The linguistic deficit in patients with Alzheimer's Disease: is there a syntactic impairment?

Inauguraldissertation
zur Erlangung des Grades eines Doktors der Philosophie
im Fachbereich Neuere Philologien
der Johann Wolfgang Goethe-Universität
zu Frankfurt am Main

vorgelegt von

Irene Caloi

aus: Noventa Vicentina – Vicenza (Italien)

2017

1. Gutachterin: Prof. Dr. Cecilia Poletto

2. Gutachterin: Prof. Dr. Adriana Belletti

Tag der mündlichen Prüfung: 16. November 2016

Contents

ABBREVIATIONS

+NP	The lexical restriction feature
+Q	The interrogative feature
+Rel	The relative head feature
+Wh	The interrogative head feature
A'-	A-bar or non-argumental
AD	Alzheimer's Disease
C	Complementizer
ClOR	Object relative with resumptive clitic
CO	Controls
CP	Complementizer phrase
DP	Determiner phrase
GG	Grammatical gender
IP	Inflectional phrase
LoE	Level of Education
MMSE	Mini Mental State Examination
NG	Natural gender
NP	Noun phrase
OR	Object relative
ORdem	Object relative with a demonstrative in the head
PAD	Patients of Alzheimer's disease
POR	Passive object relative
PP	Prepositional phrase
RC	Relative clause
RM	Relativized Minimality
SR	Subject relative
SVO	Subject-Verb-Object simple declarative sentence
WhichS	Which Subject question
WhoO	Who Object question
WhoS	Who Subject question

1 INTRODUCTION

1.0 Introduction

This study aims at enlarging the comprehension of the linguistic deficit in patients affected by Alzheimer's disease (PAD); in particular, it deals with their syntactic competence.

The interest in Alzheimer's disease (AD) is due to the relevance of the disease in contemporary society. According to the *World Alzheimer Report 2016* (Prince et al., 2016), AD and similar forms of dementia affect more than 47 millions of people in the world, a number that exceeds the inhabitants of Spain. Thanks to improved living conditions, the world population has grown considerably in the last decades, and life expectancy is increasing in many countries. The combination of these phenomena determines the fast aging of the population. According to the *Report*, the number of people affected by dementia will triplicate in the next three decades and will reach 131 millions by 2050.

The social impact of the disease is very high because patients need constant assistance in their daily life. From the economic point of view, the *Report* estimates that AD will soon cost almost one trillion US Dollars per year, an amount of money higher than the value of companies like Google and Apple. The majority of these resources is destined to daily-care. The outlook is particularly worrying because there is currently no treatment to the disease: after onset, AD never goes into remission.

One problematic aspect of Alzheimer's disease is its diagnosis: it mainly relies on the assessment of different abilities and domains. Language crucially is one of those because many patients show a linguistic deficit among the first and most evident symptoms of the disease. Therefore, a deep understanding of the linguistic impairment is necessary in order to improve the diagnostic techniques.

Moreover, as we lack drugs to treat the disease, much attention is devoted nowadays to the development of training programs for increasing the cognitive resilience to the disease. Exercises involving linguistic abilities are often part of those activities.

Unfortunately, not all aspects of the linguistic impairment in PADs are clear yet. For instance, we know much about anomia, while little has been observed with respect to sentence processing. Therefore, an in-depth analysis of the syntactic competence in PADs is needed in order to complete our understanding of their linguistic deficit. In the future, this kind of research might found a practical application in the design of diagnostic tools, as well as in the creation of training programs.

1.1 Research questions

The present study focuses on Italian-speaking PADs. In order to improve the understanding of their syntactic competence, I address the following research questions:

1) Do PADs suffer from syntactic impairment?

2) How can the impairment in PADs be accounted for?

3) At which stage of the disease are PADs affected by syntactic impairment?

Question (1) is meant to take into consideration the syntactic competence of Italian-speaking PADs from two different perspectives. On one hand, I intend to consider the syntactic information stored in the lexicon as part of the lexical entry. For this purpose, PADs complete a grammatical gender (GG) retrieval task on a list of 100 Italian nouns. The issue is interesting because it investigates an aspect of the lexicon of PADs that has not received much attention so far. While the semantic disruption behind anomia has been extensively studied (see Chapter 2), little is known about patients' ability to retrieve the GG of lexical items. The task will be helpful in order to determine whether the syntactic information tends to be spared or impaired by the linguistic deficit.

On the other hand, the question deals with syntax intended as the capacity to complete the processing of syntactic structures in sentence comprehension and production. The study focuses on sentence comprehension and includes two sentence-to-picture matching tasks: the first one focuses on Wh-questions, the second one on relative clauses (RC). The two sentence structures have been selected for the study because of the many syntactic manipulations they allow in their derivations. The possibility to compare

Wh-questions and RCs characterized by slightly different syntactic characteristics (i.e., extraction site, Wh-element in use, voice, etc.) offers the opportunity to draw comparisons, and therefore to gain a closer insight into the syntactic competence of PADs.

In case the first question receives a positive answer (i.e., signs of impairment are detected in PADs either in the retrieval of syntactic information or in the form of sentence comprehension difficulties), the following step will be the analysis of factors that influence the participants' performance. The observation of how patients perform in correspondence to different experimental conditions should unveil the characteristics of the impairment and shed light on how this can be accounted for. This is precisely the aim of Question (2).

Question (3) concerns the level of dementia at which the syntactic impairment emerges. Being AD a degenerative disease, not all symptoms appear simultaneously. Within a general progressive impairment in all relevant domains, some symptoms are more evident in the first phases of the disease, while others emerge later. With respect to this, it is crucial to observe the evolution of the linguistic deficit across the different stages of the disease. This is of primary relevance because the linguistic impairment is one of the factors in use for the diagnosis. For this purpose, PADs characterized by different levels of dementia are enrolled in the study and their performance patterns are compared.

In addition, each of the three experimental tasks in use is designed in order to answer more fine-grained research questions. These will be addressed at the beginning of Chapters 4, 6 and 7, in which the three experimental tasks are presented.

Finally, task designs and collected data will also offer the opportunity to make some considerations on theoretical aspects of the linguistic phenomena under analysis. In particular, the study will deal with the process of item retrieval from the mental lexicon and with the computation of different kinds of feature configurations in the derivation of Wh-questions and relative clauses.

1.2 Overview of the study

After a brief introduction to the research questions and the content of the present study, Chapter 2 opens with primary information on the neurodegenerative nature of Alzheimer's disease and its consequences. From the neuropsychological point of view, PADs suffer from memory loss, executive function impairment, aphasia, and behavioural

and psychiatric disorders (see Spinnler 1996, for an overview). Language is affected mainly at the lexical level. Through a long and fruitful debate, authors agreed in ascribing anomia mainly to a semantic loss process and, to a lower extent, to retrieval difficulties (Chertkow et al., 1989, Chertkow & Bub, 1990; Hodges et al., 1996; Cuetos et al., 2015, a.m.o.). Anomia also affects irregular verbal morphology: PADs are impaired at retrieving verbal forms characterized by stem changes, while the use of regular morphology is spared (Colombo et al., 2009; Walenski et al., 2009). This dissociation between regular and irregular forms supports Ullman's (2001) hypothesis of a dissociation between the declarative (mental lexicon) and the procedural (mental grammar) components of the language.

Syntax has been considered spared in PADs for a long time. Indeed, PADs are usually sensitive to syntactic violations and perform well on grammaticality judgement tasks (Kavè & Levy, 2003a; Kempler et al., 1998; MacDonald et al., 2001; Price & Grossman, 2005). However, other studies suggest that PADs might have difficulties in the computation of syntactically complex sentences (Bickel et al., 2000; Kempler et al., 1998; Small, Kemper & Lyons, 2000; Waters, Rochon & Caplan, 1998). The number of studies that address the issue is very narrow though. In addition, experimental designs lack clear theoretical frameworks for the analysis of the experimental conditions they test. This, combined with the presence of confounding factors (i.e., lack of minimal pairs, mix of animate and inanimate arguments, etc.), prevent authors from drawing clear conclusions, except for an observed asymmetry between simple declarative sentences and relative clauses. In the last part of the chapter, I will point out which aspects of syntactic competence deserve further investigation.

Each experimental study is preceded by an introduction to the theoretical framework in use and an overview of the main results from similar studies in different empirical fields (i.e., aphasics, children, etc.). The adoption of a theoretical framework allows for the design of precise research questions and experimental conditions. The overview of previous studies allows for comparisons across languages and across empirical fields. Both components enrich the interpretation of the collected data.

For this purpose, Chapter 3 deals with the language-specific characteristics of grammatical gender in lexical entries (Levelt, 1989 and much subsequent work) and in the process of GG retrieval (Longobardi, 1994; Friedmann & Biran, 2003). I will also

review studies that show how vulnerable GG is in acquired linguistic disorders (Badecker et al., 1995; Luzzatti & De Bleser, 1996). The overview will help to point out which aspects are interesting for the present research.

In Chapter 4, the competence of PADs is tested with respect to GG retrieval on transparent, opaque and irregular nouns. The task takes into consideration two further factors. First, the comparison between simple and derived nouns allows for verifying how PADs deal with derivational morphology. Second, the use of experimental nouns with natural biological gender (e.g., *madre*, 'mother') and inanimate nouns without natural gender (e.g., *cuore*, 'heart') can reveal whether PADs resort to semantic information for the retrieval of GG in opaque nouns (i.e., nouns characterized by the final marker *–e*, which is in use both for masculine and feminine nouns).

In Chapter 5, the attention shifts to the computation of sentences characterized by Wh-movement. The chapter presents an overview of the analyses put forth for Wh-questions and Relative Clauses within the framework of Generative Grammar. As for Wh-questions, I assume that the Wh-operator is extracted from its argument position and moved to the CP layer. There, it must enter into a Spec-head relation with the verb (which is also raised to CP), in order to satisfy the Wh-Critrion (Rizzi, 1996). Similarly, Relative Clauses are derived through the extraction of an argument and movement towards the CP-layer of the RC (for a recent account see Cinque, 2008, 2014).

For both kinds of sentences, the discussion also outlines the relevant factors for the evaluation of computational abilities. For instance, the overview of Italian studies on adults (De Vincenzi, 1991, 1996), aphasic speakers (Garraffa & Grillo, 2008) and language acquisition (De Vincenzi et al., 1996; Guasti et al., 2012) reveals that the relevant factors to take into consideration for Wh-question computation are the extraction site (subject vs object Wh-questions) and the kind of Wh-element in use (Who vs WhichNP questions). Similarly, studies on RCs unveil a sharp asymmetry between subject and object RCs both in adult speakers (De Vincenzi 1991; Gordon, Hendrick & Johnson, 2001; Traxler, Morris & Seely, 2002 among others), in aphasia (Caramazza & Zurif, 1976; Friedmann, 2008; Garraffa & Grillo, 2008; Grillo, 2008) and in language acquisition (Contemori & Belletti, 2014; Friedmann, Belletti & Rizzi, 2009; Kidd, Brandt, Lieven & Tomasello, 2007, among many others). What is particularly interesting in RCs then is the observation of the syntactic factors that can improve the computation

of object relatives. So far, background literature has mainly taken into consideration number and gender mismatch, lexical restriction, animacy, passive voice and resumption (Adani et al., 2010; Contemori & Belletti, 2014; Friedmann, Belletti & Rizzi, 2009, among many others). Overall, Chapter 5 sets the guidelines for the design of the experimental tasks on Wh-questions and RCs and for the interpretation of the results.

The study on Wh-question comprehension in PADs is presented in Chapter 6. PADs undergo a sentence-to-picture matching task with four experimental conditions obtained through the crossing of the two factors mentioned above (i.e., extraction site and Wh-element in use).

Chapter 7 presents results from the sentence-to-picture matching task on relative clauses. PADs listen to simple declarative sentences, subject relatives and object relatives. The comparison among the three conditions is useful in order to verify whether PADs can deal with relative clauses (simple declaratives vs relative clauses), and whether the asymmetry between subject and object relatives increases with respect to the performance of healthy controls. The task includes three more conditions for the evaluation of factors that can reduce the complexity of ORs. For instance, I will test whether PADs benefit from passive voice, object clitic resumption and mismatch in lexical restriction for the comprehension of object relatives.

Both for Wh-questions and RCs, results will be interpreted along the lines of the account based on Relativized Minimality (Rizzi, 1990). In the vein of Friedmann, Belletti & Rizzi (2009) and much subsequent work, I will claim that PADs are impaired at computing crossing movements between arguments in the relative clauses. In particular, the success of the derivation depends on the feature arrays of the two arguments involved. PADs are relatively spared at computing configurations that entail relevant mismatches in features, while they are impaired at computing sentences with argument features in a configuration of inclusion.

Results from the three experimental tasks will be further summarized and discussed in Chapter 8 in order to answer the research questions addressed above.

2 THE LINGUISTIC DEFICIT IN PATIENTS WITH ALZHEIMER'S DISEASE

2.0 Introduction

The aim of the present chapter is to illustrate the main characteristics of dementia of the Alzheimer's type. The first section (2.1) presents an introduction to the pathology and its symptoms in order to outline the framework in which the linguistic impairment develops. The overview does not examine the medical causes and the neuropsychological symptoms of the disease in depth, as that would exceed the aims of a linguistic study; however, a compendium of the main characteristics of this form of dementia (with a specific focus on what is relevant for this work) is provided.

The greater part of the chapter is devoted to an analysis of the background literature on the linguistic deficit associated to Alzheimer's disease. The most evident and the most studied linguistic disorder in AD is anomia. Many studies have addressed the issue in the last four decades. I will summarize their main outcomes in section 2.3. A much narrower number of studies focused on the status of morphology: their (consistent) results are presented in section 2.4. Finally, section 2.5 examines studies that have previously addressed the issue of a syntactic deficit in PADs.

The overview is also instrumental in illustrating the reasons why further research on the syntactic deficit of PADs is needed. I will point out two interesting aspects of the linguistic research in AD that have already been addressed, but only partially. The first one concerns the status of syntactic information within lexical entries. The second one deals with the processing of sentences characterized by extraction movements.

2.1 Alzheimer's disease

Alzheimer's disease is a degenerative illness that affects the central nervous system. It is progressive and it cannot recede. It causes an increasing impairment of cognitive abilities that compromises daily life activities. Patients gradually lose autonomy in their everyday

life. The disease is age-related (Braak & Braak, 1997) as the risk for its onset sharply increases with age.

Alois Alzheimer (1964-1915) first describes the disease in 1906: Auguste D., a 51-year old female patient, shows symptoms of cognitive impairment. First, her memory is disturbed; second, she has unexpected behaviours: e.g., she cannot recognize and use common objects; she is confused and does not know where she is and what day it is; she moves things around with no reasons, among other symptoms. After her death, Alzheimer performs an autopsy and observes remarkable changes in the brain, mainly in the form of atrophy.

After this first case, many more were observed. In 2006, on the occasion of the 100th anniversary of the first description of the disease, Braak & Del Tredici (2006) and Hyam (2006) describe the brain of patients affected by AD as showing signs of neural loss, tangles and plaques. Their formation is related to the presence of two proteins in the brain: the Tau and the Beta amyloid proteins. PADs usually produce abnormal amounts of Tau. The protein is present is the microtubules that transport substances within the brain cells. When Tau is overproduced, microtubules collapse, twist and form tangles within the cells. In turn, tangles disrupt intercellular functions. Cells characterized by tangles can survive for a long time (probably decades), but their death is nonetheless premature. When nerve cells die, tangles become extraneuronal and remain in the brain tissue.

Beta amyloids are protein fragments that surround nerve cells. Their function in the brain is not clear yet, but the human metabolism is usually able to break down fragments and eliminate them from the brain. This does not happen in PADs and beta amyloids tend to cluster and form plaques. These cause focal alterations in the synaptic organization.

In addition, neural loss increases to the point that PADs maintain only circa 50% of the neurons of a healthy brain (Gómez-Isla et al., 1997). The pace of the changes described above is very slow. According to Braak & Braak (1997), a small percentage of people (< 20%) develops the first neurofibrillary tangles already in their twenties, although the disease only very rarely manifests itself before the age of 55. Dementia emerges only after the brain has undergone sizeable changes due to neural loss, tangles and plaques (Hyman, 2006).

The distribution of lesions is rather consistent across patients (Braak & Del Tredici, 2006). Damages usually start from the transentorhonal region and subsequently spread to the entorhinal region and the hippocampus. This configuration corresponds to a pre-clinical period of AD, when changes start taking place in the brain, but patients do not show any symptoms (this stage can last decades). Later, damages harm the neo-cortex of the occipito-temporal lobes and eventually reach the frontal area. At this point, PADs start showing signs of mild cognitive impairment (Braak & Del Tredici, 2006). Ultimately, higher associative areas are also affected and dementia becomes evident through a variety of symptoms.

Spinnler (1996) provides us with an overview of the symptoms of AD at the neuropsychological level. Indeed, PADs can develop forms of impairment in different areas. The first symptom to appear is usually an impairment at the anterograde episodic memory: PADs cannot acquire new memories anymore; for instance, they do not remember new events of their life (e.g., a new person they meet, a place they visit, what they eat, etc.). The impairment usually affects prospective memory too: PADs forget to perform actions or to take part in events scheduled at a precise moment in time (e.g., they forget about taking pills, deadlines, appointments, etc.). Episodic memory related to public events and famous faces also fades away. Progressively, semantic memory is also affected, while procedural memory seems to be spared longer (but not completely). Autobiographical memory undergoes major disruptions too. In particular, PADs forget memories from more recent years and tend to transpose old memories to the present.

The supervisory attentional system is overall impaired, with negative effects on a variety of cognitive processes, such as: attentional control, inhibitions control, working memory, reasoning, etc.. The pool of executive resources that PADs can count on gradually decreases, such that task execution is quantitatively and qualitatively reduced too.

Also agnosia and apraxia are often associated to the disease. The former consists of the inability of recognizing familiar objects; the latter corresponds to difficulties at planning and carrying out actions and tasks.

All these forms of impairment deprive PADs of their autonomy in every-day life. For instance, patients become unable to prepare meals, choose appropriate clothes for the

weather, find their way in areas they used to be familiar with, take care of personal hygiene, and much more.

Spinnler (1996) also describes symptoms of the psychiatric kind: emotional distress, depression, mania and hallucinations. Behavioural disorders are often caused by the inability of patients to recognize and accept the circumstances and the location where they are. For instance, anxiety may derive by the fact of not recognizing their own house, which makes them feel in danger and eager to go home (even though they already are in an environment they should be familiar with).

Finally, PADs also suffer from language impairment. The rest of the present chapter deals with it.

But before moving forward, a word of caution is necessary. Alzheimer's disease can be diagnosed with absolute certainty only through a post-mortem inspection of brain tissues. While alive, patients receive a diagnosis of Probable Alzheimer's Disease based on three elements. First, the patient must show signs of impairment in at least two of the following domains: memory, executive function, visuospatial performance, language, and behaviour (Carrillo et al., 2013). Second, specialists collect data through clinical tests. This might include neuroimaging and blood tests. Neuroimaging techniques allow the monitoring of brain changes (i.e., lesions, atrophy). Blood tests allow the individuation of biomarkers for the presence of Beta amyloid and Tau proteins. Third, patients' medical history and further clinical exams can be used to exclude other causes of dementia (e.g., alcohol abuse, previous head injuries, etc.). These criteria were first established by the National Institute of Neurologic and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA) in the Eighties, and have more recently been revised by the National Institute on Aging and the Alzheimer's Association (Carrillo et al., 2013).

