014DN83000009

Vogel-Eyny, A., Galletta, E. E., Gitterman, M. R., & Obler, L. K. (2016). Lexical retrieval in healthy aging. In H.-O. Enger, M. I. N. Knoph, K. E. Kristoffersen, & M. Lind (Eds.), *Helt fabelaktig! Festskrift til Hanne Gram Simonsen på 70-årsdagen* (pp. 275–289). Oslo, Norway: Novus.

Wilkinson, R. (2007). Managing linguistic incompetence as a delicate issue in aphasic talk-in-interaction: On the use of laughter in prolonged repair sequences. *Journal of Pragmatics*, *39*(3), 542–569. doi:10.1016/j.pragma.2006.07.010

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., … Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069–2088. doi:10.1093/brain/awq129

Zacks, R., & Hasher, L. (1997). Cognitive gerontology and attentional inhibition: A reply to Burke and McDowd. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*(6), P274–P283. doi:10.1093/geronb/52B.6.P274

Appendix A: Transcription conventions (conversation)

(.)	Micro pause
(0.2)	Timed pause
[]	Overlapping speech
()	Uncertain transcription
(( ))	Non-verbal actions
<u>Under</u>	Emphasis
hh	Breath or laughter
.hh	Inbreath
=	Latched speech, continuation of talk
::	Elongated speech, a stretched sound

Appendix B: Semi-spontaneous narratives (concept analysis)[15]

T1: English	*T1: Norwegian*
she has asked him to get the … the … eh … the the chicks$\boxed{3}$ and the … and the … to get them down all of them and he agrees to do it and XX to do to do that and snaps and both <they> and the others are XX flat on the … and then we have the … taken to taken to the hospital$\boxed{8}$ and then the hospital$\boxed{9}$	en mann$\boxed{1}$ og … kone$\boxed{2}$ og … også det er en a man and … wife and … and there is a …ja <sette rygg på > … yes <put back on> og … så en … en kar han kommer opp og … prøver til å ta vekk … den … som som var der and …then a …a guy he comes up and … tries to take away … that … which which was there og så her så en en kom- … and then here then one one com- … ja … det … gikk ikke så god … yes. … it … didn't og so good og så så … det det kom ned begge to and then then … it it came down both of them og karen var tatt til … til … puh … mm til til til … er på er i … and the guy was taken to … to … puh …mm to to to … is on is in… *I know but I don't remember the names forget the names* så <*he was*> kjørt til … til … so <he was> driven to … to … ja … det var en som … nettopp før før du kom yes … there was one that … just befor before you came babu babu babu ba-$\boxed{8}$ babu babu babu ba- og han ligger i … i … and he lies in … in … *my wife spent fift- thirty years there bu*
T2: English	*T2: Norwegian*
there is a … first there is a fine … <X that X> chicken$\boxed{3}$ or the not chicken but eh … the- they're they're X getting they're they're getting … to to have a child	det er …to …to … there are … two … two … eh ja de som er X som hjelper der eh yes those who are X who are helping there det de er de er har sannsynligvis at de har vært X it they are they are have probably that they have been X

(*Continued*)

[15]The texts are segmented into AS-units (analysis of speech units) (Foster, Tonkyn, & Wigglesworth, 2000) divided by extra interlinear space in the transcription. The relevant key concept components (cf. Table 2) are marked by numbered boxes. Utterances in the non-target language are italicised. Three dots indicate pauses, unintelligible utterances are marked by X and X= , and uncertain interpretations are indicated by brackets.

(Continued).

T2: English	*T2: Norwegian*
and they're … putting him in a place to be	… og så så er det også sånn … and then then it is also such
and these … people are … they're trying to fix	… det de det de sitter oppi i topp oppi topp X= … it they it they are sitting up in the top up in the top X=
… waf- they should watch what they do … this	… ja og og sitter der oppe … yes and and are sitting there up
… right	og så … og de tenkt hva skal vi gjøre hva and then … and they thouht what shall we do what
… and then then … the man decides that he he will go up $\boxed{4}$…and … bring him bring this down	og så … de sitter … de bruker å … sitte … eh sånne … sitter and then … they sit … they use to … sit … eh such … sit
and the X … this is persos wants	… og og kaster den ned … and and throw it down
… the… the woman wasn't much … didn't much care one way or the other	… det X en annen X ut til den … åh … X= … it X another X out to that … oh … X=
eh but he …he's heavy and snaps eh … so that the the … the the <X others X> have to get some- somewhere else	og det har X eller X= and it has X or X=
and this is … four children … standing outside	og så … det er han X nedover <X for langt X> and then … it is him X downwards <X too long X>
there is a …	og han falt ned fra $\boxed{6}$… X= and he fell down from … X=
oh the man fell down $\boxed{6}$… and	falt til til … til fell to to … to
… I am not quite sure but I think that the woman is helping that person maybe because X eh eh because of of the <X chicks X>	… og … så det… de folk X og så under X= … and … so it … those people X and then under X=
eh this person has been … this is called fallen down on	… X=det er … X=it is … sist- … las-
and has to … eh get the … car <X now X> to	… hvert fall de har sitter og venter til … hjelp … anyway they have sit and wait to … help
the truck	og det var X= and it was X=
to take him to the … to to the …closed closed place… which is … here	til til når de er i till till when they are in
and they get they do get	… åh … oh
eh eh break a a leg and then have this X	… de er i … they are in
	… X= … X= … men i hvert fall … de kommer … ser … but anyway … they come … see

Part III

Appendices

List of appendices

Appendix A: Letters to participants

The next couple of pages include the information letters sent to participants in the different studies:

1. Information letter with consent form to persons with AD and PPA (Study II and III)

2. Additional information to persons with AD and PPA (Study II and III)

3. Information letter to control participants (Study II and III)

4. Information letter to JJ (Study IV)

UiO : **Universitetet i Oslo**
Humanistisk fakultet/Institutt for lingvistiske og nordiske studier

FORESPØRSEL OM DELTAKELSE I ET FORSKNINGSPROSJEKT OM SPRÅK, ALDRING OG DEMENS

Dette er et spørsmål til deg om å delta i et forskningsprosjekt for å undersøke hvordan språket utvikler seg når vi blir eldre og eventuelt får vansker med hukommelsen eller en demensdiagnose.

Du blir spurt om å være med i denne studien siden du nylig har vært til utredning ved hukommelsesklinikken på Ullevål sykehus.

Universitet i Oslo er ansvarlig for studien, men samarbeider med hukommelsesklinikken på Ullevål Sykehus. Vi rekrutterer også personer som har vært til utredning for litt lenger tid tilbake, og som har opplyst om at de vil delta i forskning.

FORMÅLET MED PROSJEKTET

Prosjektet undersøker språklige ferdigheter hos personer med og uten demens og i forskjellige livsfaser. I norsk sammenheng finnes det veldig lite forskning på språkbruk og språkferdigheter i et livsløpsperspektiv.

Når vi blir eldre vil mange oppleve at det er vanskelig å komme på de riktige ordene, eller det kan være vanskelig å følge med på historier andre forteller. Dette er normalt ved aldring, men det kan også forverres hvis man får en demensdiagnose. Ulike demenssykdommer kan også påvirke språkevnen ulikt.

I dette prosjektet prøver jeg å finne ut av hva som er normale språkendringer ved aldring, og om det finnes ulike «språkprofiler» for ulike demenssykdommer.

HVA INNEBÆRER PROSJEKTET?

Deltakelse i prosjektet innebærer at vi møtes én til to ganger, enten på Universitetet i Oslo, hjemme hos deg eller et annet sted du ønsker. Vi skal da gjøre noen oppgaver sammen, både på datamaskin og ved hjelp av penn og papir. I en oppgave vil jeg også filme øyebevegelsene dine ved hjelp av et lite kamera på en datamaskin. Til hvert møte setter vi av ca. 2 timer, og det er mulig å ta mange pauser underveis.

Under oppgavene vil jeg gjøre lydopptak, og noe av det du sier vil senere bli skrevet ned sånn at det kan brukes i forskningsartikler og i oppgaven min. Når materialet brukes i artikler, vil all informasjon om deg bli slettet, og ingen vil kunne kjenne deg igjen.

Jeg vil også gjennomføre et kort intervju med deg, der jeg spør litt om bakgrunnen din og om hvilke språk du snakker og har lært oppgjennom livet.

I prosjektet vil vi innhente og registrere opplysninger om deg. Hvis det ble tatt bilder av hjernen din, eller prøver av spinalvæske som ledd av utredningen, vil jeg få skriftlige beskrivelser av funnene fra disse testene fra en forskningssykepleier på Ullevål Sykehus. Jeg vil ikke få de faktiske prøvene eller bildene. Denne informasjonen oppbevares trygt på en innelåst server, og ingen andre enn meg får tilgang til denne informasjonen om deg.

Prosjektet avsluttes i august 2019, men informasjonen vil oppbevares trygt fram til juli 2024, deretter blir all informasjon om deg slettet.

MULIGE FORDELER OG ULEMPER

Det er ikke forbundet noen risiko med denne studien. Men oppgavene kan ta litt lang tid og du kan bli sliten underveis. Det er derfor viktig at du sier tydelig ifra når du ønsker pauser.

FRIVILLIG DELTAKELSE OG MULIGHET FOR Å TREKKE SITT SAMTYKKE

Det er frivillig å delta i prosjektet. Dersom du ønsker å delta, undertegner du samtykkeerklæringen på siste side. Du kan når som helst og uten å oppgi noen grunn trekke ditt samtykke. Du kan når som helst be om å stoppe lydopptak under testingen, eller si at du ikke vil gjennomføre enkelte oppgaver. Du kan når som helst trekke deg fra studien uten å oppgi noen grunn.

Dersom du trekker deg fra prosjektet, kan du kreve å få slettet innsamlede prøver og opplysninger, med mindre opplysningene allerede er inngått i analyser eller brukt i vitenskapelige publikasjoner. Dersom du senere ønsker å trekke deg eller har spørsmål til prosjektet, kan du kontakte prosjektleder Ingeborg Sophie Ribu på e-post i.s.b.ribu@iln.uio.no eller telefon 48 10 56 59.