For the sake of simplicity and uniformity to previous studies, the present work uses the expression Patients of Alzheimer's Disease (PAD) for persons who have received a diagnosis based on the criteria listed above. Nonetheless, it should be noted that a *post-mortem* diagnosis is not available for any of the enrolled patients in the studies presented in Chapters 4, 6 and 7. Unless specified differently, the same *modus operandi* is valid also for the studies reviewed below.

2.2 The linguistic impairment in PADs

The linguistic impairment of PADs is defined as a form of fluent aphasia, which is caused by major disruptions in the semantic and lexical domains. Consequently, their discourse is often described as empty, although their speech is fluent (Cummings et al., 1985; Nicholas et al., 1985; Murdoch et al., 1987; a.m.o.). The majority of PADs does not show signs of impairment at the phonological and articulatory levels (Chertkow et al., 1989; Chertkow & Bub, 1990; Patel & Satz, 1994).

Phonological and articulatory difficulties are only rarely detected. For instance, Croot, Hodges, Xuereb & Patterson (2000) report only ten cases of patients with such a form of impairment. By taking into account a variety of factors, the authors manage to point out that the phonological impairment appears more often in PADs that are characterized by at least one of the three following conditions: *i)* early onset of the disease (before the age of 65), *ii)* familial Alzheimer (at least one parent was demented), or *iii)* peculiar distribution of brain damages. As for the latter condition, they found out that patients with phonological impairment are characterized by early damages at the left temporal, parietal and frontal perisylvian areas (Croot et al., 2000:3003), while their hippocampus is relatively spared. This is precisely the opposite pattern of how damages spread in the majority of PADs. Following the results described above, we can assume that phonological and articulatory abilities are usually spared in PADs, unless specific conditions intervene.

Another frequent impairment among PADs is the progressive loss of reading and writing abilities (Platel et al., 1993; Patterson, Graham & Hodges, 1994; Harnish & Neils-Strunjas, 2008). Some authors ascribe it to the lexical loss in AD, especially for languages characterized by a low level of spelling transparency. However, inertia, apraxia and visuo-spatial impairments might also play a relevant role in the disruption of writing and reading skills. As for Italian, Luzzatti, Laiacona & Agazzi (2003) find high variability in writing proficiency across patients and cannot describe a predominant pattern. Overall, writing impairments are often observed, but the issue is far from being set. Given that the impairment is not relevant for the aims of the present study, I will set the problem aside.

In the following section, I will focus my attention on the three most interesting aspects of the linguistic competence of PADs, namely anomia, the dissociation between regular and irregular morphology and sentence processing.

2.3 Anomia

Relatives and caregivers often notice a specific difficulty at retrieving lexical items as one of the first symptoms of the disease. PADs increase their use of semantically empty words, like *thing*, *it* and *do* in order to substitute semantically specific nouns and verbs, thus failing at conveying the intended meaning. The deficit is known as *anomia*.

For instance, the affectionate readers of the British writer Iris Murdoch noticed her incapacity to use vocabulary properly in her final novel *Jackson's Dilemma* published in 1995. The book overall received poor reviews from critics. Short after the publication, Murdoch showed clear signs of dementia and was diagnosed with Alzheimer. Garrard et al. (2005) analyse Murdoch's final novel and compare it to two novels that the author had published earlier in her career. Their aim is to detect early symptoms of the disease in her writing. The lexical analysis they conduct reveals substantial differences between the two novels published earlier in her career and the last one. In particular, the range of vocabulary in use appears much more limited and the selected lexical items are characterized by higher frequency. In other words, the author resorts to a smaller pool of high frequency words in her final work. The observation is highly compatible with the diagnosis received by Murdoch and it represents a clear example of what PADs experience.

Two different sources of impairment could cause anomia. Some authors claim that PADs have difficulties at retrieving the target lexical items they need, despite their presence in the lexicon (Diesfelt, 1985; Nebes, Martin & Horn, 1984). The hypothesis overlooks the possibility of lexical disruption and calls into play factors related to impoverished cognitive abilities, e.g. the fast decay of lexical activation or the unsuccessful inhibition of alternatives.

Other authors claim that anomia is caused by a disruption in the lexical and semantic levels (Chertkow et al., 1989; Chertkow & Bub, 1990; Hodges et al., 1991, 1996; Almor et al., 2009, among others). In this case, it is not possible to retrieve lexical items precisely because they are not available in the lexicon of PADs anymore. Under this view, the disease progressively erases patients' lexical knowledge.

Overall, studies show that the disruption of semantic knowledge plays a major role in causing anomia in PADs although difficulties at lexical retrieval are also relevant (Chertkow & Bub, 1990; Chertkow et al., 1990).

In the following sections, I present three different kinds of evidence that support the idea of a semantic knowledge loss in PADs: a item-by-item analysis (2.3.1), the relation between naming and knowing (2.3.1.1), and the error analysis (2.3.1.2). I will also consider the factors that influence anomia (2.3.1.3) and the dissociation between nouns and verbs (2.3.1.4). Finally, in section 2.3.2 I review studies that found signs of preserved semantic knowledge activation in PADs.

2.3.1 Semantic knowledge loss in PADs

The study of anomia in PADs developed at the beginning of the '80s. As anticipated, the discussion mainly concerns the source of impairment in PADs. In particular, authors evaluate the status of the semantic knowledge in PADs. Nebes, Martin & Horn (1984) and Diesfeldt (1985) initially claim that the semantic knowledge of demented patients is intact and that their anomia is the manifestation of a retrieval impairment. However, in the same years, Bayles & Tomoeda (1983) and Huff, Corkin & Growdon (1986) support the exact opposite, by pointing out an impairment at the semantic level, with major consequences on the lexicon. In particular, Huff, Corkin & Growdon (1986) notice that patients tend to make mistakes on the same nouns across different tasks and suggest that anomia should be studied through item-by-item analyses.

Starting from that idea, Chertkow, Bub and Seidenberg (1989) collect sound data in favour of the semantic impairment hypothesis. They use a battery of tests based on a list of 150 nouns: all selected nouns refer to concrete objects retrieved from different semantic categories (fruits, vegetables, animals and tools, among others). Six patients with a diagnosis of probable Alzheimer's disease, who had previously undergone reading and visual tests in order to discard perceptual disorders as possible confounding factors, complete a naming and a word-to-picture matching task. As expected, their performance on both tasks is poorer in comparison to the one of healthy elderly speakers. Authors also compare data from the two tests and conduct an item-by-item analysis: results reveal that most patients either perform always correctly or always wrong on the same item in both tasks. In other words, if patients are able to match a word to the correct picture, they are also able to name the same object. In contrast, if patients fail the first task, they will most probably fail also the second one. The number of items on which patients perform well in one task, but wrong in the other one, is very narrow. The individual consistency across

tasks in patients' performance suggests that lexical items can be either spared or disrupted in the lexicon of PADs. Successful activation seems not to depend on the activation route: both visual and oral prompting lead to equivalent rates of successful retrieval.

In order to corroborate their results, Chertkow et al. (1989) repeat the naming test and ask patients twelve forced–choice questions concerning the attributes, the superordinate category and the functional characteristics of each element (e.g., for the word *scissors*: 'Is the tip made of metal or wood? ', 'Is it a tool or clothing? ', 'Do you cut things with it or lift with it?', etc.). Whenever a participant can answer correctly to at least nine questions, the item is classified as 'intact' in the subject's lexicon, otherwise it is classified as 'degraded'. A cross comparison of results confirms patients' consistency: intact items are usually named correctly and recognized correctly in a sentence-to-picture matching task. In contrast, degraded nouns correspond to lexical items on which patients had failed both previous tests.

In a follow-up study, Chertkow & Bub (1990) enlarge their experimental group up to ten PADs and enrich their battery with new tasks: semantic cue, verbal fluency and on-line semantic priming. The semantic cue technique consists in helping participants when they cannot name a picture by providing a semantic information (e.g. 'it is similar to a tiger' when the participant cannot name the picture of a lion). The procedure is effective only with words which patients had previously shown to have intact knowledge of[1]. For 'degraded' items, semantic cues are of no help for lexical retrieval.

Also results from verbal fluency pair the data presented so far: patients overall produce only 40% of the words[2] recalled by controls. Moreover, in the verbal fluency task they (almost) never spontaneously produce words that they had not been able to name previously. However, in verbal fluency tasks they do not recall all words that are otherwise classified as correct or intact in other tasks. In other words, they do not succeed at retrieving all the intact nouns that they had completed successfully in previous tasks.

[1] As describe for Chertkow et al. (1989), 'intact knowledge' is defined in the study as the capacity of answering to at least 9 out of 12 questions concerning the object. All participants in Chertkow & Bub (1990) undergo this test in a separate session from the one in which naming and semantic cues are administered.

[2] The mean number of words recalled in one minute highly varies across the prompted semantic category. PADs are sharply more successful at retrieving nouns related to body parts and clothes, most probably because they use the visual cues in the context. In contrast, they make many mistakes on the vegetable and fruit categories, by mixing up items across the two.

Finally, the semantic priming test is performed through a lexical decision task. Participants see a noun on the screen and must decide whether it is an existing word or not. Either a semantically related noun, an unrelated one or a pseudo-word precedes each experimental noun. Reaction times are recorded: patients are faster at answering to real words ('nail') when these follow semantically related nouns ('hammer'), than when they follow semantically unrelated nouns ('horse'). Authors conclude that semantic priming is spared in PADs, despite generalized slower reaction times.

In sum, Chertkow et al. (1989) and Chertkow & Bub (1990) succeed at showing that an item-by-item consistency across tasks exists in the performance of PADs. This can be explained in light of a semantic knowledge impairment. According to the author, lexical items are either intact in the lexicon, and can therefore be retrieved through different routes, or they are degraded. In the second case, the retrieval is never successful, independently of the technique in use. However, in the light of results from the last two tasks, namely verbal fluency and semantic priming, Chertkow & Bub (1990) must recognize that some retrieval difficulties are also at play. Verbal fluency shows that PADs cannot recall all the intact nouns they can count on in their lexicon, thus clearly showing an impairment in the retrieval process. Finally, the presence of enhanced priming effects in PADs indirectly speaks against the hypothesis of a semantic impairment, as it shows that the semantic knowledge network makes lexical activation flow through related items. Authors explain the priming effect by assuming an asymmetry between automatic and explicit activation. Naming tasks, word-to-picture matching and semantic probe questions all require explicit activation of lexical items, while semantic priming favours automatic (thus implicit) lexical activation. The assumed asymmetry speaks in favour of a lexical retrieval impairment, which could play a role along with semantic impairment.

2.3.1.1 Naming & Knowing

In order to strengthen the hypothesis of a semantic deficit, it is useful to verify the information that patients retain for objects they can name, and for objects they cannot name anymore. For this purpose, Hodges, Patterson, Graham & Dawson (1996) enrol 51 PADs in two different tests: a picture naming task and a definition task. In the first task, participants have to name the objects presented in 48 line drawings. In the second task, they are requested to explain what the objects are, as if they were talking to someone who

has never seen those objects before. Answers are scored by sorting the provided information into nine different categories: general physical features (e.g., for a giraffe 'It has four legs'), specific physical features (e.g., 'It has a very long neck'), general associative information (e.g., 'it eats food'), specific associative information (e.g., 'It lives in Africa'), superordinate category (e.g., 'It is an animal'), core concepts, intrusion errors (e.g., 'it lives in water'), judgements (e.g., 'It is cute') and irrelevant information (e.g., 'I saw one on TV'). The distinction between general and specific information refers to the ability to distinguish between broad properties shared by many items within the same category (e.g., having four legs) and narrow features, which contradistinguish a specific item from the similar ones. During the description task, PADs and controls recall almost the same amount of general information. As for specific information, PADs generate significantly less elements than controls, especially in the case of associative information. In total, PADs more often define elements by their superordinate category (e.g., fruit, vegetable, tool, music instrument), and provide poorer core information in comparison to controls. These asymmetries become more evident in the case of objects that patients cannot name. By comparing results on the description task to those on the naming task, authors manage to point out that the capacity to name a picture is crucially related to the amount of information the person can provide about the object. In particular, authors identify a correlation between the knowledge about specific physical characteristics and the capacity to name objects. Whenever patients can provide little specific information, it is highly improbable that they can name the element. In contrast, specific knowledge correlates with high naming scores. In order to summarize the outputs of their study, Hodges et al. (1996) use the expression *Naming & Knowing* to precisely describe the fact that the capacity to name depends on how well the speaker knows the objects.

A follow-up study shows that the capacity to name decreases in time along wth information loss. Ralph, Patterson & Hodges (1997) invite a subset of 10 PADs for a second round of tests after several months (mean: 15.6 months). Participants perform again a naming task and a description task. Answers are scored according to the same criteria and results from the two rounds are compared according to an item-by-item analysis. The authors focus in particular on items that the patients name correctly in the first round but not in the second one; the two corresponding descriptions are subsequently compared. It emerges that the description of items that move from the 'named' to the

'unnamed' category is consistently impoverished. The loss of both physical and associative information determines the incapacity to name items. However, the incidence of the two information categories is not equal across items. If items of the natural kind and artefacts are compared, Ralph et al. (1997) notice that natural items become unnamed when physical information is lost; in contrast, artefacts become unnamed when associative and functional information is lost. Through this comparison, authors manage to point out how the different categories of features do not show asymmetries *a priori*, in the sense that none of the two is systematically more impaired than the other one. Rather, the effect of their impairment depends on the items at stake. Results are captured by the Weighted Overlappingly Organized Features (WOOF) model (Ralph et al., 1997): the core information of an item is organized according to the nature of the item itself, such that physical features are more relevant for natural items than for artefacts, and functional features are more relevant for artefacts than for natural items. The capacity of naming an element crucially depends on the retention of core information about the elements. Once these fade away, the capacity to retrieve the target name is lost. This explains the lack of physical information in the description of unnamed natural items and the lack of associative/functional information for unnamed artefacts[3].

2.3.1.2 The error analysis

In addition to item-by-item analyses and to the observed relation between naming and knowing, a third kind of evidence supports the hypothesis of a progressive semantic deficit as main cause of anomia in PADs. Results from error analysis in naming tests show a gradual loss of information. Hodges, Salmon & Butters (1991) conduct a longitudinal study by following a group of PADs over three years and they observe how error patterns in naming tasks change along with the worsening of the disease. Once a year participants complete a 30-item version of the Boston Naming Test (Kaplan, Goodglass & Weintraub, 1983). In three years, the average number of mistakes increases

[3] More studies address the hypothesis of a category-specific impairment that causes asymmetries between living and non-living items (Almor et al., 2009; Cross, Smith & Grossman, 2008; Gainotti et al., 1996; Hodges et al., 1996; Montanes, Goldblum & Boller, 1996). However, Gainotti et al. (1996) find a better accuracy on living than on non-living items; Almor et al. (2009) and Montanes et al. (1996) find the opposite pattern; while Hodges et al. (1996) do not find any asymmetry. In addition, the comparison is made difficult by the use of different experimental techniques and the enrolment of different participants, such that clear conclusions cannot be drawn.

from 8.9 to 14.7 (out of 30 items), thus allowing for error analysis. Hodges et al. (1991) observe that superordinate errors are very common throughout the observation time. This means that PADs often use category labels (e.g., 'animal', ' musical instrument') when they cannot retrieve the target noun (e.g., 'rhinoceros', 'violin'). Errors of the semantic-associative kind robustly increase over the three years. By “semantic-associative” errors, it is meant that participants name an item that is clearly semantically associated to the target one, although it does not strictly belong to the same category (e.g., 'doctor' for 'stethoscope' or 'ice' for 'igloo'), such that category boundaries are often violated (i.e., they name a person for a tool and a material for a object). Accordingly, errors within the same category (e.g., 'hippopotamus' for 'rhinoceros') gradually decrease, as well as circumlocutory answers do. Results on same-category errors pair data on semantic errors: PADs gradually resort to broader semantic relations between items. Circumlocutory answers show that PADs manage to retrieve the semantic information related to the item, but they cannot retrieve the target noun. A reduction of this kind of answers further proves that PADs gradually lose fine-grained semantic information.

Besides semantic and lexical related mistakes, Hodges et al. (1991) also observe a small (but increasing over time) percentage of visual errors. Perceptual errors are nonetheless very rare in the first year. This, together with the results on perceptual abilities presented by Chertkow et al. (1989), Chertkow & Bub (1990), and Nicholas et al. (1996) allow discarding misperception as the cause for the anomia that characterizes PADs in the first stages of their disease. However, it cannot be excluded that a disruption in visual perception further affects patients later in the disease.

As for error analysis, not all authors agree in interpreting semantic errors as a sign of semantic knowledge disruption. For instance, Nicholas et al. (1996) analyse naming errors according to their semantic relatedness with the target word. In their work, every mistake receives a score according to a 5-point scale: the higher the score, the closer the error and the target word are in meaning. In contrast, lower scores are assigned to outputs that are semantically unrelated to the target. The scoring procedure applies both to PADs and controls. Results are compared, but no qualitative difference emerges between the two groups. Their performance seems to differ only in the quantity of errors, while semantic relatedness between outputs and targets is homogeneous across impaired and

unimpaired speakers. Based on these data, Nicholas et al. (1996) reject the hypothesis of a progressive semantic disruption in PADs.

Almor et al. (2009) also compare PADs and controls and find similarities and differences that, altogether, speak in favour of a gradual semantic disruption. Similarly to what Hodges et al. (1991) did, they also analyse naming mistakes according to categories rather than semantic relatedness; however, they do not follow their patients over an extended time. In order to observe how mistake patterns change along with the worsening of the disease, Almor et al. (2009) enrol two groups of patients: high-performing PADs and low-performing PADs. A comparison among results from controls and the two groups of PADs allows for the individuation of a specific pattern of disruption. The prevailing mistake category shifts from within-category outputs, to superordinate labels and finally to nonresponses. Interesting observations can be inferred from these results, if we assume that lexical activation starts from semantic feature activation (Levelt, 1989) and we interpret mistakes as the manifestation of the kind of semantic features patients can access. That is precisely what Almor et al. (2009) do. In their view, within-category mistakes represent the successful individuation of core characteristics of the item, even though the target output is not produced. The error is therefore due to the misanalysis or to the omission of fine-grained features. In contrast, superordinate errors reveal that participants can activate only general information, which apply correctly to a broad set of items. From the cognitive point of view, this can be read as a symptom of the fact that patients only retain major semantic categories for the analysis and the classification of world knowledge. Their knowledge gradually disregards specific information in favour of macro-categories. In some cases, macro-categories correspond to labels (e.g., 'musical instrument' or 'animal'), in others they are represented by one of their prototypical members (e.g., 'guitar' for all instruments or 'dogs' for all four-legged animals). Finally, in more severe stages of the disease, nonresponses prevail over all other mistakes; this reflects patients' inability to perform the task. Almor et al. (2009) also observe a relevant similarity between PADs and controls with respect to the distribution of mistakes. Neither controls nor PADs have a specific impairment in any of the noun or verb categories used in the task. Although some tendencies towards asymmetries are visible (e.g., better accuracy on non-living than on living items), it is never the case that some categories are completely spared, while others are severely impaired. Therefore, this pattern of mistake

distribution is compatible with a model of "graceful degradation of connections between features and concepts" (Almor et al., 2009: 17), which gradually reduces semantic knowledge in PADs, starting from fine-grained features and continuing with the most general ones.

Moreover, the error patterns presented in Hodges et al. (1991) and in Almor et al. (2009) are confirmed by cross-linguistic data. Cuetos, Gonzalez-Nosti & Martínez (2005), conduct a similar naming task with Spanish-speaking patients. Again, a comparison between patients' error patterns in two different rounds of tests reveals that their naming accuracy significantly reduces in two years. An item-by-item analysis confirms that many items shift from correct (in the first test administration) to semantic mistakes or to circumlocutions (in the second administration). Accordingly, experimental trials coded as semantic mistakes or circumlocutions in the first round become nonresponses in the second one.

2.3.1.3 Factors of lexical disruption

In this section, I will briefly touch upon the factors that seem to influence lexical disruption. As anticipated (see above and footnote 3), anomia affects the lexicon in a generalized fashion, with no clear asymmetries among semantic categories (e.g., living/non-living items, natural elements/artefacts, etc.). However, other factors seem to be good predictors for lexical disruption in PADs. These are age of acquisition, frequency and familiarity, and word internal complexity.

As for age of acquisition, many studies show that this factor correlates with accuracy both in picture naming (Cuetos & al., 2005; Cuetos et al., 2012; Silveri et al., 2002), in word recognition tasks (Cuetos et al., 2010, 2015), and in verbal fluency tests (Forbes-MacKay et al., 2005). Venneri et al. (2008) also investigate the neural basis of the phenomenon and find that the effect of age of acquisition increases its magnitude in correlation with loss of cortical tissue. There is sound evidence that early-acquired words (within age 6 according to Cuetos et al., 2015) are more resilient to lexical disruption in comparison to words acquired after age 8.

Frequency is a good predictor of patients' performance on lexical tasks too. Cuetos et al. (2005, 2015), Forbes-MacKay et al. (2005), and Tippett et al. (2007) find consistent results: PADs retrieve high-frequency words with more ease, in comparison to low-

frequency words. The effect emerges also when cumulative frequency[4] is considered (Small & Sandhu, 2008).

In addition, Small & Sandhu (2008) investigate the interplay of semantic and episodic memory with respect to lexical retrieval. Through a picture naming task, authors show that PADs are more successful at naming pictures that present objects from the beginning of the 20th Century, in comparison to pictures that present the same objects from the end of the 20th Century. For instance, PADs are more successful at recognizing and naming the picture of an old-fashioned phone, rather than that of a modern mobile phone. The observation is of crucial interest for the design of tasks that include visual stimuli.

Gainotti et al. (1996) take into consideration the index of familiarity of the lexical items they use in a naming and in a word-to-picture matching task and confirm that this also correlates with results: both naming and comprehension benefit from high-familiarity[5].

As for word internal complexity, PADs are particularly impaired at retrieving compounds (Chiarelli, Menichelli & Semenza, 2007). In contrast, there is no agreement on whether word length plays a role (Forbes-MacKay et al., 2005) or does not (Cuetos et al., 2005).

2.3.1.4 Verbs in anomia

Most studies on anomia in PADs focus on nouns; however, whenever verbs are inserted among the experimental materials, an asymmetry between the two lexical classes is observed. PADs are usually more impaired on verbs than on nouns (Almor et al., 2009; Bushell & Martin, 1997; Grossman et al., 1996; Robinson et al., 1996), in line with what is usually observed in aphasia (Mätzing et al. (2009) offer an overview on the topic).

One interesting piece of information concerns the syntactic information connected to verbal lexical entries. Kim & Thompson (2003) first claim that the syntactic complexity of verbs (calculated based on the number of arguments) does not influence verb retrieval

[4] Cumulative frequency of words is usually measured by subtracting the age of acquisition of the word from the speakers age. The obtained value is then multiplied by the word's mean frequency in order to obtain cumulative frequency (Cuetos et al., 2012:250).

[5] Gainotti et al. (1996) also suggest that familiarity might play as confounding factor in tasks that aim at investigating category-specific effects. For instance, in their view living items are nowadays low-familiar to most speakers.

and use, while semantic complexity does. Patients are more impaired at retrieving semantically complex verbs (e.g., 'polish') that semantically simple verbs (e.g., 'clean'). In contrast, the comparison among verbs with either one, two or three arguments does not allow for the observation of clear patterns of impairment (while it does for the agrammatic patients in the same study). Clear conclusions cannot be met though, as the statistical analysis reveals that PADs perform better on one-place and three-place verbs than on two-place verbs. The peculiar pattern casts some doubts on the data and the conclusions.

Also Grossman et al. (2007) suggest the hypothesis of a better preservation of syntactic information over semantic content in lexical entries. In their study, authors make participants familiar with the word *lour*, a very low frequency English verb[6]. Participants listen to the verb in a story-telling exercise and in a sentence-to-picture matching task. After giving them the opportunity to get familiar with the verb, its semantic meaning and its thematic structure, authors invite participants to complete some more tasks. PADs undergo acceptability tasks that concern the lexical class and the thematic structure of the verb *to lour*. A picture-word matching task samples their comprehension of the verb semantic content. Patients perform well on the former kind of tasks, but are not accurate on the latter. Thus, it emerges that they have acquired that *lour* is a verb (grammatical information), and they are partially aware of its thematic structure, but they struggle in remembering the word meaning. Tests are repeated after one week from the first exposure and results are replicated: the performance is more accurate on tasks that concern the grammatical information of the newly acquired word than on tasks that concern its meaning.

Together with the results from Kim & Thompson (2003), Grossman et al. (2007) suggest that PADs might present a dissociation between semantic and syntactic information in their mental lexicon: the former appears clearly impaired, while the latter could be spared. Unfortunately, the evidence gathered by Grossman et al. (2007) is weak, due to the fact that tasks are based only on one verb and performed only by eleven patients. For these reasons, strong claims cannot be put forth. However, the idea of a possible dissociation between semantic and syntactic information within lexical entries is

[6] Only participants who show no signs of previous knowledge of the verb *to lour* are admitted in the study (Grossman et al., 2007).

very appealing. I will further address the issue in Chapter 4, where I will provide evidence for retained syntactic information in nouns.

2.3.2 Spared semantic knowledge and the access problem

In the previous sections I reviewed studies that provide evidence of a semantic disruption in PADs as the main cause to anomia. However, not all results point in this direction: some studies also report a relative spared semantic knowledge despite dementia. For instance, Chertkow & Bub find that PADs benefit from lexical priming in a word decision task (see above). Whenever a semantically related word precedes the target one, reaction times are reduced. The observation suggests that semantic knowledge must be spared if it makes the activation flow through related nodes; otherwise, priming effects would not be visible. Nicholas et al. (1996) and Astell & Harley (1996) present more data in favour of spared semantic knowledge.

As reviewed above, Nicholas et al. (1996) analyse the errors made by PADs in a naming task by rating the level of semantic relatedness between the wrong output and the target word. Mistakes made by PADs are compared to mistakes made by controls, but no relevant difference between the two groups is found with respect to semantic relatedness: PADs get just as close to the target word as controls do. Ultimately, differences are only of the quantitative kind, with PADs making far more errors than controls.

Astell & Harley (1996) reach similar conclusion through a different task. They elicit naming through oral descriptions: participants listen to a definition and must produce the corresponding word. Their aim is to provoke tip-of-the-tongue states in their participants, in order to observe at which stage lexical retrieval is blocked (see Chapter 3). Again, PADs fail at naming the target word more often than healthy controls do. Their wrong outputs are closely related to the target one, such that it is evident that the provided description succeeds in activating the target semantic information. However, authors cannot explain the reasons why PADs do not complete the task accurately and leave the conclusions open to two alternative interpretations: there is either an impairment at the level of activation flow between the selected lemma and the corresponding phonological form of the word, or patients have difficulties at inhibiting the competing alternatives. Overall, Nicholas et al. (1996) and Astell & Harley (1996) underline the semantic relatedness of mistakes in comparison to targets and take it as a sign of spared knowledge,

even though many more studies argue against it (Almor et al., 2009; Chertkow et al., 1989; Hodges et al., 1991, 1996; Ralph et al., 1997, among others).

Furthermore, Nebes & Halligan (1996)[7] show that PADs are able to activate low-marginal characteristics of items when directly prompted to do so. Their study is based on the idea that all items are defined by a set of features; among these, not all characteristics are equally relevant. For instance, 'music' or 'sound' are high-dominant in the definition of 'piano', while 'heavy' is low-dominant. This is due to the fact that, although 'heavy' is a correct description for a piano, it is not the first thing speakers think of when they hear the word 'piano'. Nebes & Halligan (1996) aim at verifying whether PADs show the same pattern of activation as controls for high- and low-dominant characteristics of words. The experiment design includes context sentences that end in a target word. Right after hearing the context sentence, participants read a question that concerns a low-dominant characteristic of the target word and answer with 'Yes' or 'No', depending on whether the characteristic can describe the target word or not. Context sentences can be appropriate (1a), inappropriate (1b) or neutral (1c) with respect to the low-dominant characteristic:

(1) a. Mary asked John to lift her piano. Is a piano heavy?

b. The old man loved to listen to the piano. Is a piano heavy?

c. In the store window was a new piano. Is a piano heavy?

Young and older healthy controls show facilitating effects for the activation of the low-dominant feature in appropriate context (1a). This is made evident by shorter reaction times in (1a) in comparison to the neutral context in (1c). Inappropriate contexts (1b) do not cause prolonged reaction times though. Interestingly, the same pattern is found in PADs. Despite overall slower reaction times, they are equally fast on the neutral and on the inappropriate conditions, while they are faster on the appropriate context. Nebes & Halligan interpret the result as a proof of the fact that even marginal features can be

[7] See also Nebes & Brady (1990) for similar results.

activated in the semantic knowledge of PADs under the appropriate conditions. Based on the results, authors argue against the hypothesis of gradual semantic disruption in PADs.

In conclusions, leaving aside for a moment the discussion on the status of the semantic knowledge in PADs, it must be taken into account that a difficult access to the lexicon might further exacerbate anomia. Chertkow & Bub (1990) recognize signs of difficult access to the lexicon in a verbal fluency task: PADs can recall fewer words than what they actually know. For instance, when asked to name animals, they omit many names, which they can actually retrieve in other tasks. Further support to the hypothesis of difficult access to spared information is required though, in order to evaluate the relevance of this kind of impairment in anomia.

2.3.3 Interim summary

In section 2.3 I have reviewed studies that investigate the anomic deficit in PADs. It is now generally assumed that anomia in PADs is mainly due to a semantic knowledge impairment. Three kinds of evidence support the claim. First, item-by-item analyses show consistent results: PADs are always either right or wrong on a single item across tasks (Chertkow et al., 1989; Chertkow & Bub, 1990). Second, the capacity to name an object seems to correlate with the knowledge the patient retains about the object. Hodges et al. (1996) summarize the idea under the motto *Naming & Knowing*. Third, the analysis of errors from naming tasks confirms that patients progressively reduce their accuracy in answering. Their error pattern shifts from within-category items, to superordinate labels and finally to nonresponses (Hodges et al., 1991; Almor et al., 2009; Cuetos et al., 2015). The progress of anomia is influenced by age of acquisition (Cuetos et al., 2005, 2015) and frequency (Forbes-MacKay et al., 2005; Tippett et al., 2007, among others).

Finally, PADs perform significantly lower on verbs than on nouns. In addition, Grossman et al. (2007) and Kim & Thompson (2003) put forth the idea of a dissociation between syntactic and semantic information in the lexical entries of verbs, but fail at providing sound supporting data.

2.4 Verbal morphology

The number of studies that investigate verbal morphology is rather limited, but their results are sound. The interest focuses on the asymmetry between regular and irregular

verbal morphology, in accordance to the hypothesis advanced by Ullman (2001) of a dual-system for language processing.

In Ullman's (2001) model, language processing counts on two systems that replicate the dissociation between the declarative and the procedural memory: speakers count on a mental lexicon and on a mental grammar. The former belongs to the declarative system because it stores lexical entries; the latter is procedural in nature because it allows for the rule-driven derivation of different forms. In many languages, the verbal system offers the opportunity to test the status of the two components, because it includes both regular and irregular forms. According to Ullman (2001), the procedural system is responsible for producing regular forms, while irregular forms are retrieved from the mental lexicon, i.e. the declarative memory.

Walenski et al. (2009) test Italian PADs with a sentence completion task that elicits the production of regular and irregular present-tense and past-participle forms. The study aims at verifying the status of the procedural and declarative linguistic components in patients with Alzheimer.

Italian verbs distribute across three different inflectional classes. These can be recognized by the thematic vowels in the infinitive forms. Class I mainly includes regular verbs, which take the morpheme *–are* at the infinitive and *–ato* at the past-tense (*ballare* 'to dance' → *ballato*). This class is still very productive, in the sense that neologisms assume its inflectional paradigm. Class II is characterized by the *–ere* marker at the infinitive. Many verbs in this class are irregular and change stem at the past (e.g. *prendere* 'to take' → *preso*). Finally, Class III includes verbs that end in *–ire* at the infinitive. The class includes both regular (*sentire* 'to hear',→ *sentito*) and irregular past-tense forms (*aprire,* 'to open'→ *aperto*).

The sentence in (2) is an example of the experimental trials in use in Walenski et al. (2009):

(2) *A Giovanni piace ballare il tango. Allora ieri Giovanni ha _____ il tango*
'Giovanni likes dancing tango. So yesterday Giovanni has _____ tango.'

Results from the experimental task show that PADs are spared at producing the regular forms, while they are impaired at retrieving the irregular ones.

In order to test patients' ability to use their procedural knowledge productively and to recognize and exploit morphological information, the task also includes some novel verbs. Authors make up regular verbs for Class I (*carlare*) and irregular ones for Class II (*schidere*), depending on the thematic vowels (*-are* for the regular Class I and *–ere* for the irregular Class II). Patients are consistent in their behaviour in the sense that they manage to produce correct forms for novel regular verbs, while they are impaired with novel irregular verbs. In other words, patients recognize the target inflectional class by the thematic vowel in the infinitive and perform accordingly. They produce regular forms for novel verbs in *–are*, while they hesitate with novel verbs in *–ere*, thus recognizing that these belong to the irregular class. Their mistakes are crucially not of the regularization kind though. They try to produce irregular forms, but their guesses do not meet those provided by healthy controls[8].

Patients' ability to recognize and take advantage of morphological cues is further confirmed by a similar task conducted by Colombo, Fonti & Stracciari (2009). Authors design two different conditions for the task. In the first condition, verb production is elicited with infinitive forms as in (2); in the second condition, participants see the first person singular of the target verb (3).

(3) *Io ballo tutti i giorni. Quindi ieri ho* _____
'I dance everyday. So yesterday I have _____'

The first person singular form is characterized by a *–o* final morpheme, which does not allow for the recognition of the inflectional class, as the three verbal classes all share the morpheme *–o* for the first person singular. The performance of PADs in the second condition is poorer in comparison to the first condition, in which morphological cues about the inflectional paradigm are overtly realized on the infinitive. This means that patients' impairment becomes more evident when no morphological cue is at hand. In contrast, if any morphological cue is available, they use it to perform accordingly. Colombo et al. (2009) do not completely agree on the hypothesis of a strict dual system, though; they claim that the declarative and the procedural linguistic components rather

[8] Target forms of novel irregular verbs are determined according to the intuitions of healthy speakers, based on phonological similarities with real irregular verbs.

distribute along a continuum, and contribute to language processing according to variable ratios.

Cameli et al. (2005) present similar results from two bilingual English-French PADs. The pattern described above is replicated only for the first language though. In the weaker language (acquired as L2 during puberty and then used throughout adult life), the performance resembles the one on irregular verbs for the native language. This reveals that verbal forms of the L2 are all stored in the mental lexicon, and they are not produced by the mental grammar.

Fyndanis et al. (2013) investigate Greek verbal morphology with rather different results. Authors design sentence completion and grammaticality judgement tasks in order to sample the status of verbal morphology, with respect to agreement, tense and aspect. The Greek verbal system is very rich and counts on morphological markers for the three categories. Greek patients perform with good accuracy on items that concern subject/verb agreement, while they make many mistakes on items that concern tense and, most of all, aspect. Fyndanis et al. (2013) deduce that agreement is spared because it only involves syntactic information. In contrast, tense and aspect correspond to interpretable features at LF, their implementation requires the integration of morphological information as well as the integration of discourse/semantic information, thus resulting in a higher level of complexity for PADs.

The data interpretation offered by Fyndanis et al. (2013) sounds rather plausible. However, it is at odds with the results from the studies reviewed above, in which patients do not have a specific impairment at realizing past-tense forms, which entail both a tense feature (past) and an aspectual feature (perfective).

In my opinion, the asymmetry in results is due to differences in administration procedures. In the studies by Walenski et al. (2009) and by Colombo et al. (2009), the tasks in use include only one kind of manipulation at the time, such that participants see sentences with verbs at the infinitive form and have to produce the equivalent past form. Authors assume that the task directly tackles verbal morphology, because it keeps interpretative and post-interpretative demands at a minimum. In contrast, Fyndanis et al. (2013) test agreement, tense and aspect within the same task. Consequently, patients constantly have to shift the focus of their attention and to process sentences fully in order to provide a congruent answer. The execution of this kind of task is therefore much more

demanding. Finally, as the authors themselves point out, tense and aspect are rather subjective information, because they depend on the speakers' perspective and attitude towards the sentence content, which are rather difficult information to convey in out-of-the-blue sentences.

In sum, PADs seem to be spared at producing regular verbal morphology, while they are impaired with irregular verbal morphology. Data from studies with Italian speakers (Colombo et al., 2009; Walenski et al., 2009) show that their morphological competence is active and productive as they can actively use morphological cues and they inflect novel verbs accordingly. Together with studies on anomia (see above), results from production of irregular verbs further confirm that the mental lexicon of PADs is disrupted.

Finally, these conclusions hold true for verb production as a purely morphological phenomenon. When the task in use demands the integration of semantic information and interpretable features at LF, the overall performance of PADs is not accurate anymore, as suggested by the results in Fyndanis (2013).

2.5 The syntactic competence in spontaneous speech

Schwarz, Marin & Saffran (1979) first conduct a case study on an aphasic patient diagnosed with probable Alzheimer's disease; they observe that the patient possesses intact syntactic abilities, which allow for accurate comprehension and production of a variety of syntactic structures. At the time of publication, the paper contributes to a discussion concerning the existence of a dissociation between the semantic and the syntactic competence in patients with dementia (Bayles, 1982; Irigaray, 1967, 1973; Warrington, 1975; Whitaker, 1976). In its strongest version, the claim posits that the former is impaired, while the latter is usually spared in a variety of neuropsychological pathologies concerned by the decrease of cognitive abilities. Schwarz, Marin & Saffran (1979) align the discussion on Alzheimer's along these lines, such that it is initially assumed that PADs do not show any asymmetries in comparison to their healthy peers with respect to their syntactic abilities.