HVA SKJER MED INFORMASJONEN OM DEG?

Informasjonen som registreres om deg skal kun brukes slik som beskrevet i hensikten med studien. Du har rett til innsyn i hvilke opplysninger som er registrert om deg og rett til å få korrigert eventuelle feil i de opplysningene som er registrert.

Alle opplysningene vil bli behandlet uten navn og fødselsnummer eller andre direkte gjenkjennende opplysninger. En kode knytter deg til dine opplysninger gjennom en navneliste. Det er kun prosjektleder (Ingeborg Ribu) som har tilgang til denne listen.

Prosjektleder har ansvar for den daglige driften av forskningsprosjektet og at opplysninger om deg blir behandlet på en sikker måte. Informasjon om deg vil bli anonymisert eller slettet senest fem år etter prosjektslutt.

OPPFØLGINGSPROSJEKT

Hvis du synes det er greit vil jeg gjerne be om muligheten til å kontakte deg igjen ved en senere anledning for å delta i nye prosjekter senere.

FORSIKRING

Universitetet i Oslo er selvassurandør. Det er ikke forventet at det kan oppstå skade forbundet med deltakelse i prosjektet, men hvis det skulle skje vil Universitetet i Oslo dekke utgiftene.

GODKJENNING

Prosjektet er godkjent av Regional komite for medisinsk og helsefaglig forskningsetikk, REK (2016/1293).

Etter ny personopplysningslov har behandlingsansvarlig og prosjektleder Ingeborg Ribu et selvstendig ansvar for å sikre at behandlingen av dine opplysninger har et lovlig grunnlag. Dette prosjektet har rettslig grunnlag i EUs personvernforordning artikkel 6.1a og 9.2a om samtykke.

Du har rett til å klage på behandlingen av dine opplysninger til Datatilsynet.

KONTAKTOPPLYSNINGER

Dersom du har spørsmål til prosjektet kan du ta kontakt med Ingeborg Ribu på telefon 48 10 56 59 eller e-post i.s.b.ribu@iln.uio.no .

Du kan ta kontakt med institusjonens personvernombud dersom du har spørsmål om behandlingen av dine personopplysninger i prosjektet. Maren Magnus Voll er personvernombud ved Universitetet i Oslo. Hun kan nås på e-post personvernombud@uio.no eller telefon: 22 85 97 78

SAMTYKKE TIL DELTAKELSE I PROSJEKTET

JEG ER VILLIG TIL Å DELTA I PROSJEKTET

Sted og dato | Deltakers signatur

Deltakers navn med trykte bokstaver

Jeg samtykker til at det gjøres lydopptak som også kan brukes i senere forskning   Ja____   Nei____

Jeg samtykker til å bli kontaktet igjen senere for deltakelse i andre studier   Ja____   Nei____

STEDFORTREDENDE SAMTYKKE

Dersom primærdeltaker ikke lenger har samtykkekompetanse, kan pårørende gi samtykke på vedkommende sine vegne.

Som nærmeste pårørende til _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ (Fullt navn) samtykker jeg til at hun/han kan delta i prosjektet.

Sted og dato | Pårørendes signatur

Pårørendes navn med trykte bokstaver

Jeg bekrefter å ha gitt informasjon om prosjektet

Sted og dato | Signatur

Rolle i prosjektet

UiO **Universitetet i Oslo**
Humanistisk fakultet/Institutt for lingvistiske og nordiske studier

Informasjon om studien «Språk og kognisjon i et livsløpsperspektiv»

Mitt navn er Ingeborg Ribu,
Jeg er språkforsker og jobber ved Universitetet i Oslo. Der skriver jeg en doktorgrad om hvordan språket utvikler seg gjennom livet, og hvordan det kan påvirkes av hukommelsesvansker. Prosjektet har arbeidstittel «Språk og kognisjon i et livsløpsperspektiv»

Når vi blir eldre vil mange oppleve at det er vanskelig å komme på de riktige ordene, eller å følge med i samtaler der mange deltar. Dette er vanlig ved aldring, men noen opplever mer av dette enn andre. For å få god oversikt over hvor vanlig dette er, undersøker vi språkferdigheter hos ulike deltakergrupper. I studien vil det derfor inngå personer i ulike aldersgrupper, og personer med og uten hukommelsesvansker.

I undersøkelsen vil vi gjøre ulike språklige oppgaver sammen. De fleste gjennomføres med penn og papir, men det er også en oppgave på datamaskin. Det er derfor fint om deltakerne kan komme til Universitetet i Oslo for å gjennomføre undersøkelsen, men vi har også mulighet til å møtes andre steder.

Undersøkelsen tar ca. 2,5 til 3 timer, inkludert pauser underveis. Noen synes det er greit å gjennomføre alt i løpet av ett møte, mens andre vil gjerne dele det opp. Det er opp til den enkelte deltaker hvordan man ønsker det.

Det blir gjort lydopptak av alle undersøkelsene, men det er kun for at testleder skal slippe å notere så mye underveis. Noe vil også bli transkribert, men da blir all informasjon om personen som har sagt det anonymisert, så ingen kan kjenne igjen deltakeren hvis dette brukes i publikasjoner senere.

Hvis du først har sagt ja til å delta i studien, men så ombestemmer deg, er det helt OK. Husk å gi beskjed sånn at jeg får slettet all informasjon som er samlet inn om deg.

Ta kontakt dersom du har spørsmål angående studien.

Håper du ønsker å delta!

Med vennlig hilsen,

Ingeborg Ribu
Doktorgradsstipendiat
Forskergruppe for klinisk lingvistikk og språktilegnelse
Universitetet i Oslo

Tlf: 48 10 56 59
E-post: i.s.b.ribu@iln.uio.no

Postadresse: Postboks 1102 Blindern 0317 Oslo
E-post: i.s.b.ribu@iln.uio.no
Telefon: 48 10 56 59
www.uio.no

UiO **Universitetet i Oslo**
Humanistisk fakultet/Institutt for lingvistiske og nordiske studier

FORESPØRSEL OM DELTAKELSE I ET FORSKNINGSPROSJEKT OM SPRÅK, ALDRING OG DEMENS

Dette er et spørsmål til deg om å delta i et forskningsprosjekt for å undersøke hvordan språklige ferdigheter endres i løpet av livet, sammenlignet med hvordan språket kan bli påvirket ved en demensdiagnose. For deltakelse i prosjektet er vi interessert i å komme i kontakt med voksne personer over 65 år, uten en kjent demensdiagnose, samt yngre voksne mellom 18 og 30 år.

FORMÅLET MED PROSJEKTET

Prosjektet undersøker språklige ferdigheter hos personer med og uten demens og i forskjellige livsfaser. I norsk sammenheng finnes det veldig lite forskning på språkbruk og språkferdigheter hos eldre personer.

Når vi blir eldre vil mange oppleve at det er vanskelig å komme på de riktige ordene, eller det kan være vanskelig å følge med på historier andre forteller. Dette er normalt ved aldring, men det kan også forverres hvis man får en demensdiagnose. Ulike demenssykdommer kan også påvirke språkevnen ulikt.

I dette prosjektet prøver jeg å finne ut av hva som er normale språkendringer ved aldring, og om det finnes ulike «språkprofiler» for ulike demenssykdommer.

HVA INNEBÆRER PROSJEKTET?

Deltakelse i prosjektet innebærer at vi møtes en gang på Universitetet i Oslo. Vi skal da gjøre noen oppgaver sammen, både på datamaskin og ved hjelp av penn og papir. I en oppgave vil jeg også filme øyebevegelsene dine ved hjelp av et lite kamera på en datamaskin. Vi setter av ca. tre timer til gjennomføring av oppgavene, og det er satt av tid til opptil flere pauser underveis.

Under oppgavene vil jeg gjøre lydopptak, og noe av det du sier vil senere bli skrevet ned sånn at det kan brukes i forskningsartikler og i oppgaven min. Når materialet brukes i artikler, vil all informasjon om deg bli slettet, og ingen vil kunne kjenne deg igjen.

Du vil også motta et spørreskjema som du sender tilbake i posten. I skjemaet er det noen spørsmål om bakgrunnen din og om hvilke språk du snakker og har lært oppgjennom livet.

Prosjektet avsluttes i august 2019, men informasjonen vil oppbevares trygt fram til juli 2024, deretter blir all informasjon om deg slettet.

MULIGE FORDELER OG ULEMPER

Det er ikke forbundet noen risiko med denne studien. Men oppgavene kan ta litt lang tid og du kan bli sliten underveis. Det er derfor viktig at du sier tydelig ifra når du ønsker pauser.

FRIVILLIG DELTAKELSE OG MULIGHET FOR Å TREKKE SITT SAMTYKKE

Det er frivillig å delta i prosjektet. Dersom du ønsker å delta, undertegner du samtykkeerklæringen på siste side. Du kan når som helst og uten å oppgi noen grunn trekke ditt samtykke. Du kan når som helst be om å stoppe lydopptak under testingen, eller si at du ikke vil gjennomføre enkelte oppgaver. Du kan når som helst trekke deg fra studien uten å oppgi noen grunn.

Dersom du trekker deg fra prosjektet, kan du kreve å få slettet innsamlede prøver og opplysninger, med mindre opplysningene allerede er inngått i analyser eller brukt i vitenskapelige publikasjoner. Dersom du senere ønsker å trekke deg eller har spørsmål til prosjektet, kan du kontakte prosjektleder Ingeborg Sophie Ribu på e-post i.s.b.ribu@iln.uio.no eller telefon 48 10 56 59.

HVA SKJER MED INFORMASJONEN OM DEG?

Informasjonen som registreres om deg skal kun brukes slik som beskrevet i hensikten med studien. Du har rett til innsyn i hvilke opplysninger som er registrert om deg og rett til å få korrigert eventuelle feil i de opplysningene som er registrert.