Few years later, Kempler, Curtiss & Jackson (1987) strengthen the hypothesis by providing data from the speech of ten English-speaking PADs. Authors collect fifty spontaneously produced sentences per patient and they analyse the corpus according to a variety of syntactic parameters, which includes the number of simple sentences, that of

adverbial clauses, of relative clauses and of conjunctions (among others). Altogether, sentences from the PAD group do not significantly differ from those of the control group, except for the presence of a considerable larger number of mistakes of the semantic kind. In order to explain the results, authors resort to the idea that syntactic competence builds on a high level of automaticity. In their view, the number of morphosyntactic structures a language can count on is rather limited with respect to the size of its lexicon. It follows that lexical selection is a demanding process, whose outcome highly depends on the context. In contrast, the selection of the morpho-syntactic structures to be used is implemented automatically. Speakers develop automatic sequences during acquisition and reinforce their use during the life span, such that the level of automaticity is so high that the disease cannot affect the syntactic competence.

The hypothesis of a lack of difficulties at the level of syntactic planning is supported also by a more recent study: Gayraud, Lee & Barkat-Defradas (2011) observe the spontaneous speech of French-speaking patients with a specific focus on their pauses and hesitations. Author notice that PADs interrupt their speech more often than their healthy peers do. Their pauses and hesitations do not contain discourse fillers and often conclude with the utterance of low-frequency words. So far, these observations are not surprising as they are in line with the description of speakers that suffer from difficulties at discourse planning and lexical selection. However, the important contribution of the study comes from the observation of the distribution of pauses. They usually occur outside main syntactic phrases, i.e. at boundaries between phrases. Phrases are only rarely interrupted, which means that their internal planning is not problematic (provided that patients already activated the necessary vocabulary).

More studies on spontaneous speech succeed at pointing out major differences between PADs and controls. Nicholas, Obler, Albert & Helm-Estabrook (1985) first suggest that discourse analysis can be successful in the diagnostic process, in order to distinguish among PADs, Wernicke aphasics and anomic aphasics. For instance, PADs distinguish themselves from the other groups for the use of many deictic and indefinite forms, semantic paraphasias, and repetitions.

Apart from the lexical aspects, other studies focus on the syntactic ones. For instance, Altmann, Kempler & Andersen (2001) report that patients make mistakes on closed-class words. Kemper et al. (1993) and Lyons et al. (1994) observe that the speech

of PADs is also characterized by a general reduction in complexity. Major overt mistakes at the syntactic level are absent. However, patients produce simplified structures: the number of fragments and incomplete sentences increases, while the number of subordinate clauses and the mean length of utterances (MLU) decrease. For instance, Lyons et al. (1994) report that their controls have a MLU of 9.1 words, which drops to 6.4 in PADs with moderate impairment. Similarly, the mean number of clauses per utterance (MCU) drops from 1.4 to 1.0, thus meaning that patients only rarely produce multi-clausal sentences. The number of conjunction is also halved. Overall, patients resort to the use of a limited number of morphosyntactic structures. Kavè & Levy (2003b) report similar results for Hebrew-speaking patients too.

Forbes-MCKay & Venneri (2005) build on studies on spontaneous oral and written production in order to develop a picture description task for the detection of subtle declines in the speech of elderly speakers. It must be pointed out that there exist contrasting opinions on the use of this experimental technique. For instance, Duong, Giroux, Tardif & Ska (2005) argue against the use of parameters based on discourse analysis for the individuation of speakers with probable Alzheimer. Their main argument concerns the high heterogeneity found in the narratives of elderly speakers, which does not allow for the individuation of sharp and consistent differences between PADs and healthy speakers.

The possibility of developing diagnostic tools is not of primary relevance for the present discussion; however, the accuracy and relevance of results from studies on spontaneous speech are. With respect to this, more subtle considerations are needed. For instance, what appears as spared and productive in spontaneous speech, might actually hide some problems. That is the case of pronouns.

I briefly mentioned above that the use of pronouns is quite robust in PADs[9]. Almor et al. (1999) investigate the phenomenon under different perspectives. They start from the observation of spontaneous speech and they find that this is rich in pronouns, in line with what Nicholas et al. (1985) had previously observed (see also Kavè & Levy (2003b) on Hebrew-speaking patients). Authors point out that PADs overuse pronouns in context

[9] The observation concerns English-speaking PADs, but it might not hold true for speakers of other languages, in particular in the case of languages that allow subject drop and do not require the use of expletive pronouns. The speculation is based on results from Bencini et al. (2011) on sentence repetition (see below).

where controls would prefer full lexical expressions: this reveals that their use is inaccurate. In order to shed some light on the issue, authors run two experimental tasks. In the first, they ask patients and controls to listen to a sentence and to read aloud the pronoun that appears onthe screen in correspondence to the last word in the clause. In half of the stimuli, the pronoun on the screen is appropriate, in the other half it is not. Controls react to the two conditions with significant difference, in the sense that their reaction times are shorter in correspondence to appropriate pronouns and prolonged in the case of inappropriate pronouns. In contrast, in PADs the difference between reaction times in the two conditions is very narrow (and only descriptively longer for inappropriate pronouns). Almor et al. (1999) interpret the result by claming that PADs do not have strong expectations on the pronoun to appear on the screen because of poor comprehension of the sentences and lack of a semantic and syntactic analysis. Moreover, as revealed by the second task in the study, PADs are facilitated in comprehension by the presence in the sentence of noun phrases more than by the presence of pronouns. In other words, results from production reveals that patients resort to pronouns whenever retrieving the corresponding lexical item is difficult, while comprehension benefits from full lexical expressions, which facilitate the creation and retention of a mental representation of the sentence.

The study points out that speech analysis can be sometimes misleading with respect to the actual abilities of PADs. The use of rather complex morpho-syntactic elements (as pronouns are) actually hides some specific difficulties, which experimental task can unveil. In other words, the observation of spontaneous speech is not sufficient in order to achieve a comprehensive understanding of the linguistic deficit in PADs. For this reason, experimental tasks that allow for the systematic collection of data should be run. In turn, these offer an insight into the syntactic abilities of PADs.

In what follows, I will review studies that deal with the syntactic competence of PADs. In particular, the review will touch upon the three main issues that research has addressed so far: the correlation between working memory and syntactic processing, the sensitivity to syntactic violations, and the processing of sentences characterized by a non-canonical assignment of *theta*-roles. The last section collects studies that previously dealt with relative clauses: these introduce the reasons for the present study on the comprehension of Wh-questions and RCs (see Chapter 6 and 7).

2.5.1 Working memory and syntactic processing

Working memory is essential in the execution of many tasks, including language processing, both for production and for comprehension. It allows for the temporary storage or relevant information and their manipulation (Baddeley 1986, 1992). The mechanism is impaired in PADs, in particular in its central executive system, while the phonological loop is relatively spared until a moderate stage of the disease (Baddeley et al., 1991; Belleville et al., 1996; Morris, 1994; Morris & Baddeley, 1988).

The possible correlation between working memory impairment and syntactic competence impairment has been explored in a number of studies, which claim that the former is responsible for the latter; i.e. the syntactic impairment would be the reflex of impaired WM.

In the study on pronouns mentioned in the previous section, Almor et al. (1999) observe that patients are not very sensitive to the syntactic and semantic priming of pronouns, such that they react similarly both to appropriate and inappropriate pronouns. Authors ascribe the problem to patients' inability to keep the relevant information active in their working memory for the necessary time-span. While processing the sentence, information activation and integration decay soon, such that PADs cannot achieve a comprehensive understanding of the sentence. This reduces also their chances to activate expectations on how the sentences can plausibly continue. The poor performance on the task is therefore interpreted as the outcome of poor working memory[10].

Patients' poor activation of semantic and syntactic information emerges also in a study conducted by Grossman & Rhee (2001). Authors test PADs in their ability to detect a target word in sentences characterized by semantic or syntactic violations. Whenever the target word immediately follows the violation, PADs are rather fast at detecting it, with no major differences in comparisons to sentences without violations. In contrast, in the condition in which the target word follows the violation at a distance of four syllables, their reaction times are longer. In other words, PADs seem not to notice violation at first,

[10] According to Almor et al. (1999) an effect of working memory impairment emerges also at the semantic level in spontaneous speech. The spontaneous speech sequences they analyse contain pronouns in a number that correlates negatively with working memory. For each patient, lower scores on tasks tackling working memory correspond to a more frequent use of pronouns in the place of words they cannot retrieve.

while a reaction to it emerges later in the processing of the sentence. The authors therefore claim that information activation slows down in PADs.

Evidence of limited working memory resources for linguistic processing emerges also from a study based on sentence repetition. Small, Kemper & Lyons (2000) ask PADs to repeat active and passive simple sentences, and subject and object relative clauses (with both right- and left-branching). Patients' accuracy varies according to the complexity of the structure and along the sentences. They repeat correctly the first part of the experimental item, while mistakes gradually increase in the middle and in the final part of the sentence. Patients therefore show primacy effects, but not recency effects. According to the author, the pattern corresponds to a processing procedure that allocates all resources to the first section, and does not leave resources left for the subsequent strings. The result is that patients do not succeed in completing sentence repetition because of a specific limitation in working memory.

With respect to task demands and working memory, Kempler et al. (1998) point out the importance of evaluating carefully the competence and resources required for the completion of experimental tasks. The authors enrol PADs both in on-line and in off-line tasks and observe that results from the two kinds of studies sharply differ. The off-line task requests sentence-to-picture matching, while the on-line one measures reaction times to syntactic violations. In the latter, patients only slightly differ from controls, while wider asymmetries emerge in the former. In the authors' view, patients are more impaired at executing the off-line matching task because that requires post-interpretative work. After processing and comprehending the sentence, speakers need to keep the structure active in mind while analysing the visual material; the task is complete only once the participant choos the matching picture. Successful completion can take place at this point only in case PADs have managed to keep in mind the sentence meaning for all the time it took to select an image and point at it. Authors explain results under this perspective, thus underlying the role of task demands and the fact that post-interpretative workload can compromise the accuracy of execution. The claim is corroborated by the fact that results from the off-line task correlate with measures of working memory, while results from the on-line task do not. However, authors also admit that the tasks in use in their study differ in the kind of experimental sentences; in particular, those in the on-line task are of the simple active kind and do not entail any specific manipulation, while some sentences in

the off-line tasks are at the passive voice or include an embedded relative clause. Moreover, the on-line task consists in the individuation of a target word and does not require explicit syntactic analyses. The on-line task can indeed be carried out also through a shallow sentence processing, which can stop as soon as the target word is met. Overall, task results depend on a variety of factors, which cannot be reduced to impaired working memory. Nonetheless, the warning in favour of a precise evaluation of task demands in experimental studies is not to be underrated.

The discussion on experimental data correlating with working memory is the primary focus also of a study presented by MacDonald et al. (2001): in this, authors test the efficiency of a digit span task in order to simulate the workload of working memory in language processing. Participants listen to a string of numbers (7, 3, 2, 9) and repeat it in ascending order (2,3,7,9). This task tests working memory and simulates language processing because it requires to hold information active (remembering the number sequence) while manipulating it (reordering the numbers). Results on this task correlates well with results from an off-line task based on grammaticality judgement. However, authors also point out the difficulties at sharply distinguishing working memory from language processing, because the two are strongly bound and the tasks in use do not allow for a clear dissociation. Moreover, they also highlight the fact that a correlation between working memory impairment and language processing impairment does not mean that the former causes the latter. The two are firmly bound such that they can break down at the same pace, without a cause-effect relation between the two.

In sum, many studies focus on working memory as the primary source of syntactic impairment in PADs (Kempler et al., 1998; Small, Kemper & Lyons, 2000). MacDonald et al. (2001) have questioned this position, although clear conclusions cannot be met. In the present study, I will assume that working memory is involved in sentence processing and that its impairment can play a major role in it. The issue cannot be denied. However, I claim it is worthy setting it aside for the moment, in order to address the attention towards strictly syntactic aspects of sentence processing. If patients have different patterns of performance in correspondence of different syntactic conditions, the data can be highly informative on syntactic processing, independently of working memory impairment.

2.5.2 Sensitivity to syntactic violations

In absence of major overt syntactic mistakes in the spontaneous speech, a straightforward way to sample the status of syntactic competence is to verify whether patients are sensitive to syntactic violations. This can be done both automatically and explicitly. On-line tasks detect automatic individuation of grammatical mistakes by measuring reaction times or ERPs in correspondence to violations. Off-line tasks usually require explicit grammatical judgements, such that participants must judge whether the sentence is correct or not. Background literature on sensitivity to grammatical violation in PADs includes studies that imply both techniques.

In the previous section I already mentioned the studies conducted by Kempler et al. (1998) and by MacDonald et al. (2001). Both studies (which are run by the same team of researchers), sample patients' sensitivity to grammatical mistakes that violate either transitivity or subject/verb number agreement. In the on-line task, participants are requested to listen to a sentence and read aloud the last word in the structure. Half of the sentences contains a mistake, either of the number agreement kind or of the transitivity kind. In the former, singular verbs follow plural subjects (*'The children was in the garden') or *vice versa* (*'The child were in the garden'). Sentences with transitivity violations present either transitive verbs followed by indirect objects (*'Susan attended to meeting'), or intransitive verbs followed by direct objects (*'Susan went meetings'). Reaction times at reading the last word in the sentence reveal sensitivity to syntactic violations in both PADs and controls. All participants have longer reaction times in correspondence of sentences including a violation. The data therefore reveal that both PADs and controls are engaged in the process of violation reparation, with no significant statistical difference between the two groups.

Off-line tasks reveal a group effect, with PADs overall less accurate than controls at detecting grammatical violations. All controls perform with percentages of accuracy in-between 88% and 100%. In turn, PADs' accuracy ranges from 58% to 95%, thus revealing a high variability within the group. As mentioned in the previous Section, Kempler et al. (1998) and MacDonald et al. (2001) ascribe the asymmetry to post-interpretative demands set by off-line tasks.

Apart from impaired Working Memory, another on-line study detects slowed-down information processing in PADs. Grossman & Rhee (2001) enrol their patients in a word

detection test. Sentences in the stimuli distribute across three conditions: target words in correct sentences, target words right after the mistake, and target words after the mistake at a distance of four syllables. Mistakes are of the semantic or of the grammatical kind (number misagreement, mass/count discrepancy or negation violations), but the different kinds do not produce relevant asymmetries. Controls have similar reaction times in the first condition (correct sentence) and third condition (the target word follows the mistake by several syllables). This means that they have prolonged reaction times only in the second condition, when the target word immediately follows the mistake. Thus, their sentence reparation takes place as soon as a violation is met. In contrast, PADs perform similarly in the first (correct sentence) and in the second condition (the target word immediately follows the violation), and have prolonged reaction times only in correspondence of target words that follow the mistake by several syllables. The pattern of performance reveals that PADs are slower at activating the reparation operation and engage in it later in sentence processing.

The observation on delayed reparation of syntactic violation is not robust though, as Price & Grossman (2005) find prolonged reaction times to target words that immediately follow a transitivity violation. Thus, doubts are casted on the timing of syntactic reparation, despite clear sensitivity to violations.

As for the asymmetry between on-line and off-line tasks, results are not replicated in a study conducted by Kavè & Levy (2003a). Authors find that patients are sensitive to violations with an accuracy that resembles the one of controls in both kinds of study. Their work is of high interest because it enrols speakers of Hebrew, a language characterized by a rich inflectional system. In the on-line study authors sample patients' sensitivity to gender agreement[11]. In the off-line task they request explicit grammatical judgements for sentences including tense and person violations. In both conditions, PADs perform as accurate as controls, with percentages of target answers that reach above 90%. Moreover, no high variability within the group is attested. The results therefore do not replicate those from Kempler et al. (1998) and MacDonald et al. (2001) with respect to the asymmetry between on-line and off-line tasks and the high variability in the former. The difference in results might be ascribed to demographic differences between the two experimental groups (i.e. different stages of the disease and/or different levels of

[11] In Hebrew, subject and verb agree for gender (Kavè & Levy, 2003).

cognitive impairment). As an alternative account, Kavè & Levy (2003a) mention the fact that speakers of different languages might have different levels of sensitivity in accordance to the richness of the inflectional system of their native language. The hypothesis is appealing, although authors lack stronger evidence in its favour.

The issue of sensitivity to language-specific syntactic constraints is the focus of the study by Bencini et al. (2011). Authors compare English-speaking and Italian-speaking PADs in order to verify whether their repetition abilities interact with the properties of the native language. Italian is a null subject language (Chomsky, 1981; Rizzi, 1982), which allows the covert realization of the subject. In contrast, English grammar does not allow this option. Participants from both groups are requested to repeat aloud sentences characterized by increasing complexity (one-clause vs two-clause sentences) and increasing length. Results report that both groups show the expected pattern: their accuracy is higher in simple and short sentences, and lower in longer sentences. However, Italian-speaking patients tend to contain the complexity of longer sentences by omitting the repetition of subject phrases. The same pattern does not emerge in English-speaking PADs. Authors conclude that patients retain a good knowledge of the syntactic options offered by their native language and are able to resort to it as a strategy for reducing workload.

Overall, results from the reviewed studies speak in favour of a spared syntactic competence. Overt violations do not appear neither in spontaneous speech (Kemper et al., 1993; Lyons et al., 1994; Nicholas et al., 1985, among others) nor in sentence repetition (Bencini et al., 2011); and patients retain sensitivity to grammatical violations (Kavè & Levy, 2003a; Kempler et al., 1998; MacDonald et al., 2001; Price & Grossman, 2005).

2.5.3 Canonicity of theta-role assignment

Despite sensitivity to syntactic violations and absence of overt syntactic mistakes in spontaneous speech, it was noticed that PADs show a generalized simplification of their morpho-syntactic structures (Kemper et al., 1993; Lyons et al., 1994). Based on this observation, studies attempt to look for signs of syntactic deficit in details. Their attention focuses on syntactic manipulations that might entail a certain degree of complexity. Background literature offers a series of studies, which take into consideration two major factors: canonicity of *theta*-role assignment and embedding (especially in the form of

relative clauses). The present section deals with the former, the following section deals with the latter.

In order to verify patients' reliance on the canonicity of *theta*-role assignment, authors take into consideration either passive structures or verbs characterized by peculiar argument grids. For instance, Bates et al. (1985) elicit short narratives containing passive structures. The analysis of the output counts only few passives though, among which, the majority is of the periphrastic *get*-type. Passive is therefore underrepresented in production. However, comprehension of passive seems to be spared. Grossman & White-Devine (1998) ask comprehension questions concerning active and passive sentences, and patients answer accurately on both conditions. No difference between active and passive sentences emerges also in a sentence repetition task run by Small et al. (2000).

Bickel et al. (2000), Waters et al. (1998), and Kemper et al. (1998) test patients' mastery of passive structures through sentence-to-picture matching tasks, and obtain rather sound results. Bickel et al. (2000)[12] do not find any significant difference between active and passive sentences. However, within the latter condition, they notice that reversible passives challenge PADs more than non-reversible passives do. That means that patients' processing of passive structures benefits from the presence in the passive clause of an inanimate referent, which is less probable to be assigned the agent *theta*-role[13].

Waters et al. (1998) gather similar results with a sentence-to-picture matching task and a video judgment task. In the latter, participants must judge whether the stimulus sentence they hear correctly describes the video content. In both tasks, PADs equally master the comprehension of active and passive sentences.

The only study to report an asymmetry between active and passive sentences is the one run by Kempler et al. (1998): their PADs are less accurate on passive sentences than on active ones. However, the asymmetry is reduced in the case of active sentences with two conjoined NPs in the object position (e.g., 'The cow chases the dog and the cat').