Alle opplysningene vil bli behandlet uten navn og fødselsnummer eller andre direkte gjenkjennende opplysninger. En kode knytter deg til dine opplysninger gjennom en navneliste. Det er kun prosjektleder (Ingeborg Ribu) som har tilgang til denne listen.

Prosjektleder har ansvar for den daglige driften av forskningsprosjektet og at opplysninger om deg blir behandlet på en sikker måte. Informasjon om deg vil bli anonymisert eller slettet senest fem år etter prosjektslutt.

OPPFØLGINGSPROSJEKT

Hvis du synes det er greit vil jeg gjerne be om muligheten til å kontakte deg igjen ved en senere anledning for å delta i nye prosjekter senere.

FORSIKRING

Universitetet i Oslo er selvassurandør. Det er ikke forventet at det kan oppstå skade forbundet med deltakelse i prosjektet, men hvis det skulle skje vil Universitetet i Oslo dekke utgiftene.

GODKJENNING

Prosjektet er godkjent av Regional komite for medisinsk og helsefaglig forskningsetikk, REK (2016/1293).

Etter ny personopplysningslov har behandlingsansvarlig og prosjektleder Ingeborg Ribu et selvstendig ansvar for å sikre at behandlingen av dine opplysninger har et lovlig grunnlag. Dette prosjektet har rettslig grunnlag i EUs personvernforordning artikkel 6.1a og 9.2a om samtykke.

Du har rett til å klage på behandlingen av dine opplysninger til Datatilsynet.

KONTAKTOPPLYSNINGER

Dersom du har spørsmål til prosjektet kan du ta kontakt med Ingeborg Ribu på telefon 48 10 56 59 eller e-post i.s.b.ribu@iln.uio.no .

Du kan ta kontakt med institusjonens personvernombud dersom du har spørsmål om behandlingen av dine personopplysninger i prosjektet. Maren Magnus Voll er personvernombud ved Universitetet i Oslo. Hun kan nås på e-post personvernombud@uio.no eller telefon: 22 85 97 78

SAMTYKKE TIL DELTAKELSE I PROSJEKTET

JEG ER VILLIG TIL Å DELTA I PROSJEKTET

Sted og dato	Deltakers signatur
	--
	Deltakers navn med trykte bokstaver

Jeg samtykker til at det gjøres lydopptak som også kan brukes i senere forskning Ja____ Nei____

Jeg samtykker til å bli kontaktet igjen senere for deltakelse i andre studier Ja____ Nei____

Jeg bekrefter å ha gitt informasjon om prosjektet

Sted og dato	Signatur
	--
	Rolle i prosjektet

MultiLing Dementia – en studie av kommunikasjon hos flerspråklige

Kjære deltaker,

Når man blir eldre kan man ofte oppleve at det er vanskelig å finne de riktige ordene eller setningene, og at det påvirker kommunikasjon med andre. Dette er veldig vanlig for personer med demens, men det er også normalt for personer uten demens. Per i dag finnes det ikke noe verktøy som enkelt kan skille mellom vanlige og uvanlige språkvansker hos eldre. Vi er derfor interessert i å undersøke språket hos personer med og uten demens. Ansatte i helsevesenet og pårørende til personer med demens vil også ha nytte av å lære mer om språkvansker hos personer som har en demensdiagnose.

I den sammenheng ønsker vi å gjøre opptak av en- og flerspråklige personer, både med og uten demens. For de flerspråklige personene trenger vi opptak på norsk og på morsmålet. Deltakerne i undersøkelsen vil bli bedt om å fortelle en historie, og å svare på noen oppgaver. Det vil også bli avholdt et intervju med deltakerne og deres pårørende. Det kan hende vi trenger å treffe deg mer enn en gang.

Hvis du er interessert i å delta i undersøkelsen, vil vi gjerne at du signerer samtykkeskjemaet under og sender det tilbake til oss.

All informasjon er konfidensiell og vil bli anonymisert i formidling fra prosjektet. Det gjelder også alle opptakene vi gjør. Data fra prosjektet vil bli analysert og gjengitt på en slik måte at ingen kan kjennes igjen, og ingen kan vite hvem som har deltatt i undersøkelsen. Alt vi samler inn vil kun være tilgjengelige for forskere tilknyttet prosjektet. Alle dokumenter vil slettes etter prosjektslutt for å ivareta konfidensialitet.

Prosjektet avsluttes i løpet av 2015. De redigerte opptakene vil bli oppbevart i inntil 20 år etter prosjektslutt for sammenlikning med senere studier.

Din deltakelse i prosjektet er svært verdifull for oss for å forstå mer om hva som skjer med språket når vi blir eldre. Dette vil være med på å øke vår teoretiske kunnskap rundt temaet, samtidig som det kan være til nytte for personer som av profesjonelle eller personlige grunner har daglig omgang med personer med en (mulig) demensdiagnose.

Prosjektet er meldt inn til Personvernombudet for forskning, Norsk Samfunnsvitenskapelig Datatjeneste. Hvis du skulle ha noen spørsmål rundt undersøkelsen, eller vite noe om resultatene fra prosjektet, er det bare å ta kontakt med prosjektgruppa via adressen nederst på arket. Deltakelse i prosjektet er frivillig, og du kan når som helst (fram til prosjektslutt) trekke deg uten å måtte oppgi noen grunn om du ikke lenger ønsker å ta del i undersøkelsen. Dette gjøres på e-post eller telefon til prosjektleder Professor Bente Ailin Svendsen. Dersom du ønsker å trekke deg vil all data om deg bli slettet.

Tusen takk!

Vennlig hilsen (på vegne av forskerteamet)

Forskerteam:

Prosjektleder:	Professor Bente Ailin Svendsen (Universitetet i Oslo)
Medlemmer av teamet:	Professor Hanne Gram Simonsen (Universitetet i Oslo)
	Professor Jan Svennevig (Universitetet i Oslo)
	Dr. Art. Marianne Lind (Statped)
Prosjektassistent:	Ingeborg Sophie Ribu (Universitetet i Oslo)

Bente Ailin Svendsen
MultiLing, Institutt for lingvistiske og nordiske studier, Universitetet i Oslo
Postboks 1102, Blindern
0317 Oslo
b.a.svendsen@iln.uio.no
22856966

Appendix B: Notification to the data protection officials

On these pages, the responses from the data protection officials are included:

1. Acceptance for studies II and III from the Regional committees for medical and health research ethics

2. acceptance after changes to the project plan for studies II and III were implemented

3. Remit Assessment from the Regional committees for medical and healthy research ethics for study IV. The committee deemed the project to not be approved by them, but by the Norwegian Centre for Research Data.

4. Approval from the Norwegian Centre for Research Data (NSD)

REK
REGIONALE KOMITEER FOR MEDISINSK OG HELSEFAGLIG FORSKNINGSETIKK

Region:	Saksbehandler:	Telefon:	Vår dato:	Vår referanse:
REK sør-øst	Mariann Glenna Davidsen	22845526	20.12.2016	2016/1293 REK sør-øst B
			Deres dato:	**Deres referanse:**
			10.11.2016	

Vår referanse må oppgis ved alle henvendelser

Ingeborg Sophie Ribu
Universitetet i Oslo

2016/1293 Demens og språklig aldring

Forskningsansvarlig: Universitetet i Oslo
Prosjektleder: Ingeborg Sophie Ribu

Vi viser til tilbakemelding på ovennevnte forskningsprosjekt. Tilbakemeldingen ble behandlet av Regional komité for medisinsk og helsefaglig forskningsetikk (REK sør-øst) i møtet 30.11.2016. Vurderingen er gjort med hjemmel i helseforskningsloven (hfl.) § 10, jf. forskningsetikkloven § 4.

Prosjektleders prosjektbeskrivelse
Prosjektets formål er å undersøke hvordan språk påvirkes av normal og patologisk aldring, både når det gjelder produksjon og forståelse av ordforråd og grammatikk i talt språk. Prosjektet har fire forskningsspørsmål: 1- Hvordan påvirkes forståelse ordforråd og grammatikk av en aldersrelatert kognitiv svekkelse (f.eks. demensdiagnose)? 2- Hvordan påvirkes produksjon av ordforråd og grammatikk av en aldersrelatert kognitiv svekkelse? 3- Planlegger personer med og uten demens tale annerledes? 4- Kan ulike lingvistiske trekk brukes som klinisk markør ved aldersrelatert kognitiv svekkelse? Det vil bli utført psykolingvistiske tester med ulike deltakergrupper for å undersøke ordgenerering og spontantale, og for å måle forståelse og gjenkjenning av ord.

Saksgang
Komiteen behandlet søknaden første gang i møte den 17.08.2016. I sitt brev datert 14.09.2016, skrev komiteen at den utsatte å treffe et endelig vedtak i saken da den ønsket utfyllende opplysninger på enkelte punkter. Prosjektleder sendte inn en tilbakemelding mottatt den 10.11.2016, inkludert revidert protokoll samt reviderte informasjons- og samtykkeskriv.

Komiteens vurdering ved første gangs behandling (gjengitt fra vedtak av 14.09.2016)
Komiteen forstår det slik at man i dette prosjektet ønsker å undersøke forskjeller i språkprosessering mellom personer med og uten kognitiv svikt. Det skal rekrutteres 90 personer fordelt på 3 grupper (personer over 65 år med og uten demens, samt unge voksne som utgjør kontrollgruppen).

Metoder
Det skal innhentes tidligere helseopplysninger fra pasientjournal om demensdiagnose, enten fra fastlege eller utreder ved hukommelsesklinikk («journal fra demensutredning»). Alle deltakerne vil gjennomgå 4 språkoppgaver for å samle inn data om produksjon og forståelse av enkeltord og setninger. Det skal brukes eye-tracker i noen oppgaver for å registrere øyebevegelser og blikkvarighet. Under oppgaveløsningene gjøres lydopptak. Disse anonymiseres og vil kunne bli gjort tilgjengelige for andre forskere gjennom

databasen "Disordered Speech Bank". I tillegg skal deltakerne screenes med tanke på global kognitiv funksjon (Mini-Mental State Evaluation), eksekutivfunksjoner og arbeidsminne. De skal også besvare spørreskjema om hukommelse, kommunikasjon og språkbakgrunn.

Komiteen kan ikke se det er beskrevet nærmere i protokoll hvilke metoder som skal benyttes for testing av eksekutive funksjoner og arbeidsminne. Det er også uklart hva som skal innhentes av informasjon fra pasientjournal.

Rekruttering
Det vises til at deltakere med demens vil rekrutteres gjennom sykehjem og dagsentere i Oslo, samt gjennom kontakt med hukommelsesklinikk og Nasjonalforeningen for folkehelsen samt demenskoordinatorer i de bydelene som har det. De eldre deltakerne uten demens vil i stor grad rekrutteres gjennom personlige nettverk og andre interesse-nettverk. De yngre deltakerne vil rekrutteres gjennom personlige nettverk og oppslag ved Universitetet i Oslo.

Det er ikke tydelig beskrevet i protokoll hvordan og hvem som skal vurdere samtykkekompetanse når det gjelder rekruttering av pasienter med demens diagnose.

Informasjon- og samtykkeskriv
Det er beskrevet at alle deltakerne vil bli gitt skriftlig og muntlig informasjon, samt at alle skal gi skriftlig samtykke til å delta i studien (for noen deltakere vises det til at pårørende skal spørres).

Det er kun vedlagt ett informasjon- og samtykkeskriv. Det kan synes som dette informasjonsskrivet skal dekke alle de tre deltakende gruppene. Det er i så fall ikke tilstrekkelig. Det må utarbeides egne, tilpassede informasjons- og samtykkeskriv til hver av gruppene da de har ulike forutsetninger for deltakelse.

Før komiteen tar endelig stilling til prosjektet vil det være nødvendig å få en nærmere redegjørelse av følgende:

1. Rekrutteringsprosedyren/ hvem skal vurdere samtykkekompetanse

2. Hvilke metoder som skal benyttes for testing av eksekutive funksjoner og arbeidsminne

3. Hvilke info skal hentes fra journal.

4. Utarbeide infoskriv til alle de deltakende gruppene. Disse må sendes komiteen for gjennomgang

Komiteens beslutning
Vedtak i saken utsettes. Komiteen tar stilling til prosjektet ved mottatt svar.

Prosjektleders tilbakemelding
Prosjektleders tilbakemelding av 11.10.2016 gjengis i sin helhet:

1. Samtykkekompetanse: Dersom forsøkspersonene med en demensdiagnose har manglene samtykkekompetanse, vil dette framkomme av informasjon fra utredende lege. Jeg har inngått et samarbeid med hukommelsesklinikken ved Oslo Universitetssykehus, Ullevål om at de kan gi informasjon om prosjektet, og henvise aktuelle kandidater til mitt prosjekt. Disse personene vil være utredet ved hjelp av et større klinisk verktøy, og utredende lege vil vurdere personenes samtykkekompetanse. I det nye samtykkeskjemaet (vedlegg 1) er det også satt av plass til at pårørende gir samtykke.

2. Kognitive funksjoner og arbeidsminne: I samarbeid med hukommelsesklinikken ved Oslo Universitetssykehus, Ullevål, har jeg utarbeidet en protokoll basert på deres testmateriale og utredningsprotokoll (vedlegg 2 og 3). Testene som inngår i testprotokollen for kognitive ferdigheter vil være Mini Mental State Evaluation, 10-ordstest fra CERAD, verbal flyt, forkortete Boston Naming Test, test for abstrakt tenking (likheter mellom ord), Klokketegning, figurkopiering, Trail Marking A+B. Videre bruker jeg en elektronisk versjon av Flanker (Eriksen, 1972) og Stroop (D-KEFS), tallrangering og baklengs *tallhukommelse, to tester for prosesseringshastighet (letter comparison og "boxes"), og pluss/minus-regning for set-shifting.*

Det er også blitt gjort endringer når det gjelder de språklige testene. Som før vil jeg benytte meg av en bildebeskrivelsesoppgave for å ellisitere spontantale/narrativproduksjon, men bildet er byttet ut. I den opprinnelige søknaden var det tenkt at jeg skulle bruke "kaketyveriet" fra Boston Naming Test, men jeg

har nå bestemt meg for "fuglehistorien", en tegneserie fra Bilingual Aphasia Test. Bildebenevnelsestesten forblir den samme, men jeg har lagt til en Time-Reference test (TRT) for produksjon av setninger. TRT innebærer grammatisk manipulasjon av setninger, fra fortid til framtid og omvendt.

Ordgjennkjennelsesoppgaven er byttet ut med en ord-og-bildematchingsoppgave for å undersøke forståelse av enkeltord, mens eye-trackingoppgaven forblir som tenkt i den opprinnelige søknaden. Videre har jeg lagt til to tester for å undersøke semantisk kunnskap, nemlig en ordassosiasjonstest der deltakerne blir forelagt 90 ord og skal skrive ned det første ordet de tenker på til hvert ord, og Pyramide- og palmetesten som ofte brukes for å undersøke semantisk kunnskap hos personer med afasi. I pyramide- og palmetesten får deltakerne se tre bilder, og de må avgjøre hvilke to som hører sammen.

3. Informasjon fra journal: Siden jeg har inngått et samarbeid med hukommelsesklinikken ved Ullevål sykehus, og de vil hjelpe meg med å finne personer som kan være med i studien, kan jeg være nokså sikker på at alle forskningspersonene som inngår i gruppa med demens har vært utredet med det samme verktøyet. Informasjon jeg vil få tilgang til fra journalene deres er hva slags demensdiagnose de har fått og når denne ble stilt, informasjon om hvilke tester de er blitt utredet med og resultatet på disse. Dersom det foreligger spinalprøver og MR-bilder, vil jeg få beskrivelse av resultatene fra disse, men jeg vil ikke få tilgang til de faktiske bildene og prøveresultatene. Jeg vil også søke NorKog om å få tilgang til pasienter som er registrert i dette registeret. Det etter oppfordring fra Hukommelsesklinikken ved OUS Ullevål.

4. Nye informasjonsskriv og samtykkeskjemaer er vedlagt. Vedlegg 1 er et nytt samtykkeskjema til gruppen med demenspasienter, og deres pårørende. Vedlegg 4 er et nytt samtykkeskjema til de to gruppene med forsøkspersoner uten demens, og vedlegg 5 er et informasjonsskriv som følger med samtykkeskjemaene og leveres til alle deltakere. Det er tenkt som mer informasjon enn det som framkommer i samtykkeskjemaene, og kan derfor fungere fint som supplement også til pårørende av forsøkspersonene som deltar i demensgruppen.

Komiteens vurdering
Komiteen mener prosjektleders tilbakemeldinger er tilfredsstillende besvart, men vil kommentere at samtykkeskrivene bør følge REK sine anbefalte maler. Det vil ikke være nødvendig med egne avkrysninger i samtykkeskrivet for annet enn spørsmål om å bli kontaktet for andre prosjekter i fremtiden. Komiteen setter dermed følgende vilkår til prosjektet:

1. Utbedre informasjon- og samtykkeskriv slik at de følger REK mal. I tillegg fjerne unødvendig avkryssingsalternativer.

Vedtak
Komiteen godkjenner prosjektet i henhold til helseforskningsloven, med forutsetning om at ovennevnte vilkår oppfylles, jf. § 9 og § 33.

Når vilkåret er oppfylt gjelder godkjenningen under forutsetning av at prosjektet gjennomføres slik det er beskrevet i søknaden.

Tillatelsen gjelder til 14.04.2019. Av dokumentasjonshensyn skal opplysningene likevel bevares inntil 14.04.2024. Opplysningene skal lagres avidentifisert, dvs. atskilt i en nøkkel- og en opplysningsfil. Opplysningene skal deretter slettes eller anonymiseres, senest innen et halvt år fra denne dato.

Forskningsprosjektets data skal oppbevares forsvarlig, se personopplysningsforskriften kapittel 2, og Helsedirektoratets veileder "*Personvern og informasjonssikkerhet i forskningsprosjekter innenfor helse- og* *omsorgssektoren".*

Sluttmelding og søknad om prosjektendring
Dersom det skal gjøres endringer i prosjektet i forhold til de opplysninger som er gitt i søknaden, må prosjektleder sende endringsmelding til REK. Prosjektet skal sende sluttmelding på eget skjema, se helseforskningsloven § 12, senest et halvt år etter prosjektslutt.

Klageadgang
Komiteens vedtak kan påklages, jf. forvaltningslovens § 28 flg. Klagen sendes til REK sør-øst B. Klagefristen er tre uker fra du mottar dette brevet. Dersom vedtaket opprettholdes av REK sør-øst B, sendes klagen videre til Den nasjonale forskningsetiske komité for medisin og helsefag for endelig vurdering.

Komiteens avgjørelse var enstemmig.

Med vennlig hilsen

Grete Dyb
professor, dr. med.
leder REK sør-øst B

Mariann Glenna Davidsen
rådgiver

Kopi til:
- *Universitetet i Oslo ved øverste administrative ledelse*

REK
REGIONALE KOMITEER FOR MEDISINSK OG HELSEFAGLIG FORSKNINGSETIKK

Region:	Saksbehandler:	Telefon:	Vår dato:	Vår referanse:
REK sør-øst	Ingrid Dønåsen	22845523	15.06.2018	2016/1293 REK sør-øst B
			Deres dato:	**Deres referanse:**
			16.05.2018	

Vår referanse må oppgis ved alle henvendelser

Ingeborg Sophie Ribu
Universitetet i Oslo

2016/1293 Demens og språklig aldring

Forskningsansvarlig: Universitetet i Oslo
Prosjektleder: Ingeborg Sophie Ribu

Vi viser til søknad om prosjektendring datert 16.05.2018 for ovennevnte forskningsprosjekt, samt e-post med supplerende opplysninger mottatt 14.06.2018. Søknaden er behandlet av sekretariatet i REK sør-øst på delegert fullmakt fra REK sør-øst B, med hjemmel i helseforskningsloven § 11.