[12] Among the reviewed studies, Bickel et al. (2000) is the only one that samples the asymmetry between active and passive sentences in a language different from English, namely German.
[13] Chapter 5 includes a discussion on the reasons why passive structures with inanimate subject are processed with more ease than passive structures with two animate arguments.

Grossman & White-Devine (1998) shift the attention towards another aspect of non-canonical *theta*-role assignments, namely verbs characterized by peculiar thematic grids. In particular, Grossman & White-Devine (1998) analyse the case of causative verbs:

(4) John drowns the swimmer

In (4), the actual agent of the drowning action is the swimmer, as made explicit by the use of the verb in its intransitive form:

(5) The swimmer drowns

The same meaning as in (4) can further be conveyed by a periphrastic causative structure:

(6) John made the swimmer drown

In their study, Grossman & White-Devine (1998) sample patients' comprehension of active sentences with transitive verbs, with causative verbs (4) and with periphrastic causative structures (6). Participants listen to the sentences and then answer Yes/No questions (e.g. 'Did the swimmer drown?' or 'Did John drown?'). Based on patients' answers, authors claim that PADs are spared at processing transitive verbs, while they are impaired at processing lexical causative verbs. However, comprehension of causatives improves in the case of periphrastic structures, which make the peculiar *theta*-role distribution explicit. Authors deduce that the specific difficulties on causative verbs stem from the fact that the Agent role must be assigned to the object in the clause structure.

Manouilidou et al. (2009) further investigate non-canonical distributions of *theta*-roles and their effects on PADs' processing abilities. The object of the study are psychological predicates of the 'fear' and 'frighten' kind:

(7) John fears the storm

(8) The storm frightens John

Both kinds of verbs do not respect the thematic hierarchy, which prescribes the Agent > Experiencer > Theme order. 'Fear' verbs entail an atypical *theta*-role assignment because the first element in the hierarchy (the agent) is not assigned. However, the canonical order is respected as the experiencer (John) precedes the theme (the storm). In contrast, verbs of the 'frighten' kind entail a non-canonical *theta*-role order as the theme (the storm) is assigned prior to the experiencer (John).

Manouilidou et al. (2009) use a sentence completion task, in which the argument structure is provided to participants, along with three alternative verbs. Participants' task is to select the target verb that can complete the sentence in a meaningful way. Overall, PADs perform well on transitive verbs, while they have many uncertainties with contexts requiring psychological verbs (with no major asymmetries between 'fear' and 'frighten' verbs). The authors interpret the result as the manifestation of a specific deficit: PADs are impaired when *theta*-role assignment does not strictly respect the canonical hierarchy.

Taken together, results from the studies reviewed in the present section show that patients do not experience difficulties with thematic roles, when these are canonically realized at the deep structure level. Passive structures with a theme/agent reversed order are not problematic. This is because at the deep structure the two roles are canonically assigned: the agent corresponds to the external argument and the theme to the internal one. At the surface level, the order is reversed but patients can nonetheless process the sentences. In contrast, causative and psychological verbs entail hierarchical violations at the deep structure, which challenges PADs.

2.5.4 Relative clause comprehension

In order to investigate sentence processing in PADs, linguists take into consideration syntactic complexity as a relevant factor. In particular, they aim at verifying whether the performance of PADs decreases in correlation with factors of syntactic complexity. Two structures have often be implied in order to test the effects of syntactic complexity: passive voice and relative clauses. I considered the former in the previous section; the latter will be the object of the present one.

Background literature provides at least four studies that include relative clauses among their experimental conditions: Bickel et al. (2000); Kempler et al. (1998); Small, Kemper & Lyons (2000); and Waters, Rochon & Caplan (1998). Those works have

already been reviewed in the present study because of their contribution to the investigation of the computation of passive. In what follows, I will therefore review the papers again with a specific focus on relative clauses, in order to summarize what is known so far about the computation of relative clauses in PADs, and what the results can tell us about the syntactic competence in Alzheimer's disease.

Waters, Rochon & Caplan (1998) deny the presence of a syntactic deficit in PADs. However, in my opinion, the reasons for this strong claim stem from authors' definition of syntactic complexity and the design of the experimental conditions in use. Waters, Rochon & Caplan (1998) engage participants in a sentence-to-picture matching task, which includes nine different conditions: active sentences (9a), active sentences with a conjoined DP in the object position (9b), active sentences with a direct object and a prepositional phrase (9c), passive sentences (9d), truncated passive sentences (9e), objet clefts (9f), right-branching subject relative clauses (9g), two conjoined sentences (9h) and left-branching object relatives (9i)[14]:

(9) a. The lion kicked the elephant

b. The pig chased the lion and the cow

c. The elephant pulled the dog to the horse

d. The elephant was pushed by the cow

e. The pig was touched

f. It was the dog that the horse passed

g. The horse kicked the elephant that touched the dog

h. The elephant followed the lion and pulled the dog

[14] Examples are reported from Waters et al. (1998: 369).

i. The dog that the pig followed touched the horse

Patients repeat the test several times in slightly different conditions, which vary according to the supplied visual material. They see either two or three pictures and must point to the one that correctly represents the action described by the experimental sentence. In two different versions of the test, patients see either one picture or one short video and must decide whether these correctly depict the sentence meaning. All versions of the test provide the same results: with respect to the baseline condition (9a), patients are significantly less accurate in the comprehension of sentences that include two conjoined clauses (9h) and in sentences with a relative clause (in (9g) and (9i)).

Authors consider three factors for the interpretation of the results. These are canonicity of *theta*-role assignment, number of arguments in the clause and number of predicates. The first two factors represent syntactic complexity and their manipulation allow the distinction between active and passive sentences, and between sentences with two arguments (subject and object as in (9a)) or with three arguments (one subject and two objects as in (9b)), or subject, object and indirect object as in (9c)). The third factor, namely the number of predicates in the clause, is considered only as a factor for sentence length, but not for syntactic complexity. Authors intend to use it as a test for working memory, thus assuming that two predicates imply longer sentences and therefore higher processing workload for working memory. Based on these premises, Waters et al. (1998) conclude that PADs deal well with syntactic complexity because they show a good comprehension both of active and passive sentences, both with two or three arguments. However, patients' poor comprehension on relative clauses is interpreted as the reflex of sentence length and it is ascribed to poor working memory. In their discussion, authors completely disregard the syntactic derivation that characterizes relative clauses and the asymmetry between matrix and embedded clauses. They only evaluate sentence length as a predictor for sentence comprehension: in their view, sentences including relative clauses are more difficult to comprehend because they are usually longer than simple sentences. Contra to authors' definition of syntactic complexity, I will claim that relativization is precisely a factor of syntactic complexity. In my view, the comprehension of relative clauses posits some difficulties, which stem from the syntactic derivation *per se*, rather

than from the number of words in the sentence (see Chapter 5 on the derivation of Relative Clauses).

Unfortunately, the only two relatives in use are a right-branching subject relative and a left-branching object relative. The two conditions in use differ in two factors, namely relativization site (subject vs object relative clauses) and direction of branching (left- vs right-branching clauses). The study therefore lacks minimal pairs for the evaluation of the single components, such that it is not possible to compare subject and object relatives and left- and right-branching as separate factors. Based on the experimental material and the results at hand, the only information we can infer from Waters et al. (1998) is that PADs are more impaired at comprehending relative clauses than at comprehending simple sentences[15].

Kempler et al. (1998) also investigate patients' ability with a sentence to picture matching task. Authors aim at evaluating two factors by including four kinds of sentences. The four experimental conditions comprehend active sentences (10a), active sentences with two conjoined DPs in the object position (10b), passive sentences (10c) and relative clauses (10d)[16]:

(10) a. The boy pushes the girl

b. The boy scratches the dog and the cat

c. The boy is kissed by the girl

d. The dog chases the girl that chases the boy

Kempler et al. (1998) design the experimental conditions above in order to observe two factors: syntactic complexity (simple active vs passive, and simple active vs relative clauses) and the number of arguments (two in (10a) and (10b) vs three in (10b) and (10d)).

[15] The study also reveals that the number of arguments in the experimental sentences and the number of pictures displayed in the visual stimuli have an impact on patients' successful comprehension. However, the two issues are not of primary interest for the present discussion and therefore will not be further addressed in this section.

[16] Examples are reported from Kempler et al. (1998: 301).

Results report the highest level of accuracy on simple active sentences and the lowest one on relative clauses. Intermediate levels of successful comprehension correspond to the two remaining conditions, with no significant differences between them. According to the described pattern of performance, authors conclude that both considered factors are relevant and influence patients' rate of comprehension. Relative clauses are more difficult than simple active sentences because they are complex from the syntactic point of view; moreover, sentences with three arguments are more difficult than sentences with two arguments. Unfortunately, the experimental material does not allow the evaluation of fine-grained syntactic factors within relative clauses. For instance, authors report to have included subject relatives both of the centre-embedded and of the right-branching kind in their experimental material, but then they collapse the two kinds into one single condition for data analysis.

Bickel et al. (2000) attempt to evaluate different factors internal to relative clauses in their sentence to picture matching task performed by German-speaking PADs. Their study includes fourteen different sentences, among which there are four kinds of relative clauses: centre-embedded subject relatives (11a), centre-embedded object relatives (11b), centre-embedded subject relatives with intransitive verbs (11c), and right-branching subject relatives (11d)[17]:

(11) a. *Das Abendkleid, das auf dem Anzug liegt, ist aus blauer Seide*
'The robe that lies on the suit is made up of blue silk.'

b. *Das Abendkleid, auf dem der Anzug liegt, ist aus blauer Seide*
'The robe, on which the suit is lying, is made up of blue silk. '

c. *Der Junge, der im Auto sitzt, unterhält sich mit dem Mann*
'The boy, who is sitting in the car, is talking with the man.'

d. *Der Junge unterhält sich mit dem Mann, der im Auto sitzt*
'The boy is talking with the man sitting in the car.'

[17] Examples are adapted from Bickel et al. (2000: 435).

Bickel et al. (2000) subdivide their patients into two groups according to their level of dementia. Patients who score above 20/30 in the Mini Mental-State Examination Test (Folstein et al., 1975)[18] are classified as mildly impaired. Participants with scores lower than 20 on the same test form the group of patients with moderate impairment. Both groups show a preference for subject relatives over object relatives. In particular, centre-embedded object relatives correspond to the condition with the lowest level of accuracy (at-chance performance). In contrast, comprehension of centre-embedded subject relatives is good: moderately impaired PADs perform above chance level and mildly impaired PADs do not differ from controls. What actually surprises in the results is the fact that within the subject relative condition, PADs perform better on centre-embedded than on right-branching clauses. The result is rather unattended from the syntactic point of view because it seems to indicate that the processing of matrix clauses benefits from the interposition of a centre-embedded clause (thus leaving the initial NP stored in the working memory without an assigned role in the matrix clause). Bickel et al. (2000) interpret the result as an effect of task performance. In the case of centre-embedded relative clauses, the task can be completed in a shorter time because the information necessary for the picture selection is provided by the embedded clauses. Participants can therefore perform the task without completing the processing of the matrix clauses. In contrast, right-branching relative clauses require the complete processing of both the matrix and the subordinate clause prior to the selection of the target image. Based on this observation, Bickel et al. (2000) also resort to limited working memory in order to explain the pattern of performance.

Overall, the main result of the reviewed study consists in the differences in performance between mild and moderate patients. The first group performs similarly to controls, while the second one performs poorly. I will try to examine in depth the issue in Chapter 6 and 7, in order to achieve a better understanding of how patients' syntactic competence changes along with the worsening of the disease.

The fourth study I am going to review in this section represents a further attempt to disentangle two factors of syntactic manipulation in relative clauses, in order to comprehend their role in sentence processing. Small, Kemper & Lyons (2000) enrol their

[18] The Mini Mental State Examination Test (Folstein et al., 1975) is a test for the evaluation of cognitive impairment. Section 4.2.2 contains more information on it.

patients in a sentence repetition task. The factors at stake for relatives are canonicity of *theta*-role assignment (subject vs object relatives) and branching direction (centre-embedded vs right-branching relatives). Both factors are crossed twice in order to obtain four conditions: centre-embedded subject relatives (12c), centre-embedded object relatives (12d), right-branching subject relatives (12e) and right-branching object relatives (12f). The task also includes simple active sentences (12a) and passive structures (12b)[19]:

(12) a. The circus at the convention centre attracted thousands of children

b. The motorist was stopped by the highway patrol

c. The tornado that swept through the town destroyed several homes

d. The job that the woman wanted required a college education

e. The operator assisted the merchant who made a long distance call

f. The angry parent disciplined the student who the teacher sent home

Small et al. (2000) calculate percentages of accurate repetition for each condition and find an unattended pattern of performance. PADs repeat passive sentences and right-branching subject relatives better than simple active sentences. In turn, simple active sentences are repeated more accurately than the three other conditions with a relative clause. Such a poor repetition of simple active sentences is most probably due to the prepositional phrase, which modifies the subject DP and delays verb processing. Moreover, the PP increases the mean length of simple active sentences, a rather relevant factor for a sentence repetition task.

As for relative clauses, authors observe that subject relatives are usually better processed than object relatives. However, the facilitating effect disappears in case of centre-embedding. PADs usually prefer right-branching over centre-embedding, but the

[19] Examples are reported from Small et al. (2000:236).

asymmetry is nullified in case of object relatives. The interpretation of results is made difficult by the crossing of the two factors, which seem to interact but not in an additive way. If the two factors played an additive effect, results would report the lowest accuracy on centre-embedded object relatives, but this is not the case (the poorest performance is actually registered on centre-embedded subject relatives). Overall, Small et al. (2000) fail at clearly disentangling the effect brought in by the two factors. They observe recency effects in patients' performance as these accurately repeat the first part of the sentence, while they fail at repeating the final one. The discussion therefore brings into play once again working memory effects, but lacks the analysis of strictly syntactic factors. Authors ascribe the pattern of performance to limited resource capacity. When the request for information activation exceeds the available pool of resources, the task can be executed only partially.

Altogether, the reviewed studies point out a specific impairment in the processing of relative clauses in comparison to simple active sentences. The observation immediately turns relative clauses into an interesting point of observation to gain insights into the syntactic competence of PADs. However, the studies mentioned above fail at offering the opportunity to analyse the impairment in details. This is partially due to characteristics of the experimental designs in use. Conditions often lack minimal pairs for the evaluation of specific factor manipulations (Kempler et al., 1998; Waters et al., 1998). For instance, in Kempler et al. (1998) centre-embedded and right-branching relatives are collapsed into one condition although they are also characterized by different extraction sites. In the latter case, right-branching SRs are compared to centre-embedded ORs, but the analysis is made difficult by the impossibility of disentangling the two factors (extraction site and branching direction). Moreover, it is not always clear whether authors control for potential confounding factors like, for instance, animacy and word frequency. Animacy (and therefore sentence reversibility) is crucial for the computation of passive sentences and of ORs (see Chapter 5 on this). Word frequency cannot be overlooked in studies with patients that suffer from severe anomia, because low-frequency words might compromise sentence comprehension and the results. For instance, Small et al. (2000) report the full list of sentences they use: each sentence contains different lexical items, such that the differences between active sentences and relatives might be due to lexical items rather than different syntactic derivations. Finally, the mentioned studies do not provide a

syntactic analysis for the experimental sentences in use. The lack of a theoretical framework impoverishes both the task design and the result interpretation. For all these reasons, further investigation into the processing of relative clauses in PADs is needed.

2.6 Conclusions

As summarized in the first part of the present chapter, dementia of the Alzheimer's type is a neurodegenerative disease, which affects millions of elderlies. It is determined by major changes in the brain, which consist of neural loss, neurofibrillary tangles and plaques (Braak & Del Tredici; 2006; Hyam, 2006).

A variety of symptoms characterizes the disease: memory loss, apraxia, agnosia, anomia, behavioural disorders and psychiatric disorders are the most common ones (Spinnler, 1996). Usually, patients do not experience all of them at the same time. Different symptoms can appear in different stages of the disease and their incidence can vary. However, the overall impairment progressively worsens to the point that PADs lose their autonomy in everyday life.

As for the linguistic impairment, the main deficit is anomia. Most authors now agree in ascribing it to a gradual loss of semantic knowledge, in association with retrieval difficulties (Chertkow & Bub 1990; Cuetos et al., 2015). The hypothesis of a disruption of semantic knowledge as the main cause to anomia has been supported by three kinds of evidence. First, item-by-item analyses show that PADs tend to perform consistently on single words across tasks: either they perform always correctly or they perform always wrong on the same word in different tests (Chertkow et al., 1989; Chertkow & Bub, 1990). Second, their ability to name objects correlates with the amount and the quality of information that they retain about the objects (Hodges et al., 1996; Ralph et al., 1997). Third, the analysis of errors in naming tasks reveals a progressive pattern of disruption at the semantic level: the main error type shifts from with-in category errors, to superordinate labels and finally to nonresponses (Almor et al., 2009; Hodges et al., 1991). In this process of lexical disruption, frequency (Cuetos et al., 2005, 2015; Forbes-MacKay et al., 2005; Tippett et al., 2007), age of acquisition (Cuetos et al., 2005, 2012; Silveri et al., 2002), and grammatical category (Almor et al., 2009; Grossman et al., 1996; among others) seem to be good predictors for the words that keep spared longer. PADs perform

better with high-frequency and early-acquired words; and verbs are more impaired than nouns.

With respect to verb retrieval and use in PADs, Kim & Thompson (2003) observe that PADs perform equally on verbs characterized by different numbers of arguments. Moreover, Grossman et al. (2007) notice that PADs can acquire syntactic information about new words but not their meaning. Taken together, the two studies suggest that there might be a dissociation between the syntactic and the semantic information in the lexicon of PADs. The idea is very appealing; however, the collected data are poor (and not very consistent), and robust conclusions cannot be met. For this reason, further research on the issue is needed.

Lexical disruption also affects irregular verbal morphology, while regular morphology is usually spared. This supports the hypothesis put forth by Ullman (2001) of a dissociation between declarative knowledge and procedural mechanism: the former is impaired, while the latter is relatively spared in PADs (Colombo et al., 2009; Walenski et al., 2009).

The last part of the chapter contains an overview of studies on sentence comprehension. So far, two factors have been taken into account for impaired comprehension: impaired Working Memory (Baddeley et al., 1991; Kempler et al., 1998; Small et al., 2000) and the effects of non-canonical assignments of *theta*-roles (Grossman & White-Devine, 1998; Manouilidou et al., 2009). Some studies also point out an impaired comprehension of relative clauses, with respect to main clauses (Bickel et al., 2000; Kempler et al., 1998; Small et al., 2000; Waters et al., 1998).

Unfortunately, the tasks in use contained some confounding factors and did not allow for more subtle observations. For this reasons, as pointed out at the end of section 2.5.4, more investigation is needed with respect to the syntactic deficit of PADs.

3 GRAMMATICAL GENDER RETRIEVAL

3.0 Introduction

It is generally agreed upon that grammatical gender (GG) is a linguistic phenomenon that plays a major role at the lexical level and at the syntactic one (Cacciari and Cubelli, 2003; Caramazza and Miozzo, 1997; Carstens, 2000; Friedmann and Biran, 2003; a.m.o.): information on GG is independently stored in the lexicon[20], attached to the lexical item, and is required for syntactic agreement.