Endringene innebærer:
- Endring i inklusjonskriterier: "*Fra mai 2018 vil også pasienter med diagnosen Primær progressiv afasi (PPA) inngå i studien dersom de har patologi som tilsvarer Alzheimers sykdom.*". Prosjektleder opplyser at ved å inkludere pasienter med Primær progressiv afasi i tillegg til pasienter med Alzheimer vil man kunne finne ut mer om hvordan språkferdigheter påvirkes i de ulike diagnosene, og hvordan man eventuelt ved hjelp av språkprofiler kan skille dem fra hverandre.
- Ny test for semantisk assosiasjon: The Semantic Association test.
- Protokollen (mottatt på e-post 14.06.2018) er oppdatert i henhold til endringene
- Ny kontaktperson ved forskningsansvarlig institusjon: Gunn-Elin Aa. Bjørneboe

Vurdering
Sekretariatet i REK har vurdert de omsøkte endringene, og har ingen forskningsetiske innvendinger til endringene slik de er beskrevet i skjema for prosjektendring.

Vedtak
REK godkjenner prosjektet slik det nå foreligger, jfr. helseforskningsloven § 11, annet ledd.

Godkjenningen er gitt under forutsetning av at prosjektet gjennomføres slik det er beskrevet i søknad, endringssøknad, oppdatert protokoll og de bestemmelser som følger av helseforskningsloven med forskrifter.

Klageadgang
REKs vedtak kan påklages, jf. forvaltningslovens § 28 flg. Eventuell klage sendes til REK sør-øst B. Klagefristen er tre uker fra mottak av dette brevet. Dersom vedtaket opprettholdes av REK sør-øst B, sendes klagen videre til Den nasjonale forskningsetiske komité for medisin og helsefag for endelig vurdering, jf. forskningsetikkloven § 10 og helseforskningsloven § 10.

Vi ber om at alle henvendelser sendes inn på korrekt skjema via vår saksportal: http://helseforskning.etikkom.no. Dersom det ikke finnes passende skjema kan henvendelsen rettes på e-post til: post@helseforskning.etikkom.no.

Vennligst oppgi vårt referansenummer i korrespondansen.

Med vennlig hilsen

Knut Ruyter
Avdelingsdirektør
REK sør-øst sekretariatet

Ingrid Dønåsen
Rådgiver

Kopi til: *i.s.b.ribu@iln.uio.no; universitetsdirektor@uio.no*

Emne: Sv: Demens hos flerspråklige - språklig kompetanse, praksis og forvaltning
Fra: post@helseforskning.etikkom.no
Dato: 10.11.2014 14:25
Til: b.a.svendsen@iln.uio.no
Kopi:

Vår ref.nr.: 2014/1993 D Demens hos flerspråklige - språklig kompetanse, praksis og forvaltning

Viser til skjema for framleggingsvurdering mottatt 30.10.2014. Henvendelsen er vurdert av komiteens leder på fullmakt.

Formålet med prosjektet er å undersøke språklig kompetanse og språklig praksis hos flerspråklige som nylig har blitt diagnostisert med demens.

Deltakerne i studien skal rekrutteres i samarbeid med NAKMI, og utvalget vil bestå av fire personer med demensdiagnose og fire personer uten noen slik diagnose. Halvparten av deltakerne skal være flerspråklige.

Prosjektet søker å svare på hvordan demens manifesterer seg språklig hos enspråklige og flerspråklige personer, hva som kjennetegner samtaler der en av deltakerne har demens, og hvordan personer med demens og deres pårørende forholder seg til endringer som oppstår som følge av sykdommen, blant annet i møter med det offentlige, f. eks. helsevesenet

Basert på opplysningene som gis, oppfatter REK at formålet med prosjektet ikke er å gi ny kunnskap om sykdom og helse som sådan. Prosjektet faller dermed utenfor helseforskningsloven, som forutsetter at formålet med prosjektet er å skaffe ny kunnskap om helse og sykdom.

Prosjektet kan derfor gjennomføres uten godkjenning av REK. Det er forskningsansvarlig institusjons ansvar på å sørge for at prosjektet gjennomføres på en forsvarlig måte med hensyn til for eksempel regler for taushetsplikt og personvern.

Jeg gjør oppmerksom på at konklusjonen er å anse som veiledende jf. Forvaltningsloven §11. Dersom du likevel ønsker å søke REK vil søknaden bli behandlet i komitémøte, og det vil bli fattet et enkeltvedtak etter forvaltningsloven.

Med vennlig hilsen
Anne S. Kavli
Førstekonsulent
post@helseforskning.etikkom.no
T: 22845512

Regional komité for medisinsk og helsefaglig forskningsetikk REK sør-øst-Norge (REK sør-øst)
http://helseforskning.etikkom.no

REK
REGIONALE KOMITEER FOR MEDISINSK OG HELSEFAGLIG FORSKNINGSETIKK

Norsk samfunnsvitenskapelig datatjeneste AS
NORWEGIAN SOCIAL SCIENCE DATA SERVICES

Harald Hårfagres gate 29
N-5007 Bergen
Norway
Tel: +47-55 58 21 17
Fax: +47-55 58 96 50
nsd@nsd.uib.no
www.nsd.uib.no
Org.nr. 985 321 884

Bente Ailin Svendsen
Institutt for lingvistiske og nordiske studier Universitetet i Oslo
Postboks 1102 Blindern
0317 OSLO

Vår dato: 11.12.2014 Vår ref: 40523 / 3 / HIT Deres dato: Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 30.10.2014. All nødvendig informasjon om prosjektet forelå i sin helhet 08.12.2014. Meldingen gjelder prosjektet:

40523	*MultiLing Dementia - competence, practices, policies*
Behandlingsansvarlig	*Universitetet i Oslo, ved institusjonens øverste leder*
Daglig ansvarlig	*Bente Ailin Svendsen*

Personvernombudet har vurdert prosjektet, og finner at behandlingen av personopplysninger vil være regulert av § 7-27 i personopplysningsforskriften. Personvernombudet tilrår at prosjektet gjennomføres.

Personvernombudets tilråding forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, ombudets kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/meldeplikt/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://pvo.nsd.no/prosjekt.

Personvernombudet vil ved prosjektets avslutning, 31.12.2015, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Katrine Utaaker Segadal

Hildur Thorarensen

Kontaktperson: Hildur Thorarensen tlf: 55 58 26 54
Vedlegg: Prosjektvurdering

Personvernombudet for forskning

Prosjektvurdering - Kommentar

Prosjektnr: 40523

BAKGRUNN
Prosjektet er forhåndsvurdert av REK sør-øst (ref. 2014/1993), som oppfatter det slik at prosjektet faller utenfor helseforskningslovens virkeområde.

Prosjektet er en internasjonal samarbeidsstudie. Universitetet i Oslo er behandlingsansvarlig institusjon for den norske delen. Personvernombudet forutsetter at ansvaret for behandlingen av personopplysninger er avklart mellom institusjonene. Vi anbefaler at det inngås en avtale som omfatter ansvarsfordeling, ansvarsstruktur, hvem som initierer prosjektet, bruk av data og eventuelt eierskap.

MultiLing Dementia søker å forene psykolingvistiske og sosiolingvistiske tilnærminger til studiet av demens hos flerspråklige personer. Hovedfokuset ligger på tre temaer: 1: Flerspråklig kompetanse (hvordan manifesterer demens seg språklig?), 2:Flerspråklig praksis (Hvordan preges samtaler av at en deltaker har demens?), og 3: Forvaltning av flerspråklighet (Hvordan opplever og forholder personer med demens og deres pårørende seg til livsendringene som sykdommen medfører, i familien og i møter med helsevesenet?

UTVALG
Utvalget vil bestå av eldre personer med demens og deres pårørende. Ansatte vil også bli intervjuet om sin erfaring med gruppen som helhet, men ikke om enkeltpersoner.

INFORMASJON OG SAMTYKKE
Utvalget informeres skriftlig og muntlig om prosjektet og samtykker til deltakelse. Informasjonsskriv mottatt 08.12.2014 er godt utformet.

Det oppgis at enkelte av brukerne vil kunne ha noe redusert samtykkekompetanse. Personvernombudet finner at opplysninger innhentet fra personer uten full samtykkekompetanse, kan behandles med hjemmel i personopplysningsloven § 8 d) og § 9 h).

Det anses ikke som potensielt belastende for vedkommende å delta i prosjektet. Det opplyses at vedkommende vil bli gitt tilpasset informasjon, samt at hjelpeverge/nærmeste pårørende informeres om prosjektet, og eventuelt gir en uttalelse om hvorvidt opplysninger om vedkommende kan anvendes i studien.

Det vurderes at den valgte fremgangsmåten for inklusjon av personer uten full samtykkekompetanse, bidrar i betydelig grad til å redusere personvernulempen ved deltakelse. Det vurderes videre at opplysningene vil kunne komme gruppen som helhet til gode. På bakgrunn av dette finner personvernombudet at samfunnsinteressen i at behandlingen finner sted, overstiger ulempen den medfører for den enkelte registrerte.

INFORMASJONSSIKKERHET
Personvernombudet legger til grunn at forsker etterfølger Universitetet i Oslo sine interne rutiner for

datasikkerhet. Dersom personopplysninger skal lagres på mobile enheter, bør opplysningene krypteres tilstrekkelig.

PROSJEKTSLUTT
Forventet prosjektslutt er 31.12.2015. Ifølge prosjektmeldingen skal innsamlede opplysninger fra dem som har samtykket til dette da lagres i 20 år for oppfølgingsstudier. Data fra dem som ikke samtykker til lagring vil bli anonymisert.

Vi gjør oppmerksom på at all ny bruk av datamaterialet må meldes til personvernombudet.

Appendix C: Test protocols for the studies II and III

1. The short protocol for persons recruited via the memory clinic
2. The long protocol for all control participants

Date: Age: Tested by:		ID:
Semantic association test		
Pyramids and palm trees		
Free word association		
Picture naming	Nouns: Verbs:	
Verbal fluency	F: A: S: Animals:	
BAT cartoon description		
Digit span forwards	Longest: Score:	
Digit span backwards	Longest: Score:	
Digit span ordering	Longest: Score:	
Word-Picture matching	Nouns: Verbs:	
BNT short		
Stroop	Errors 1: corrected 1: time 1: Errors2: corrected 2: time 2:	
Sentence comprehension (visual world)	Active: Subject cleft: Object cleft:	

Date: Age: Tested by:		ID:
MMSE		
CERAD 10-words	1: 2: 3:	
Semantic association test		
Recall CERAD 10-ord	A: B:	
Pyramids and palm trees		
Trail making	A: B:	
Free word association		
Picture naming	Nouns: Verbs:	
Verbal fluency	F: A: S: Animals:	
BAT cartoon description		
CERAD figure copying	1: 2: 3: 4:	
Digit span forwards	Longest: Score:	
Digit span backwards	Longest: Score:	
Digit span ordering	Longest: Score:	
Recall CERAD figures	1: 2: 3: 4:	
Word-Picture matching	Nouns: Verbs:	
BNT short		
Stroop	Errors 1: corrected 1: time 1: Errors2: corrected 2: time 2:	
Sentence comprehension (visual world)	Active: Subject cleft: Object cleft:	

Appendix D: Items for the picture-naming and word-to-picture matching tests in studies II and III

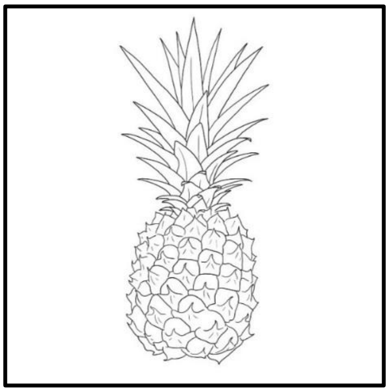
1. The final list of nouns for the Picture Naming task and the word-to-picture matching task
2. The final list of verbs for the Picture Naming task and the word-to-picture matching task
3. Example images from the Picture Naming task
4. Example images from the word-to-picture matching task

Norwegian	English (translation equivalent)
Et belte	A belt
En sekk	A backpack
En kenguru	A kangaroo
Et tog	A train
En fiolin	A violin
En løk	An onion
En pingvin	A pinguin
Et termometer	A thermometer
Ei fjær	A feather
Ei vekt	A (pair of) scales
En elg	A moose
Ei tønne	A barrel
En radio	A radio
Et skjerf	A scarf
Ei flue	A fly
En linjal	A ruler
En dress	A suit
Ei skilpadde	A turtle
Ei skjorte	A shirt
En sopp	A mushroom
En gitar	A guitar
En hval	A whale
En sitron	A lemon
En hanske	A glove
En komfyr	A stove
En sjiraff	A giraffe
En ananas	A pineapple
En maur	An ant
En hatt	A hat
En traktor	A tractor

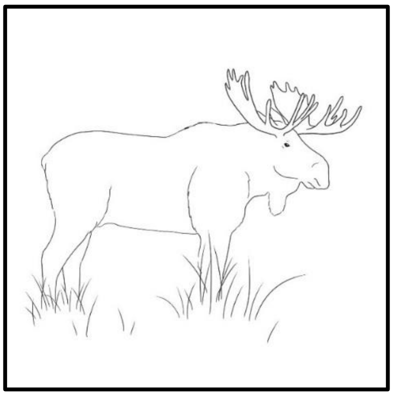
Norwegian	English (translation equivalent)
Å stryke	To iron
Å brenne	To burn
Å barbere	To shave
Å fekte	To fence
Å skrive	To write
Å skyte	To shoot
Å plystre	To whistle
Å spikke	To chisel
Å lese	To read
Å blåse	To blow
Å hugge	To chop
Å ro	To row
Å bore	To drill
Å ake	To sled
Å mure	To lay bricks
Å krabbe	To crawl
Å bokse	To box
Å danse	To dance
Å kjevle	To roll (dough)
Å dryppe	To drip
Å vaske	To clean
Å kjøre	To drive
Å smelte	To melt
Å male	To paint
Å brekke	To break / to snap
Å seile	To sail
Å dusje	To shower
Å løpe	To run
Å ri	To ride (horseback)
Å fryse	To freeze

Picture-naming: Nouns

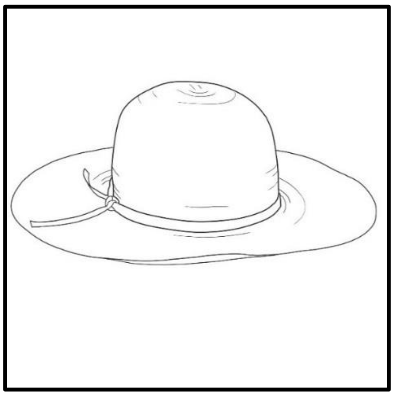
Ei fjær
(a feather)



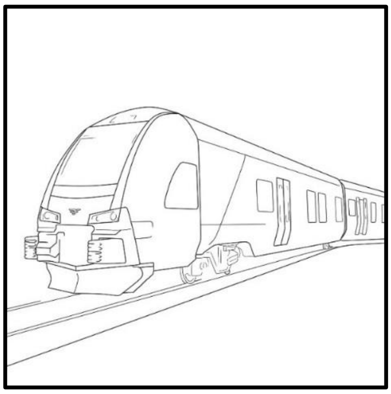
En annanas
(a pineapple)



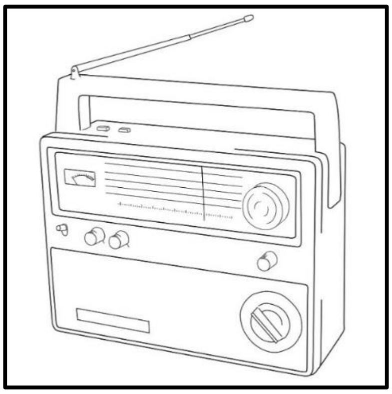
En elg
(a moose)



En hatt
(a hat)



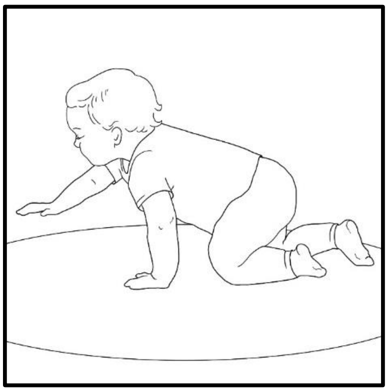
Et tog
(a train)



En radio
(a radio)

Picture-naming: Verbs

Å danse
(to dance)

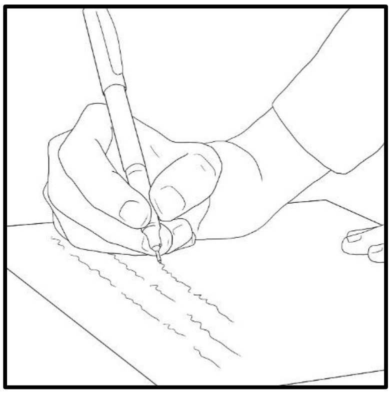


Å krabbe
(to crawl)

Å fryse
(to freeze)

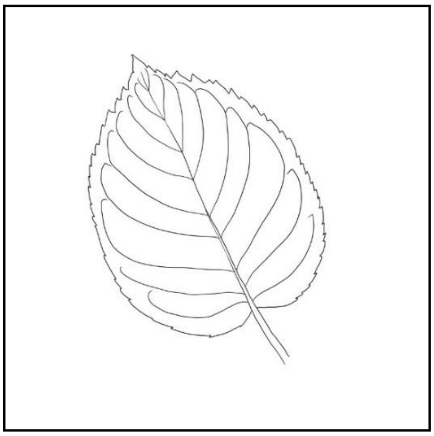
Å hugge
(to chop)

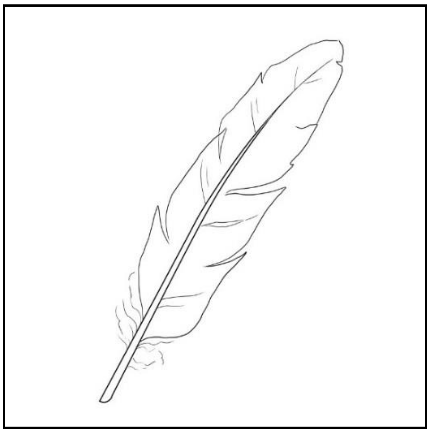
Å løpe
(to run)

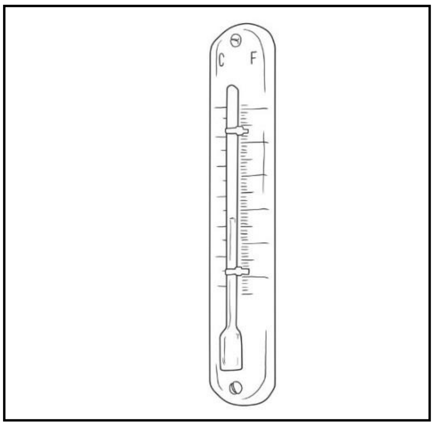


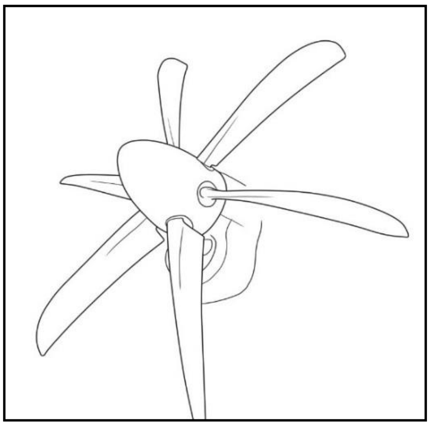
Å skrive
(to write)

Hvor er det en fjær?
Where is there a feather?