GG retrieval is described as a phenomenon that can be classified under either the declarative or the procedural components of linguistic faculty, depending on how, and for which purpose, the lexicon is accessed (Goodglass, 2000). For example, speakers show procedural knowledge of GG whenever they automatically and accurately use gender agreement in a phrase; in contrast, declarative knowledge is retained in conscious metalinguistic judgement on gender class. Under a different perspective, grammatical gender is procedurally retrieved whenever this can be inferred from the word form; while in case this is not possible, speakers must resort to their declarative knowledge about grammatical gender (see below in this chapter). Previous research on the linguistic competence of PADs has shown hints of a possible dissociation between procedural and declarative components, with a better preservation of the former in comparison to the latter, much more likely to undergo erosion (cf. Colombo et al. (2009), and Walenski et al. (2009) on the asymmetry between regular and irregular verbal morphology). For the present research on PADs, GG is then particularly interesting in consideration of its transversal function and relevance. If the hypothesis of a dissociation in PADs between procedural and declarative knowledge is correct, as well as the premises on grammatical gender are, I expect to find a clear pattern of performance in PADs, with a good performance whenever they can rely on procedural retrieval, in comparison to cases in which declarative knowledge about GG is required.

[20] Exceptions to this analysis will be discussed in what follows.

To the writer's knowledge, the number of studies that have questioned the status of grammatical gender in Italian-speaking PADs is very narrow (Manenti et al., 2004; Paganelli et al., 2003). I will discuss thoroughly the results of those studies in the next sections; for the moment it is sufficient to mention that in one study, PADs have a certain sensitivity to GG (Manenti et al., 2004), both in the form of facilitation and of interference in a semantic and gender priming task. In the other study, PADs fail at providing proof of gender retention in semantic errors (Paganelli et al., 2003), thus suggesting an impairment at the level of GG activation in the process of lexical retrieval.

The present chapter is organized as follow: in 3.1 I present some relevant information on grammatical gender with the purpose of highlighting how this morphosyntactic phenomenon can bare different characteristics across languages. As a consequence of this first observation, I would like to remark that any conclusions reached in the present work about GG only concern the Italian linguistic system and therefore should not be extended to other languages. In 3.2 I illustrates the Italian GG system in order to make the reader familiar with it. Section 3.3 deals with the role of GG at the syntactic level. The process of GG retrieval is illustrated in 3.4. In 3.5 I review experimental studies that unveiled the vulnerability of GG in bilingual as well as in aphasic speakers. Finally, previous studies on the status of GG in PADs are summarized in 3.6. In 3.7 I recap the relevant information for the reasons behind the experimental task I present in Chapter 4.

3.1 Grammatical gender

Following Comrie (1999) and Matthews (1997), grammatical gender can be defined as a system in which each noun is assigned to a class; the assignation to a class depends on the form that other elements must take when they are syntactically related to the noun at stake. For instance, a noun is assigned to the masculine grammatical gender if determiners and adjectives take the masculine form when they combine with it. The number of grammatical genders in the system varies across languages (see below).

Different lexical, syntactic, morphological and phonological constraints (e.g., whether gender triggers syntactic agreement) shape how GG is implemented in languages. It follows that the phenomenon is not homogeneous, even among close-related languages. Indeed, a brief look at few Indo-European languages reveals that GG is not a

uniform phenomenon, but it rather assumes different characteristics depending on the specific language: English only maintains residuals of GG, Italian has two genders (masculine and feminine), just like Danish (neuter and common), while German nouns distribute across three grammatical genders (masculine, feminine and neuter). Many non-Indo-European languages lack GG at all, among these Chinese, Turkish and Japanese. It is evident that GG must be treated as a language-specific phenomenon and that it is necessary to be very careful with generalizations.

According to Corbett (2006) and Goodglass (2000), a language is endowed with gender when there are recognizable families of words that share common agreement patterns. Noun distribution across families or categories is often based on commonalities, such as animacy or biological gender. Yet, these commonalities are not rigid constraints to gender assignment and represent mere tendencies, i.e. characteristics shared by a considerable number of nouns endowed with the same gender.

Comrie (1999) also warns from confusing grammatical gender with natural gender. The first one is a linguistic phenomenon, while the second explicitly refers to biological sex and it is related to the animate feature that some nouns acquire from their referents.

One crucial characteristic of GG is that there is no constraint to how grammatical gender is assigned; i.e. the relationship between a noun and its grammatical gender is arbitrary. In the case of nouns that lack natural gender, arbitrariness does not come as a surprise; indeed, an element can be named by nouns endowed by different genders in different languages. A good example is the English word *blood*, which happens to be masculine in Italian ('sangue.M'), feminine in Spanish ('sangre.F') and neuter in German ('blut.N'). Cases of lexical items endowed with natural gender as part of their semantic value but with contrasting information with respect to GG, further confirm the fundamental arbitrariness that contradistinguishes gender assignment. An example of this last described phenomenon is the German word *Weib* ('woman.N'): the neuter GG clearly clashes with the information on feminine natural gender.

In the next sections, I will try to point out the main characteristics of GG with respect to Italian, the language spoken by the PADs involved in the present study; occasional references to other language will be used for the same purpose, i.e. a better understanding of how GG works.

3.2 Grammatical gender in Italian: an overview

Along its diachronic change, Italian simplified the threefold Latin system by dissolving the neuter gender and distributing nouns across the two classes left, namely masculine and feminine (Chini, 1995). Still, the simplification did not lead to a completely transparent system, as various declension classes are still in use (although only few of them are still productive). A declension class is defined (Comrie, 1999) as a set of words assuming the same morphological forms, for instance for the realization of the singular and the plural.

Chini (1995) pointed out that Italian nouns subdivide into seven different declension patterns, which differ according to the inflectional characteristics they take. The repartition into seven classes derives from the five declensions across which Latin nouns distributed. However, major changes took place (Rohlfs, 1968, § 350) and it is now difficult to individuate a strict correspondence between the Latin and the modern Italian system (except for the first and second Latin declensions, which correspond to Class I and II in the system of modern Italian).

Table 3.1 shows the three main declension classes; in total, they cover the 91.2% of Italian nouns (Chini, 1998) and account for the first-hand impression of a language provided with a rather simple system (although this is not completely true, as I will show).

Table 3.1. Italian declension classes I, II and III (Based on Chini, 1995).

Class	Ending sg.	Ending pl.	Gender	Example	Translation
I	-o	-i	M	libro/libri	book/books
II	-a	-e	F	carta/carte	paper/papers
III	-e	-i	M	cane/cani	dog/dogs
			F	ape/api	bee/bees

Class I and II are homogenous for what concerns gender, as they either include only masculine nouns (class I) or only feminine ones (class II). Class III includes both masculine and feminine nouns sharing the same inflectional properties, namely the *-e* ending for singular and the *-i* ending for plural.

Table 3.2 presents the remaining four Italian declension classes; the reason for presenting them apart is that they include a much lower number of nouns (7,6% of Italian

nouns, according to Chini, 1998) in comparison to Classes I, II and III; in particular, class VII is very narrow.

Table 3.2. Italian declension classes IV, V, VI and VII (based on Chini, 1995).

Class	Ending sg.	Ending pl.	Gender	Example	Translation
IV	Various	[=sg.]	M	re/re	king/kings
			F	città/città	city/cities
V	-a	-i	M	problema/problemi	problem/problems
VI	-o M	-i M/-a F	M/F	uovo/uova	egg/eggs
VII	-o	-i	F	mano/mani	hand/hands

Class IV includes monosyllabic nouns (e.g. *re* 'king.M'), oxytones[21] (e.g. *città* 'city.F') and loanwords in use in Italian (e.g. *bar*, *computer*), which are all invariable, as their plural forms are syncretic to the singular ones and no morphemic change is to be detected in accordance with number change[22].

Most words in Class V derive from Ancient Greek[23] and maintained a -(*m*)*a* ending (*problema* 'problem.M', *teorema* 'therorem.M', *monarca* 'monarch.M') for singular masculine nouns, although the plural regularizes in *-i* (*problemi* 'problems.M', *teoremi* 'theorems.M', *monarchi* 'monarchs.M'). As pointed out by De Martino et al. (2011), nouns included in Class V can provide with interesting insights into gender retrieval processes due to the fact that their gender suffix *-a* (otherwise typical for feminine words in the sizable Class II) is inconsistent with the assigned syntactic grammatical gender, thus generating a sort of information clash. I will come back to this issue later in the discussion.

What is special to Class VI is plural. Words included in this class have a *-o* suffix in the singular, which corresponds to masculine GG; however, their plural turns feminine with an *–a* suffix (*uovo*/*uova* 'egg.M'/'eggs.F'). Next to the feminine plural, some items are also equipped with a masculine plural in *-i*; in that case, the two plural forms specialized for different meanings, according to generalized semantic oppositions: “abstract” vs. “concrete” or “common” vs. “individual” (e.g. *muro*/*mura*/*muri*

[21] Some of the words in this class (e.g. virtù 'virtue') once belonged to Class III in Old Italian because of their ending in *-e* (*virtude* 'virtue'), but moved to Class IV after losing the last syllable. Nowadays speakers do not retain knowledge about this process.
[22] In Class IV, information about Number (singular/plural) is to be retrieved from the elements within the DP that agree with the noun, namely determiners, adjectives, etc..
[23] A few nouns in this class come from Latin, though (e.g. *poeta* 'poet.M', *transfuga* 'deserter.M').

'wall.M'/'walls(around a city).F'/'walls(of a room).M'). Plural forms in *-a* are visible residuals of Latin neuter gender (whose plural was precisely marked by an *-a* suffix), but still it is not possible to talk about neuter gender in Italian for at least two reasons (Chini, 1995). First, the declension pattern is not productive anymore and, second, it has no syntactic counterpart, in the sense that elements agreeing with the N in the DP take either masculine suffixes (in the singular) or feminine ones (in the plural), but there is no specific agreement pattern pairing neuter. Rohls (1968, §419) claimed that Southern Italian dialects retain manifestations of neuter gender, in particular with mass nouns referring to goods or substances (e.g., wine, salt, iron, etc.), but this is certainly not the case for standard Italian, the variety at stake here. In the Southern varieties, mass nouns pair with neuter determiners and this allows for the individuation of a neuter GG, although its use is very limited.

Despite the presence of only two genders in Italian, the picture is quite puzzling in what concerns the mapping of GG into morphological markers. For instance, one single marker can be associated to different genders (*-e* can mark masculine and feminine singular nouns, as well as feminine plural nouns) and a gender may enlist different markers (singular masculine can be marked with *-o*, *-e* or *-a*). As a result, univocal relations between GG and morphological markers are not to be found.

In the light of the seven declension classes described above and the very uneven distribution of nouns across them (Chini, 1998), it is now evident why the Italian GG system should be considered as a mixed one. Some of its components, in particular Class I and II, are transparent; indeed, the ending suffixes provide clear information about GG: masculine nouns typically end with an *-o*; while feminine nouns are recognizable through their *-a* ending. Nonetheless, a considerable number of nouns is opaque, in the sense that suffixes do not provide any disambiguating information about gender, as in the case of Class III and IV, which include both masculine and feminine words. Last, nouns included in Class V, VI and VII can be labelled as irregularities, as their patterns are very mixed and only apply to a restricted number of lexical items. The description is supported by statistical data: Chini (1998) verifies the distribution of Italian nouns through the different declensions classes and reports that 71.5% of Italian nouns (Class I and II) is transparent with respect to the gender information encoded in the word final suffix. Another 26% consists of opaque nouns (Class III and IV), carrying no disambiguating morphology

towards either masculine of feminine; while a remaining 2.2% of Italian nouns is to be considered as irregular.

As described above, the relationship between a noun and its gender is arbitrary in principle. Nonetheless, it is possible to recognize some tendencies in gender assignment in Italian: for example, semantics play a role. GG very often pairs natural gender, especially in the case of lemmas carrying a [+human] feature (e.g. *uomo*/*donna* 'man.N'/'woman.F', *padre*/*madre* 'father.M'/'mother.F'). Moreover, GG comes into help in order to refer explicitly to natural gender in word pairs like *gatto*/*gatta* ('cat.M'/ 'cat.F') or *ragazzo*/*ragazza* ('boy'/ 'girl'). Chini (1995) also pointed out that correlations emerge between gender and semantic fields; here a few examples: metals and trees are often named with masculine nouns (e.g. *ferro* 'iron.M', *oro* 'gold.M', *acero* 'maple.M', *abete* 'fir.M'). Feminine names are in use for continents (*Europe* 'Europe.F', *Asia* 'asia.F'), disciplines (*linguistica* 'linguistics.F', *biologia* 'biology.F') and fruits (*mela* 'apple.F', *fragola* 'strawberry.F'). Still, the examples only represent tendencies, as counter-example are also available: e.g., *vite* ('vine.F') is feminine and not masculine, differently from most tree names in Italian; and *diritto* ('jurisprudence.M') is a counter-example to feminine names for disciplines.

When it comes to gender assignment for nouns resulting from a derivational process, Italian often follows a transparent pattern (Chini, 1998; Luzzatti, De Bleser, 1996; Scalise, 1994). This is due to the fact that derivation is a lexical process that allows the transformation of a lexical stem into a new word. As a result of this process, a new syntactic category applies to the element. Nominalization implies the assignation of GG to a newly formed word, an operation which is constrained by the suffix in use. Indeed, derived Italian nouns inherit GG from the suffix. This process is allowed by the fact that suffixes result in being the grammatical head of the new form (Scalise, 1994; Luzzatti, De Bleser, 1996) and therefore they transfer their syntactic features to the derived formed word through percolation. Italian can count on a number of derivational morphemes endowed with either masculine or feminine GG. Examples of suffixes carrying masculine GG are -*iere* and –(*t*)*ore*; while -*aggine* and -*udine* stand for feminine gender (Scalise, 1994):

(1) Masculine derived nouns: *cantiere* ('construction site.M'), *quartiere* ('neighborhood.M'), *malore* ('sudden illness.M'), *motore* ('motor.M').

(2) Feminine derived nouns: *stupidaggine* ('nonsense.F'), *sfacciataggine* ('insolence.F'), *solitudine* ('loneliness.F'), *latitudine* ('latitude.F').

Derived nouns in the examples are to be ascribed to Class III because of their *-e* ending. Yet, it is legitimate to wonder whether the opaqueness of words marked with *-e* is reduced in case the final vowel is actually part of a derivational suffix carrying a clear disambiguating cue towards either masculine (*-iere*, *-ore*) or feminine gender (*-aggine*, *-udine*). In other words, the question is whether speakers perceive words like *malore* and *solitudine* differently from *cane* and *ape* and are facilitated at gender retrieval by suffixes. Luzzatti and De Bleser (1996) already addressed this question and actually found that aphasic speakers are more proficient at retrieving GG for derived words, rather than for Class III words. Results from the quoted study were highly influential for the design of the present work and therefore will be presented in details in Section 3.5.1.

3.3 Grammatical gender at the syntactic level

In Italian, GG plays an important role at the syntactic level as it triggers agreement both within and outside the DP. Gender agreement affects different elements: definite and indefinite determiners, pronouns (both personal pronouns and object clitics), adjectives in a modifier position (3) or as predicates; also (some) quantifiers are marked for gender (4). At the syntactic level, Italian has gender agreement on past participles with the auxiliary *essere* ('be') and on past participles with the auxiliary *avere* ('have') when the object is realized as a clitic pronoun (5). In the former case (with *essere*), participles agree in gender and number with the grammatical subject in the sentence; in the latter (with *avere*), agreement is triggered by the presence of an object clitic, thus the participle agrees with the object, rather than with the subject in the clause.

(3) *La nuova casa*
the.F new.F house.F

(4) *Molte case*
Many.F.Pl houses.F.Pl

(5) *La casa l'ho affittata subito*
The.F house.F clit.F rented.F immediately

Moreover, Longobardi (1994) claimed first that bare nouns are not licit arguments in the syntax of Italian language. This means that every time a noun is retrieved from the lexicon and is introduced in a syntactic derivation, a full DP is necessarily constructed. The assumption is very relevant here: the construction of a DP automatically calls into play GG as definite and indefinite articles are both marked for gender and number. In other words, for languages like Italian, in which bare arguments are not admitted in the derivation, access to grammatical gender information takes place automatically, any time a noun joins the syntactic derivation.

Friedmann and Biran (2003) capitalize on Longobardi (1994) and suggest that the lack of bare nouns in syntactic structures is a crucial factor, capable of determining important differences among languages with respect to GG retrieval. In particular, the authors suggest the idea that speakers of languages in which bare nouns are admitted in the syntactic derivation, e.g. Hebrew, activate the information about GG only if this is requested by the syntactic context in which the noun is inserted, otherwise GG remains silent and non-active. In contrast, languages of the Italian type always require the activation of information about GG, as any time a noun is retrieved from the lexicon, a full DP is constructed, independently of the contingent syntactic conditions. According to Friedmann and Biran (2003), constraints on the use of bare nouns are able to shape speakers performance even in single word tasks: speakers of Hebrew, for example, show signs of gender preservation neither in cases of paraphasias nor in tip-of-the-tongue states[24] (TOT). In contrast, speakers of Italian show clear signs of correct gender activation despite paraphasias and TOTs (Caramazza and Miozzo, 1997; Vigliocco et al., 1997). Friedmann and Biran (2003) explain the different behaviours across languages as a by-product of differences concerning bare nouns: speakers of languages in which bare

[24] The Tip-of-the-tongue phenomenon is a state in which the speaker fails at successfully retrieving the desired word and has access only to partial information; he or she also has a feeling of imminent activation.

nouns are not allowed (i.e. Italian) overuse DPs, which means that they always automatically retrieve gender, even in single word tasks (e.g. picture naming tasks). Moreover, gender retrieval is such a core characteristics of lexicon processing in Italian, that it takes place even despite incomplete activation of words. In contrast, no gender retrieval is entailed in single word tasks for speakers who admit bare nouns in their language (e.g. Hebrew). In the view of Friedmann and Biran (2003), automaticity of gender retrieval can therefore be considered as a language specific phenomenon, constrained by syntactic factors: syntax plays a crucial role in shaping the cognitive and linguistic process of lexical retrieval, distinguishing between languages that automatically perform gender retrieval whenever a noun is selected and others that do not. According to Friedmann and Biran (2003), Italian belongs to the former group.

This view is only partially supported by Paganelli at al. (2003): their unimpaired speakers of Italian show gender congruency effects in picture naming only if the task explicitly requests participants to name objects by providing a full DP. In contrast, in the condition in which experimenters request participants to name objects by speaking aloud only bare nouns, no such effect can be detected, in the sense that the non-target words provided by participants very often have a different gender with respect to the target word. In other words, on the one hand Paganelli et al. (2003) find convergent evidence of the crucial difference played by DPs versus bare nouns with respect to gender retrieval: the former requires automatic activation of GG, the latter does not. On the other hand, their results speak against the hypothesis of language specific effects: the performance of unimpaired speakers of Italian is characterized both by gender congruence and incongruence, depending on the experimental conditions in use (DPs vs. bare nouns). Syntax plays a role then, but only with respect to the specific conditions and not as a language specific phenomenon. The typological difference pointed out by Friedmann and Biran (2003) between languages that admit bare nouns and others that do not is therefore questioned, while the relevance of DP construction for GG retrieval is further supported.