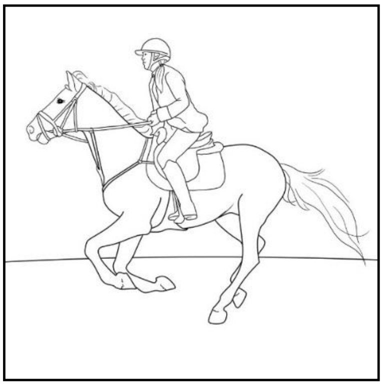


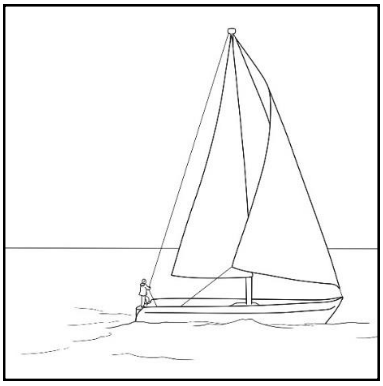






Hvor er det en som rir?
Where do you see someone riding?





Appendix E: Items on the free word association task in study II

Please say the first word that comes to mind when you hear the words I read aloud.

Please try to answer with only <u>one</u> word per response

Remember that there are no right or wrong responses.

sol – *sun* (N)

gå – *walk* (V)

hverdag – everyday (N)

forventning - expectation/anticipation(N)

ansette- employ (V)

servere- serve/wait/deliver (V)

håndtere- handle/ manage (V)

anledning – occasion/opportunity(N)

religion- religion (N)

sykehus – hospital (N)

virksomhet-firm/operation/activity(N)

konsert – concert (N)

innføre- introduce/ import (V)

tolke- interpret (V)

industri- *industry* (N)

gidde- be bothered (V)

nyte – enjoy (V)

konsentrere – concentrate (V)

understreke –emphasize/underline (V)

transport- transport (N)

avis- newspaper (N)

motstander- opponent (N)

litteratur- litterature (N)

drømme- dream (V)

overføre – transfer/transmit (V)

struktur- structure (N)

landbruk- agriculture/ farming (N)

middel – means/ resource (N)

økning- increase (N)

vindu- window (N)

informere – inform (V)

variant – variant/ version (N)

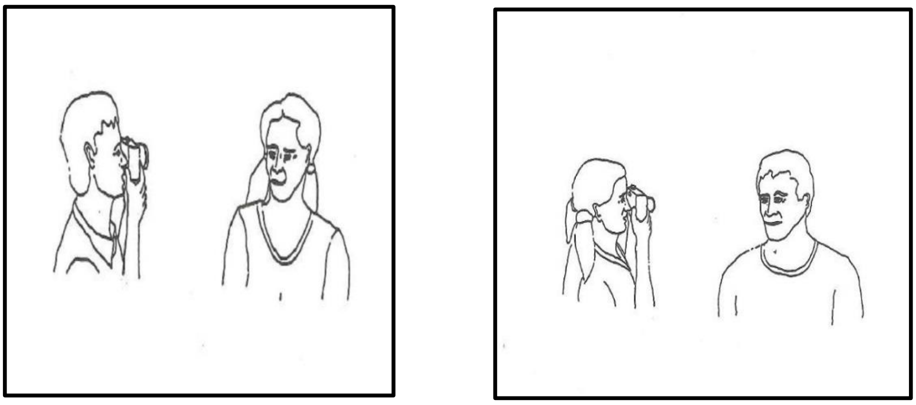
Appendix F: Items on the sentence-to-picture matching task in study III

Test set 1	
Norwegian	**English**
Jenta fotograferer gutten	The girl photographs the boy
Det er barnet damen dytter	It is the child the woman pushes
Det er kua som biter hesten	It is the cow that bites the horse
Jenta bærer mannen	The girl carries that man
Det er damen som truer mannen	It is the woman who threatens the man
Det er gutten jenta kveler	It is the boy the girl strangles
Jenta dytter gutten	The girl pushes the boy
Det er kua hesten biter	It is the cow the horse bites
Det er damen som vasker barnet	It is the woman who washes the child
Mannen klyper damen	The man pinches the woman
Det er hesten som sparker kua	It is the horse that kicks the cow
Det er jenta som filmer gutten	It is the girl who films the boy
Jenta kysser mannen	The girl kisses the man
Det er mannen damen redder	It is the man the woman rescues
Det er gutten jenta fotograferer	It is the boy the girl photographs
Mannen slår damen	The man hits the woman
Det er gutten jenta maler	It is the boy the girl paints
Det er jenta som klorer gutten	It is the girl who scratches the boy

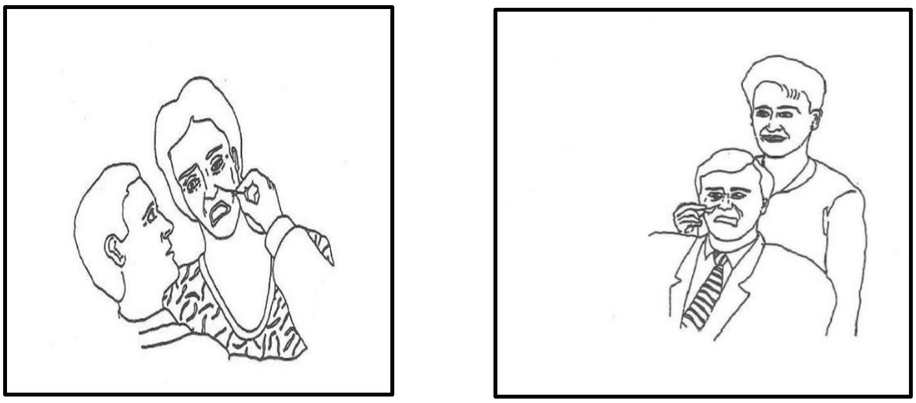
Test set 2	
Norwegian	**English**
Mannen redder damen	the man rescues the woman
Det er barnet damen dytter	It is the child the woman pushes
Det er mannen som bærer damen	It is the man who carries the woman
Mannen truer damen	The man threatens the woman
Det er gutten som kveler jenta	It is the boy who strangles the girl
det er barnet damen vasker	It is the child the woman washes
Gutten klorer jenta	The boy scratches the girl
Det er hesten kua sparker	It is the horse the cow kicks
Det er barnet moren slår	It is the child the mother hits
Hesten biter kua	The horse bites the cow
Det er jenta som fotograferer gutten	It is the girl who photographs the boy
Det er jenta som kysser mannen	It is the girl who kisses the man
Gutten maler jenta	The boy paints the girl
Det er jenta som klyper gutten	It is the girl who pinches the boy
Det er damen mannen filmer	It is the woman the man films

Test set 3	
Norwegian	**English**
Damen vasker barnet	The woman washes the child
Det er jenta gutten klyper	It is the girl the boy pinches
Det er mannen barnet kysser	It is the man the child kisses
Damen kveler mannen	The woman strangles the man
Det er mannen som redder damen	It is the man who saves the woman
Det er damen mannen truer	It is the woman the man threatens
Jenta fotograferer gutten	The girl photographs the boy
Det er gutten som dytter jenta	It is the boy who pushes the girl
Det er gutten som maler jenta	It is the boy who paints the girl
Kua sparker hesten	The cow kicks the horse
Det er barnet mannen klorer	It is the child the man scratches
Det er damen som slår mannen	It is the woman who hits the man
Mannen filmer damen	The man films the woman
Det er jenta mannen bærer	It is the girl the man carries
Det er jenta som biter gutten	It is the girl who bites the man

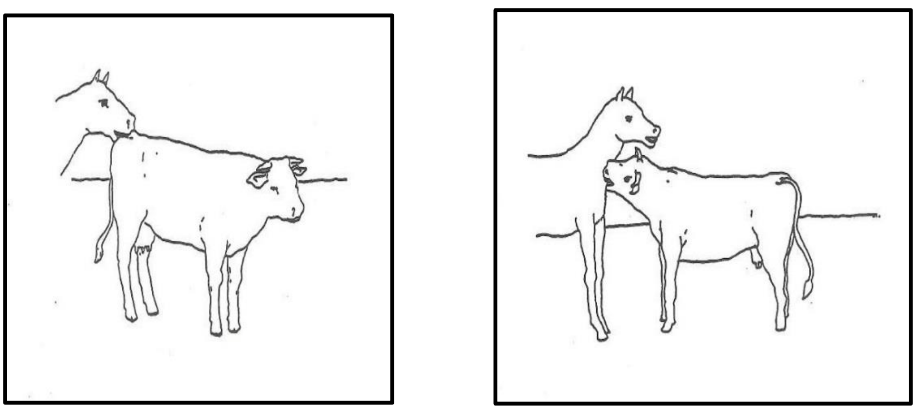
ACTIVE



The girl photographs the boy



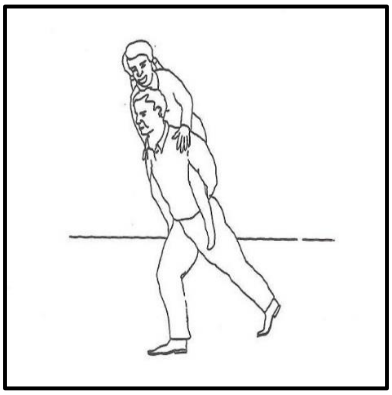
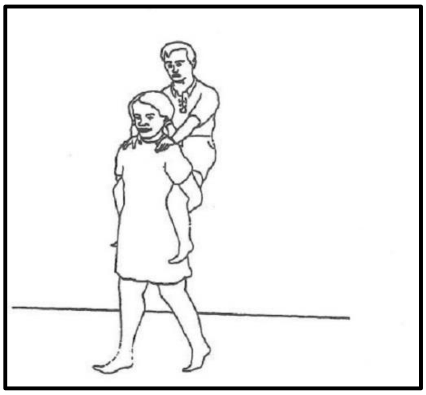
The man pinches the woman



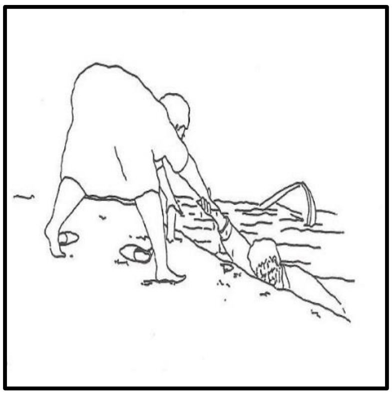
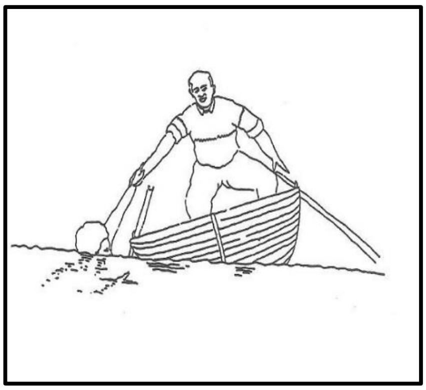
The cow bites the horse

SUBJECT CLEFT


It is the woman who pushes the child

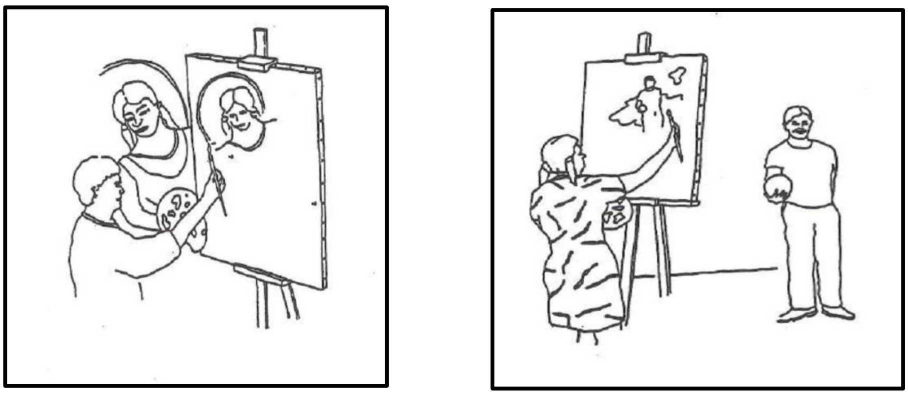



It is the girl who carries the man

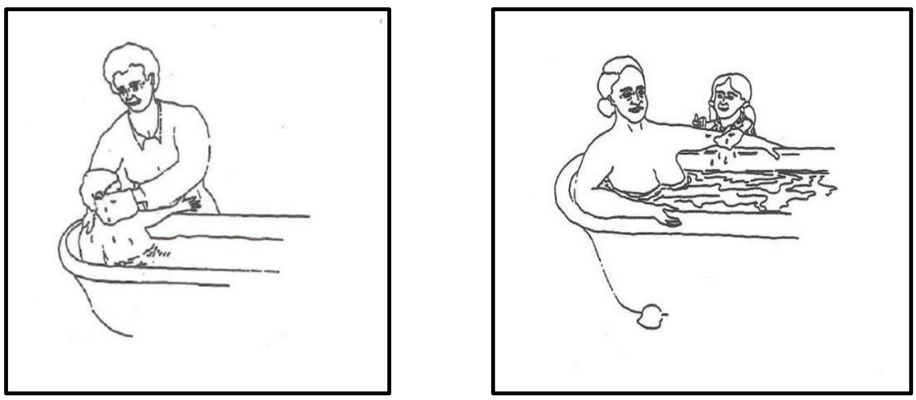



It is the woman who saves the man

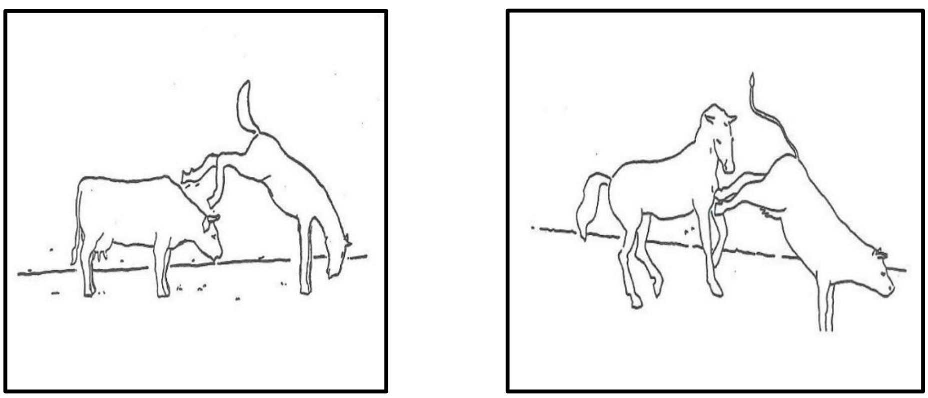
OBJECT CLEFT



It is the boy the girl paints



It is the child the woman washes



It is the horse the cow kicks

Appendix G: Background questionnaire for studies II and III

Questionnaire linguistic and social background

ID:_______________

(Added by test leader)

Date:______________

1. Gender: Are you **male** or **female**? (encircle)

2. Date of birth:

3. Are you **right handed** or **left handed**? (encircle)

4. What is you highest completed education?
    a. Elementary school (7 years)
    b. Middle school (10 years)
    c. High school (12 or 13 years)
    d. Vocational high school (3 or 4 years after middle school)
    e. Undergraduate
    f. Postgraduate
    g. PhD
    h. Total years of schooling______

5. What is your current profession (last profession, if retired)

6. Do you have normal hearning?
    Yes
    No
    a. If no, do you use a hearing aid?
    Yes
    No
    b. If yes, is your hearing normal with hearing aid?
    Yes
    Nno

7. Is your vision normal?
    Yes
    No
    a. If no, do you wear glasses or contact lenses?
    Yes
    No

b. If yes, is your vision corrected to normal with glasses/contact lenses?
   Yes
   No

8. Are you color blind?
   Yes
   No

9. Do you have difficulties with understanding and processing numbers or mathematics?
   Yes
   No

10. Do you have difficulties with reading/writing?
   Yes
   No

11. Have you ever had a brain injury?
   Yes
   No
   a. If yes, what kind of injury?

12. Do you have any known neurological diseases (epilepsi, MS etc.)?
   Yes
   No
   a. If yes, what kind?

13. Do you currently use psychofarmace?
   Yes
   No

14. What is your mothertongue?
   a. Norwegian
   b. Other

   c. If other, which?

15. In which country did you grow up?
   a. Norway
   b. Other

   c. If other, which?
   d. If other, how old were you when you moved to Norway?

16. Have you ever lived in another country than Norway for a period of six months or more?

a. Yes
b. No

c. If yes, which country, for how long and when was this?

17. Which language did you speak in your household when you grew up?
    a. Norwegian
    b. Other

    c. If other, which language and to whom did you speak this?

18. Have you learned any foreign languages in school?
    Yes
    No
    a. If yes, which one(s)?

19. Have you ever learned a foreign language outside of school (i.e., by living in a country where the language is spoken, self studies, language games or programs)
    Yes
    No
    a. If yes, which one(s) and in what manner?

20. Do you ever **speak** other languages than Norwegian in your daily life?
    Yes
    No
    a. If yes, how often?
    Every day
    A couple of times a week
    A couple of times a month
    less

21. Do you ever **read** other languages than Norwegian in your daily life?
    Yes
    No
    a. If yes, how often?
    Every day
    A couple of times a week
    A couple of times a month
    less

22. Do you ever **write** in other languages than Norwegian in your daily life?
    Yes
    No

a. If yes, how often?
   Every day
   A couple of times a week
   A couple of times a month
   less

23. Do you ever **hear** in other languages than Norwegian in your daily life?
    Yes
    No
    a. If yes, how often?
       Every day
       A couple of times a week
       A couple of times a month
       less

24. Range your skills in Norwegian, on a scale from 0-10 for the following activities (tick the box you feel corresponds to your skill level)

	1	2	3	4	5	6	7	8	9	10
Speak										
Understand										
Read										
Write										

25. Compared to a native speaker, range your skills in your foreign languages on a scale from 0-10:

Language:

	1	2	3	4	5	6	7	8	9	10
Speak										
Understand										
Read										
Write										

Language:

	1	2	3	4	5	6	7	8	9	10
Speak										
Understand										
Read										
Write										

Language:

	1	2	3	4	5	6	7	8	9	10
Speak										
Understand										

Read										
Write										

Appendix H: Items on the picture-naming task in study III

Norwegian	**English (translation equivalent)**
En gris	A pig
Ei dør	A door
En sopp	A mushroom
Ei pil	A dart
Ei geit	A goat
Et hjul	A wheel
En katt	A cat
En blyant	A pencil
Et belte	A belt
Ei tunge	A tounge
Et eple	An apple
En kylling	A chick(en)
En buss	A bus
Et hjerte	A haert
Et tog	A train
Et skjell	A shell
Ei bok	A book
Et brød	A bread
En hval	A whale
En fisk	A fish
Ei flaske	A bottle
En elefant	An elephant
Ei saks	A pair of scissors
Ei tavle	A back board
En vante	A mitten
En kikkert	Binoculars
Ei høne	A hen/chicken
En kjole	A dress
En skilpadde	A turtle
En banan	A banana

Norwegian	English (translation equivalent)
Å klippe	To cut
Å sove	To sleep
Å sitte	To sit
Å huske	To swing
Å stupe	To dive
Å sage	To saw
Å (spille) fløyte	To (play the) flute
Å flette	To braid
Å sykle	To ride a bike
Å filme	To film
Å melke	To milk
Å drikke	To drink
Å plukke (blomster)	To pick (flowers)
Å massere	To give a massage
Å støvsuge	To vacuum clean
Å fiske	To fish
Å danse	To dance
Å file (negler)	To file (nails)
Å bade	To bathe
Å fryse	To freeze
Å kjøre	To drive
Å bokse	To box
Å poste	To mail
Å hoppe	To jump
Å skrive	To write
Å fekte	To fence
Å kjevle	To roll (dough)
Å sy	To sow
Å stryke	To iron
Å trylle	To do magic