When dealing with grammatical gender in syntax, it is necessary to address the issue concerning the presence of a functional projection specifically dedicated to the realization of grammatical gender. Bernstein (1993) proposed that languages that have overt morphological markers for gender according to different inflectional classes, also dispose of a dedicated functional projection in the DP. According to the author, the presence of a

GenderP is parametrized. In the author's view, languages that set a positive value, do not only manifest overt gender markers, but also admit syntactic phenomena like noun ellipsis and require head noun movement towards a higher position in the DP functional domain (resulting in the Noun/Adjective order, rather than in the Adjective/Noun one).

Similarly, Picallo (1991, 2008) posits that GG is a functional element represented by a feature [CLASS] in a dedicated functional projection, the same in use for the realization of noun classifiers in Mayan languages and for noun class markers in languages of the Bantu type. Picallo (1991) also assumes that the noun morphology of Romance languages like Spanish and Catalan mirror the cyclic upwards movement of the noun. Given a word like the Spanish *muchachas* ('girls.F.PL.'), the gender marker *-a* linearly precedes the plural number marker *-s*, which means that NumP should occupy a functional projection higher than the one dedicated to gender (GenP or WordMarkerP in Bernstein (1993), classP in Picallo (2008)).

There is no agreement on the presence of a dedicated functional projection for gender in the DP (Alexiadou, 2004; Alexiadou, Haegemann, Stavrou, 2007; Carstens, 2000; Di Domenico, 1997). Alexiadou (2004) rejected Bernstein's (1993) generalization by showing that the correlation between presence of gender markers, noun movement and noun ellipsis (correlation that Bernstein takes as a manifestation of gender projection in the DP) does not hold when more languages are observed and compared (e.g., Italian, Spanish, Greek and Hebrew in her work). In Alexiadou´s (2004 and Alexiadou et al. 2007) view, (most)[25] names enter the derivation with their inherent gender feature, which is stored in the lexicon as part of the lexical nodes, contrary to number, which is always syntactically assigned. For what concerns the structure of the DP, number and gender then clearly differ: the former is a syntactically active feature, which dispose of a functional projection within the DP, the latter is represented on the NumP and is therefore parasitic to the structure (De Vincenzi and Di Domenico, 1999; Di Domenico, 1997). Finally, an operation of feature coping and sharing realizes concord of the head noun with its modifiers and the determiner (Alexiadou et al. 2007, Carstens, 2000; Giusti, 2009).

I will not go further into the detailed reasons for rejecting the hypothesis of a functional head dedicated to gender feature projection, what really matters to the present

[25] Alexiadou (2004) distinguishes between nouns endowed with an intrinsic gender feature and nouns that get their gender feature assigned at the syntactic level. The difference will be brought up in the discussion in the next paragraphs.

work is the relevance of gender in the syntactic derivation and at PF (Alexiadou, 2004). The studies reviewed above all agree on the presence of GG in syntax and the necessity for its retrieval whenever a noun enters a syntactic computation (at least in Italian). This suggests that the mapping between grammatical gender and its overt realization through morphological markers on determiners, modifiers and any other element sharing the feature with the noun should be quite robust in languages like Italian, and therefore GG should play a relevant role in language processing. The hypothesis is founded on Longobardi (1994) and is further supported by Friedmann and Biran (2003).

In the next section, I will present some major information on lexical (and therefore gender) retrieval. Background information will be indeed particularly useful for the evaluation of how PADs retrieve GG in comparison to unimpaired speakers.

3.4 Grammatical gender retrieval

Grammatical gender retrieval is a complex task, due to the high amount of information coming from different sources and cognitive spheres that has to be integrated in order to successfully activate the correct gender of target nouns.

Given the amount of information activated in GG retrieval, it is plausible to expect that extra-linguistic cognitive abilities are relevant to the process too, in particular in what concerns inhibition of alternatives and integration of mismatching information. Inhibition of alternatives, for example, is decisive in sentence processing when either gender or semantic cues prime the target noun (Manenti et al., 2004). Paraphasias and TOTs are examples of simultaneous failures at reaching the threshold level for the activation of the target word, and at inhibiting eligible alternatives (Badecker et al., 1995; Paganelli et al., 2003).

Coping with mismatching information retrieved from different sources can be challenging, as claimed by De Martino et al. (2011). In their study, authors unveiled that unimpaired speakers of Italian are slower and less accurate at processing nouns like *problema* ('problem.M'), which are characterized by information mismatch between grammatical gender (masculine, in this case) and morpho-phonological information on the suffix (the *-a*, otherwise typical of feminine nouns). Therefore, there is no doubt that GG retrieval is a complex process.

In this process of word activation, in which the speaker must activate and retrieve lexical, morphological and syntactic information, the role of semantic information with respect to natural gender has not received a clear evaluation yet. Bates et al. (1995, a.o.) addressed the problem and did not find a specific effect of natural gender on a grammatical gender decision task. According to their results, semantic information concerning natural gender does not take a major role in GG activation. In contrast, Andonova et al. (2004) proved that natural gender elicit faster and more accurate performance in a gender classification task in Bulgarian. According to the authors, the effect might be triggered by the fact that Bulgarian GG system is based on three genders: feminine, masculine, and neuter, and only a very narrow number of animate items belongs to the neuter class. This might result in an enhanced correlation between biological status and grammatical gender, as compared to two-gender languages. The hypothesis lacks further data from cross-linguistic studies to be confirmed. Nonetheless, I think it should be assumed that information on animacy and natural gender is activated as part of the semantic contents connected to the target lemma, independently of how they interact with GG in different languages.

3.4.1 A double route to grammatical gender retrieval

Many studies (Alexiadou, 2004; Badecker et al., 1995; Cantone and Müller, 2008; Carstens, 2000; Caramazza, Miozzo, 1997) agree on the hypothesis that GG is stored at the lexical level attached to the lexical entry. The hypothesis is particularly plausible in the case of irregular words, as those belonging to Italian declension Classes V, VI and VII. Although these irregularities are actually the result of linguistic changes, it is highly unlikely that common speakers retain awareness of the Latin or Ancient Greek origin of those nouns, for whom it is more efficient to assume that GG is stored in the lexicon. Nouns in Class I and II, on the other hand, might imply the automatic activation of morpho-phonological information provided by transparent ending suffixes, clearly distinguishing between masculine and feminine entries (as well as between singular and plural). The hypothesis is supported by data on gender decision: unimpaired speakers are faster and more accurate at judging transparent nouns than opaque items; moreover, their performance is at minimum on irregular nouns. This means that morphology plays a role

in gender checking, at least in perception (Bates et al., 1995; Padovani and Cacciari, 2003).

Researchers capitalized on the results of psycholinguistic studies in order to suggest the existence of a double route to lexical access. On one hand, irregular nouns lead to the necessary postulation of a direct access to the abstract information on GG, independently of any other phonological cue or semantic value: that is the case of words like *Weib* ('woma.N') in German or the irregular Italian noun *mano* ('hand.F'), which is feminine, despite the *-o* ending marker otherwise typical of masculine words. How to retrieve then their correct GG if not from the lexicon?

On the other hand, effects of regular morphology in gender retrieval or grammatical judgement tasks unveil the existence of a morpho-phonological access to GG for regular nouns (Bates et al., 1995).

Gollan and Frost (2001) posit direct access to the abstract gender node associated to each lexical entry as the fundamental strategy for GG retrieval, and the inference of GG through noun forms as a secondary, but more concrete, access to gender. Badecker, Miozzo and Zanuttini (1995) also claim that the former strategy, namely direct access to GG in the lexicon, is superior to the form-based one in that it allows the correct implementation of irregular nouns. Moreover, authors assume that the form-based strategy (i.e. direct mapping from morpho-phonology cues to gender) plays a greater role in language acquisition, but its relevance decreases along with the consolidation of lexical entries. Various researchers therefore assume the existence of a double route to the activation of GG, although it might be the case that the two play different roles in different languages. Each language has its specific system with respect to GG and factors like the gradient of transparency or opaqueness might have a high impact on how gender is processed.

3.4.2 Lexical retrieval and grammatical gender

In order to better exemplify how speakers retrieve GG from the lexicon, it is useful to spend a few words on the psycholinguistic model of lexical activation. There is now wide agreement on some fundamental ideas first presented by Levelt (1989)[26]: speakers

[26] While some core ideas of Levelt´s (1989) proposal have received much support, others were criticized and rejected within the scientific community; in particular, I will briefly touch upon the discussion about

progressively access lexical entries through semantic and lexical selection, finally resulting in the activation of the corresponding phonological nodes. In Levelt´s (1989) view, activation flows through at least three different types of nodes (or levels) of representation: first, the speaker must decide which kind of conceptual information he or she intends to communicate and activate the corresponding semantic representation. As a second step, activation flows to the lexical level, where lexical selection takes place: more lexical nodes might activate simultaneously in response to the semantic intention, but only the most relevant lemma will be selected. At this point, grammatical information attached to the lemma is retrieved: among it, lexical class and gender in the case of nouns, or argument structure in the case of verbs. The last step corresponds to the activation of the fitting phonological nodes composing the retrieved lexeme.

The original proposal formulated by Levelt (1989) was actually modified in a number of subsequent works (see Indefrey & Levelt, 2004). For instance, much debate concerned whether *lemma* should refer to the entries in the syntactic or in the semantic lexicon (or both), and whether the syntactic content is always activated in lexical retrieval. The base-line model described above was enriched in order to account for the process of word retrieval as well as for word reading and for word oral comprehension. However, for the sake of the current discussion, it is not necessary to enter the details of the discussion. It will be sufficient to consider that most proposals reproduce the relevant partition between the conceptual system, the semantic, the syntactic and the phonological levels, and maintain the core structure proposed by Levelt for the activation of the different levels (see Friedmann, Biran & Dotan, 2013, for an overview).

A much-debated aspect in Levelt's proposal concerns the role of syntactic information in lexical retrieval. With respect to this, Caramazza (1997) adopts the idiom *syntactic mediation* in order to describe the intermediate position of syntactic information within the lexical retrieval process proposed by Levelt (1989), thus explicitly referring to the role of syntactic information as mediator between semantic and phonology: syntactic retrieval follows semantic activation and precedes phonological activation.

Results from studies based on the elicitation and observation of TOT-states, were indeed interpreted as proof of grammatical activation preceding phonological activation.

---

the number of levels and nodes that should be activated in order to succeed in the correct activation of the target noun.

For Italian, Vigliocco et al. (1997) and Badecker et al. (1995) observe that both unimpaired and anomic speakers show gender congruency or gender knowledge despite ineffective lexical retrieval. Authors also explain the phenomena as due to the privileged status of GG with respect to phonology in the process of lexical retrieval, and they ascribe speakers' failure to an insufficient activation of the phonological level, while semantic and syntactic activation are completed successfully.

Caramazza (1997) and Caramazza and Miozzo (1997) reject the idea of phonological activation depending on the previous successful retrieval of grammatical information. Authors further support their view by calling into cause cases of healthy and aphasic speakers that have level-specific impairment in different tasks requiring the retrieval of phonological or morpho-syntactic information. Authors analyse results from naming tasks, and focus on the information participants can provide in TOT-states: in some cases they provide both GG and phonological information (e.g., number of syllables, first and/or last phoneme, etc.), in others only either one of the two. The case in which they provide GG but no phonological information is compatible with the hypothesis of a syntactic mediation between the semantic and the phonological level; however, the opposite pattern with phonological information but no clue about GG is not. Researchers propose then a slightly modified version of Levelt´s (1989) model for lexical access, in which the difference between lemma and lexeme is neutralized, and they claim the exclusive existence of single lexemes in the form of abstract nodes mediating between the semantic and the phonological or orthographic forms of a word and connecting to grammatical information.

The debate on the actual structure of lexical entries and the process for their retrieval is very rich. However, for the sake of the present discussion, what really matters is the presence of information record for GG within the lexicon. Most researchers now agree on it (Alexiadou, 2004; Badecker et al., 1995; Caramazza, 1997; Caramazza and Miozzo, 1997; Friedmann and Biran, 2003; Levelt, 1989, Vigliocco et al., 1997).

Alexiadou (2004) agrees with the general assumption that GG is recorded in the lexicon and retrieved from it along with the lexemes, but she also points out that exceptions to this general mechanism exist. The author in particular refers to word pairs that receive their gender specification contextually. That is precisely the case of word pairs of the type *ragazzo*/*ragazza* ('boy'/'girl'), which realize a gender opposition based

on the final vowel marker: *-o* for nouns denoting masculine entities, and *-a* for their feminine counterparts. The kind of noun that shows this type of gender alternation is usually specified for the [+animate] feature. Non-animate nouns do not have such alternations. For those noun pairs, gender is assigned contextually, depending on the animate referent they are linked to: the natural gender of the animate referent is then copied on the noun in form of grammatical gender. The locus of gender assignment for nouns referring to animate or human referents remains unspecified in Alexiadou (2004) and in Alexiadou et al. (2007). According to Di Domenico (1997) gender assignment for otherwise unspecified nouns takes place in the numeration, short after lexical items are selected and before they enter the syntactic derivation.

The subdivision of nouns into two categories, those with specified intrinsic gender and those with referent-dependent gender finds support in the psycholinguistic research: Franzon et al. (2014) administrate a phrase completion task to 24 unimpaired speakers of Italian. Participants examine adjective-noun pairs; in each couple, either one of the two elements lacks the ending morpheme and participants are requested to complete it. Their performance is measured according to reaction times and accuracy: a significant difference between nouns with intrinsic GG and those with contextually dependent GG emerges. Namely, subjects perform faster and more accurately on nouns that contextually receive gender specification, even though this is considered to be the least common procedure among nouns, being the majority of words intrinsically specified for gender (Franzon et al., 2014).

Summing up, in the present paragraph, I reviewed a number of studies on gender that point out the complex amount of abilities and factors involved in the process of GG retrieval. For instance, the capacity to inhibit alternatives and to address selective attention in case of incongruent information with respect to natural gender both co-operate in order to achieve the activation of the target GG. I also presented the double route that speakers can take to reach their goal: either they retrieve GG from the word-form or they recover it directly from the lexicon. This last route builds on the assumption that the feature for GG is stored in the lexicon at the lemma level and is activated, according to Levelt's model (1989), along with grammatical information and prior to the phonetic word form. In the next paragraphs, I will further explore GG retrieval from a different point of view, namely the neural substrates that carry out the task.

3.4.3 Neural correlates of grammatical gender retrieval

One further perspective from which researchers have looked at grammatical gender is the neural correlates involved in the retrieval process. By using neuroimaging and physiological studies, researchers have managed to point out which brain areas are mostly involved. As already illustrated, the process of GG retrieval is rather complex; this is due to the need for integrating different linguistic spheres, namely phonetics, syntax, morphology, etc., the levels at which GG plays a role. Considering all these, it is not difficult to imagine that GG retrieval relies on a bilateral network of focal areas (Padovani et al., 2005), playing complementary roles in the process.

A comparison among studies reveals some dissimilarities across results with respect to the neural foci that show significant activation in correspondence to GG retrieval. This comes as no surprise if we consider the high incidence that task designs can have on the process to be performed and therefore on the areas that speakers must activate in order to comply with task requirements. Nonetheless, a clear result emerges from all studies, namely the significant activation of Broca's area (Heim et al., 2006; Hernandez et al., 2004; Miceli et al., 2002; Padovani et al., 2005): areas BA44 and/or BA45 in the left hemisphere systematically activate in all tasks involving grammatical gender retrieval. Other areas that are often reported to activate in the left hemisphere are BA47 and B9. What is intriguing, moreover, is the fact that some studies report activation of the inferior portion of BA44, while others accentuate the role of the superior part of BA44 together with BA45/BA47.

Heim's (2008) precise review of a considerable number of studies on the issue succeeds in discriminating what determines the different roles played by BA44 and BA45. First, the robust presence of activation of BA44 across studies confirms that this is the main locus devoted to grammatical gender retrieval; therefore, further variation must be ascribed to the different techniques through which speakers can access GG. In particular, Heim et al. (2005) claim that tasks requiring retrieval of GG without verbalization of the determiner are responsible for activating BA44 in its inferior portion. Tasks allowing for explicit retrieval (i.e. determiner verbalization) use a different and more spread network, which has its core nodes in the following foci: BA 45/BA47 and the superior part of BA44. Authors ascribe the observed differences in patterns of activation to the asymmetry between automatic verbalization of the determiner or its mere

implicit retrieval, an issue that is shaped by tasks' requirements and constraints, rather than by word-specific characteristics.

This interacts with another important factor concerning GG: as described in previous paragraphs, speakers can retrieve GG either through a form-based route (and that is the case for morphologically transparent nouns) or through direct access to the syntactic information in the lexical entry, the only available option for opaque and irregular nouns. It is therefore plausible to suppose that the two routes bring along different patterns of activation of neural substrates. Padovani et al. (2005) address the issue in an fMRI study based on gender classification with transparent, opaque and irregular Italian nouns but fail at unveiling clear differences in the neural substrates that activate in correspondence to one or the other route, except for an enhanced activation of foci in the left hemisphere in the case of non-transparent words. Of course, the lack of sharp differences cannot be automatically used to claim that there are not specialized neural circuits that support the different access routes to GG gender, and the issue certainly deserves further research because the cited study only involved six speakers of a single language.

For the moment, and for the sake of the present study, it will be enough to keep in mind that GG retrieval recruits a neural circuit made of several areas, in which BA44 and BA45 play a major role.

3.5 Vulnerability of grammatical gender

Previous research has pointed out how vulnerable grammatical gender is under specific conditions, among these: bilingualism, second language acquisition and acquired linguistic disorders. This is no surprise in the light of the complexity of the retrieval operation and of the uneven performance of unimpaired speakers with respect to different declension classes, as described in the previous paragraph.

Interestingly, correct gender assignment does not represent a challenging goal in the L1 acquisition of Italian[27] (Chini, 1995; Kupisch et al., 2002), despite the presence in the system of opaque words. Kupisch et al. (2002) report good mastery of GG by a native

[27] Kupisch et al. (2002) showed nonetheless that the gradient of difficulty of GG acquisition in L1 could be language specific, depending on the morpho-phonological characteristics of the observed system: they found discrepancies between acquisition paths and rates, while comparing L1 acquisition of Italian and French.

Italian child even at a very early age: between age 1;7 and 2;7 the child´s rate of mistakes is as low as 1.9% in gender assignment. Chini (1995) monitored a child from age 2;2 until age 4 and found only one mistake over this long span, although this was repeated several times.

Bianchi (2013) recorded gender mis-assignment in the language of adult early Italian-German bilinguals (2L1): participants provided signs of vulnerability in their gender assignment process, despite good agreement accuracy. In other words, mistakes in gender assignment are frequent in the language of speakers who, in contrast, master native-like performance in gender agreement both with elements internal and external to the DP. Bianchi (2013) ascribed the performance pattern to both language-internal and language-external factors: irregular nouns like those belonging to declensions Classes IV and V are language-internal sources of difficulties, while reduced exposure to linguistic input can play a major role in making GG assignment vulnerable. More on the vulnerability of GG, previous research on gender retrieval and agreement also put forth the hypothesis that GG can undergo a process of attrition in case of interrupted acquisition of L1 (Anderson, 1999; Håkansson, 1995): thus suggesting the idea of a rather instable status of GG within linguistic competence.

I will not get here into the issue of GG in second language acquisition, as the topic is not strictly relevant for my research. It will be enough to mention that the acquisition path of GG can be highly influenced by the amount of similarities (or dissimilarities) shared by the L1 and the target language (Franceschina, 2005). Moreover, L2 learners can manifest difficulties both on GG assignment (Bianchi, 2013; Carroll, 1989) and on agreement configurations (Hawkins and Franceschina, 2004; Matteini, 2014; Montrul, Foote and Perpiñan, 2008), probably as a major effect of frequency.

In the following paragraph, I will first present results from previous research on the status of GG in aphasic speakers; second, I will review available studies on PADs.

3.5.1 Grammatical gender retrieval in aphasic patients

A milestone in the investigation of GG retrieval in aphasia are the data presented in Luzzatti and De Bleser (1996), (see also De Bleser, Bayer, Luzzatti (1996)): authors report the case of two native Italian speakers with acquired language disorders due to an aneurysm, in one case, and a head injury, in the other. Both subjects undergo the tests a

few years after the onset of their aphasia, which manifests with agrammatic spontaneous speech. Moreover, participants have a medium-to-mild impairment in sentence repetition and naming, and disrupted comprehension of both active and passive sentences. The study comprehends eight tasks concerning lexical morphology; in particular, participants perform the following ones: gender production in simple nouns and proper names, gender agreement in noun-adjective phrases, gender production in derived nouns, gender production in compounds, pluralization and singularization of simple nouns, pluralization of compound nouns, production of prepositions in prepositional compounds and production of adjectival derivational suffixes. The material in use in the study is very extensive, as well as the amount of collected data is; therefore, in what follows, I will limit my attention to the results from the tasks on gender retrieval, the most interesting one for the sake of the present discussion.

Overall, De Bleser and Luzzatti (1996) report that their patients have different gradients of proficiency depending on the nature of the experimental material in use, with a clear dissociation between simple and compound nouns, being the former better preserved than the latter. Roughly speaking, both subjects seems to retain a good competence on simple nouns, while compounds represent a weak point in their lexical competence; nonetheless, a closer look at the data actually reveals some signs of impairment on simple nouns too, especially in the case of the patient suffering from a severe brain injury. Problems arise in particular in the following domains: gender retrieval is often unsuccessful for opaque nouns ending in *-e* (declension Class III in Chini (1995)), which are classified as masculine. The same happens for irregular feminine words ending in *-o* (e.g., *mano,* 'hand.F', from declension Class VII in Chini, 1995). Data suggest the onset of a rough gender generalization, allowed by the disregard of lexical knowledge and the reliance on morphological rules. As for the task on gender-agreement in simple noun-adjective phrases, both participants achieve high rates of accuracy, most probably thanks to their capacity to rely on gender cues, provided by the determiner of the phrases to be completed. It is also interesting to point out that the performance improves in derived nouns ending with *-e*: this suggests that subjects are sensitive to the grammatical information contained in the derivational suffix and actually process nouns like *malore* ('sudden illness.M') differently from nouns like *cane* ('dog.M'). In the former, the suffix *-ore* is recognized as a bearer of masculine GG, while nouns like *cane* provide no

morphological cue. This last result is not surprising, especially in the light of other studies, which proved umimpaired speakers' capacity to break down words into their constituent morphemes during word recognition (Marslen-Wilson et al., 1994; Meinzer et al., 2009).

The status of GG retrieval in agrammatic patients is still debated though, given that attempts to replicate the results presented in Luzzatti and De Bleser (1996) were only partially successful for other languages. For example, Bastiaanse et al. (2003) collected data from ten German and nine Dutch speakers with Broca´s aphasia, and found virtually at-ceiling performance on GG retrieval in task requiring either sentence completion or production. This is a particularly surprising result if we consider the involvement of Broca's area in GG retrieval, as exemplified in paragraph 3.4.1. However, as anticipated, GG assumes very specific characteristics depending on the language, and this should dissuade researchers from relying on cross-linguistic generalizations.

Moreover, as cited above, failures in word recall attempts can be a rich source of information on how lexical retrieval works and how proficient speakers are at it. I refer in particular to studies examining paraphasias and TOT states. Badecker et al. (1995) describe the performance of Dante, a young Italian speaker affected by anomia[28], in picture naming and in sentence completion tasks. When the subject is in an anomic state, experimenters ask him questions about the word he is trying to retrieve. Questions concern GG, first and final phonemes, length of the word and similarity to other words. Through questions, Badecker et al. (1995) unveil a clear dissociation in Dante´s ability between grammatical gender (very accurately produced) and information about the phonological and orthographic forms (poorly recalled), with no effects of regularity, meaning that GG is equally well retrieved for both regular and irregular items. Dante also completes a gender-decision task based on the nouns he has not retrieved in previous tasks. Again, his performance is very accurate: he correctly retrieves the grammatical gender of words he failed to activate in previous tasks (98%). Authors read the results as a proof of the fact that anomic patients have access to grammatical information about target nouns, meaning that activation of the target lemma is successful, while they fail at selecting the proper phonological nodes, necessary for the word spell out. With reference to the model of lexical retrieval first suggested by Levelt (1989), Badecker et al. (1995)

[28] Dante was affecteded by amnesia and anomia as a consequence of a long comatose state, most probably due to a meningoencephalitis. Badecker et al. (1995) describe his spontaneous speech as fluent, despite anomia.

claim that syntactic information (the lemma) precedes (and maybe it even is a prerequisite for) the activation of phonological information (the lexeme). Based on these assumptions, authors set the locus of impairment for anomic patients in-between the lemma and the lexeme levels: that is where the flow of activation is interrupted and word retrieval is blocked.

Not all researchers agree on the hypothesis above. The capacity to retrieve GG despite unsuccessful word activation is ascribed by Friedmann and Biran (2003) to language specific phenomena. Basing their speculation on data from Hebrew anomic speakers, who actually do not show gender congruency on their paraphasias, the authors claim that the capacity to retrieve GG is biased by syntactic constraints, which are language-specific. In doing so, they refuse to take into account impairments more directly linked to the word retrieval process, for which a universal (and therefore homogeneous across languages) model should be assumed. In Friedmann and Biran´s (2003) view, it is therefore incorrect to claim that agrammatic speakers usually preserve GG knowledge. Rather, GG retrieval depends on how nouns are usually implemented in the syntactic derivation of the language at stake. The lower the acceptance of bare nouns in the syntax is, the more frequently speakers produce well-formed DPs rather than bare nouns, and therefore the better GG is retrieved.

It is important to consider that anomia does not strictly imply semantic deficits. That is a macroscopic difference with respect to PADs, who actually have been extensively reported to suffer from severe semantic deficit too: in their case, anomia is most probably the ultimate consequence of erosion of semantic information. It is therefore not possible and fruitful to directly compare speakers affected by different kinds of impairments. The reviewed studies are nonetheless interesting for the purpose of pointing out information concerning GG retrieval and how vulnerable (or not) it might be.

In order to complete the overview in GG retrieval and get gradually closer to the main focus of the present study, in the following paragraph I will review two studies from background literature, both of which deal with the competence and sensitivity of PADs on linguistic tasks involving grammatical gender.

3.6 Grammatical gender in PADs

Among the numerous symptoms of AD, anomia appears quite early and persists along with the worsening of the disease. Researchers devoted much attention to the phenomenon and many agree on recognizing a number of causes to anomia, as well as few consequences affecting different levels of linguistic competence. Among these, the ability to compute complex clause structures decreases (Small et al., 2000; Waters, Rochon and Caplan, 1998). Overall, no major disruption has been reported concerning gender, but I should also say that the phenomenon has not received much attention so far. Previous literature provides two studies addressing the status of GG in PADs, with the purpose of investigating language processing at the lexical and sentential levels. The two studies I am going to review both focus on Italian-speaking subjects, while they tackle the issue from different perspectives: Paganelli et al. (2003) investigate whether PADs manage to retrieve GG despite failing at completing lexeme activation; their purpose is to find out at which stage the process of lexical retrieval is impaired. Manenti et al. (2004) focus on sentence processing and investigate the effects of gender priming in combination with semantic priming.

Paganelli et al. (2003) enrol seven Italian-speaking PADs and compare their performance on two picture-naming tasks to that of unimpaired controls. Both tasks include 170 stimuli, selected according to a combination of regular (declension Class I and II) and opaque items (declension Class III): in the first task, authors ask participants to name objects by their bare noun; in the second, they ask participants to produce complete DPs (determiner + NP). Controls perform differently in the two experimental conditions: in the former (bare noun naming), they do not have any consistent gender congruency effect: when they fail at retrieving the target word, their outputs is not of the same GG as the expected word. A gender effect (i.e., the mistake bears the same GG as the target word) arises in the latter condition, when a DP is expected. PADs never have a gender congruency effect in their naming failures, independently of the conditions in use. In other words, the activation of syntactic frames (the DPs) influence the performance of normal controls, but not the one of PADs, meaning that the two observed groups experience failures at different levels. Normal controls succeed in activating the proper semantic nodes and the activation flows up to the required syntactic frames; failures ultimately arise just before the selection of the target lexeme. At that point, only

alternatives compatible with the already activated syntactic frame are admitted. The mistake is then constrained by the activated syntactic information, namely gender. On the other hand, patients mainly produce semantic errors with no higher-than-chance gender congruency, which means that their mistakes are generated pre-lexically, at the conceptual level, before the activation of syntactic information. The authors posit then the locus of disruption at the semantic level and investigate gender only as a marker of the border in-between two different stages of lexical retrieval. In my view, the phenomenon deserves more attention; we should indeed consider that it could also be the case that, on top of the mis-selection of semantic alternatives, ineffective construction of adequate syntactic conditions co-occurs, further favouring naming mistakes.

Manenti et al. (2004) adopt an on-line testing paradigm in order to investigate PADs´ ability to take advantage of semantic and gender priming cues while processing orally presented sentences. They also address the converse capacity to inhibit erroneous alternatives, with interesting results. Stimuli are short two-sentence narratives, with a target word presented visually. Participants listen to the narrative (in 6) and complete it by reading out aloud the provided target nouns (in 7). The semantic cue to target nouns is given by the sentence meaning, while gender cue is provided by the determiner, which is always the last word the participants hear before seeing the visual target. The task includes four conditions: in the first one, both semantic and gender cues are congruent with the target (7a); in the second (7b) and in the third (7c) conditions, only one of the two cues is congruent (either the semantic or the gender one); in the fourth condition (7d), neither of the two factors contributes to correctly priming the target word.

(6) *Quando vado a letto prima di addormentarmi leggo sempre*
'When I go to bed before falling asleep I always read'

(7) a. un LIBRO
'a.M book.M'
b. un TOPO
'a.M mouse.M'
c. una LIBRO
'a.F book.M'

d. una TOPO

'a.F mouse.M'

Reading times are measured and results reveal different performance patterns across controls and PADs. Overall, young and elderly controls clearly have a facilitation effects in reaction to positive semantic and gender priming (7a). PADs, instead, experience interference too, on top of facilitating effects. In other words, PADs benefit from combined semantic and gender priming (7a) and they can process faster and more accurately the first condition. Moreover, incongruence of combined priming cues is hard for them to inhibit and it results in slower RTs (in 7b and 7c). The performance is therefore compatible with preservation and integration of grammatical knowledge in sentence processing. What actually appears problematic is the inhibition of alternatives and incongruent priming cues.

In conclusion, the two studies I reviewed above brought to different insights into the processing of gender features by PADs. On the one side, Paganelli et al. (2003) underline the relative irrelevance of grammatical information activation in the process of lexical selection, given that lexical disruption takes place at the preceding semantic level; on the other, Manenti et al. (2004) confirm the active role played by gender priming in combination with semantic priming in lexical activation during sentence processing. Indeed, their patients read a sentence's final word faster if this is primed both by semantic and syntactic cues (a determiner bearing the correct grammatical gender), while their reaction times are longer in correspondence to incongruent primers.

3.7 Conclusions

The review of background literature was useful to observe how GG is a very uneven category across languages. Still, with respect to Italian, studies pointed out the following information:

a) Italian has a mixed system with respect to transparency of final morpheme of nouns: the majority of words is transparent, many are opaque, few are irregular (Chini, 1995).
b) GG does not have a dedicated projection in the functional domain of DP (Alexiadou, 2004; Carstens, 2000); nonetheless, it plays a relevant role both at the

syntactic level and at PF: it allows feature-sharing and agreement both inside the DP (e.g., on determiner and modifiers), and outside the DP (e.g., on past participle and pronouns).

c) Most nouns have intrinsic GG, which is assigned arbitrarily and independently of semantic, morphological and phonological information; the relevant feature is stored in the lexicon attached to the lemma node and is retrieved whenever the noun is activated (Caramazza, 1997; Caramazza and Miozzo, 1997; Levelt, 1989; Vigliocco et al., 1997). In some other cases, GG is assigned inherently through a checking operation with the [+human/+animate] referent in the context (Alexiadou, 2004; Di Domenico, 1997; Franzon et al., 2014). That is the case of word pairs like *ragazzo*/*ragazza* ('boy.M'/'girl.F').

d) There exists a double-route to gender retrieval. Results from psycholinguistic research provide evidence both in favour of direct access to the abstract feature in the lexicon and of form-based retrieval of GG (Badecker et al., 1995; Bates et al., 1995; Gollan and Frost, 2001)). The relevance of the two routes to GG activation might depend on the specific linguistic system at stake and, in particular, on how transparent or opaque this is.

e) Broca's area plays a prominent role in the neural circuits engaged in GG retrieval (Heim, 2008).

f) Gender is a vulnerable piece of knowledge. The claim is based on the observation of GG in specific linguistic conditions like unbalanced bilingualism (Bianchi, 2013), interrupted L1 acquisition (Anderson, 1999; Håkansson, 1995), and acquired linguistic disorders (Luzzatti and De Bleser, 1996 a.m.o.).

4 GRAMMATICAL GENDER RETRIEVAL IN PATIENTS WITH ALZHEIMER´S DISEASE

4.0 Introduction

In the present Chapter, I present results from a gender retrieval task performed by Italian PADs. The reasons for investigating this field mainly resides in the nature of grammatical gender (GG). GG is a very interesting linguistic element that involves a number of linguistic levels and competences: it is stored in the lexicon as part of the lexical entry and it brings along a syntactic value, which is relevant in the syntactic derivation. Moreover, it also has a morphological and phonological manifestation. To a certain extent, it belongs to the class of procedural components of linguistic competence, but it also implies declarative knowledge, as in the case of opaque and irregular nouns (see Chapter 3). Investigations on GG retrieval could therefore be highly informative on the linguistic competence of PADs.

The chapter is organized as follows: in 4.1 I report the research questions I intend to address in the present study; section 4.2 presents information concerning the materials in use, participants, procedure and coding guidelines. In 4.3 I provide results; these are further discussed in 4.4.

4.1 Research questions

In light of the background literature I have reviewed in the previous chapter, I would like to address three research questions concerning GG in PADs.

The first focus of my interest will be the kind of strategy PADs adopt in order to access GG. There exist two different, but complementary, routes to GG: one consists of direct access to the GG information stored in the mental lexicon, the second exploits form-based GG retrieval. The former always guarantees successful retrieval, while the latter works safely only on transparent nouns. I investigate whether PADs change the way they employ these two routes as a reaction to altered linguistic and cognitive abilities. As discussed in Chapter 2, PADs have shown to be more spared on procedural rather than

on declarative components of the language. For instance, PADs show a clear tendency to modify irregular verb forms by applying regular morphology, thus performing an operation of over-regularization (Colombo et al., 2009; Walenski et al., 2009). In order to answer the question in the present study, PADs complete a gender retrieval task including transparent, opaque and irregular nouns. If the generalization about spared procedural mechanisms with respect to declarative components is correct, I should find a specific impairment on opaque and irregular nouns, in association with strategies of systematic regularization. If not, the three categories (transparent, opaque and irregular nouns) should be equally (un)affected by mistakes on GG retrieval.

Second, I intend to investigate the role of natural gender in grammatical gender retrieval. In unimpaired speakers, natural gender is not relevant, or at least Bates et al. (1995) do not find any facilitation effects for words endowed with natural gender with respect to the inanimate ones. Still, there are reasons to assume that this might not be the case for PADs. The doubt is worth trying to shed light on the topic. It is well-known that PADs suffer from severe anomia, most probably due to semantic erosion and loss of world knowledge (see Chapter 2). This process of progressive semantic information loss first starts by affecting fine-grained information and then spreads on. Consequently, speakers increasingly regress to core semantic macro-categories. This disruption and reorganization at the semantic level might have consequences at the syntactic one. In other words, a plausible hypothesis to be tested is whether natural gender could be employed as a compensating strategy in case of difficulties on GG retrieval. If the hypothesis is correct, PADs should present a dissociation between nouns whose referents are endowed with natural gender, and nouns that refer to inanimate referents. That is the reason for including both kinds of nouns in the material in use in the present study.

The third question I want to address concerns the role of derivational morphology in GG retrieval. I recalled in Chapter 3 that Luzzatti and De Bleser (1996) found a dissociation between words belonging to the opaque class of nouns ending in *-e*: the presence of derivational morphology seems to facilitate the performance in comparison to simple nouns. This is most probably due to the fact that derivational morphemes carry a GG feature, although this is not as transparent as the *-a* for feminine. In order to verify whether PADs have the same kind of sensitivity for derivational morphology, the test will also include derived forms as well as baseline ones.

Summing up, the present work takes into consideration three lexical factors:

a) Transparent, opaque and irregular nouns (Chini, 1995) will help analyse PADs' use of GG retrieval strategies, namely direct access to gender in the lemma versus form-driven retrieval. This will help to verify the status of the two possible routes to GG in PADs;
b) The role of natural gender will be addressed by introducing nouns characterized by natural gender, either masculine or feminine animates, and by comparing them to inanimate nouns, which are not specified for masculine or feminine biological gender;
c) Derivational morphology will also be represented in a subset of experimental items and compared to simple nouns, in order to test participants' ability to analyse nouns into their constituent morphemes.

4.2 The experimental study

In what follows, I will present results from a grammatical gender retrieval task performed by a group of PADs. The task material is designed in order to allow the investigation of mechanisms of access to GG and to evaluate the role played by different factors, like transparency of ending markers, influence of biological gender (as a semantic feature) and derivational morphology (in comparison to base forms).

Participants perform gender classification overtly, by speaking aloud the appropriate determiner for each noun. This procedure is more advisable than a silent classification performed through pointing to two symbols (either for masculine or for feminine gender) for a few reasons. First, Heim et al. (2005) report in previous literature that silent classification leaves open the opportunity to develop individual strategies for task completion. Post-task interviews to participants in their study reveal that subjects split into two groups: some participants implicitly retrieve the target GG and directly point to the appropriate symbol; others actuate a verbalisation strategy by silently producing a DP in order to meet a decision. By choosing overt classification, I therefore intend to reduce the variability across participants. Moreover, I find that pointing to symbols for feminine or masculine grammatical gender can be ambiguous and represent a potential confounding factor in the task: it is indeed well known that PADs have reduced selective

attention, and the two symbols for GG could therefore be easily taken as symbols for natural gender, thus confounding task performance and results. Overt verbalisation of GG does not leave any doubt on how participants are performing the task and allows for a better monitoring. The experimenter can make sure that participants do not randomly point to either of the two symbols and that they are actually performing the requested task.

4.2.1 Material

The present work owes a lot to the study realized by Luzzatti and De Bleser (1996) on aphasic speakers. Luzzatti and De Bleser (1996) evaluated the performance of two Italian agrammatic speakers on a series of tasks concerning GG retrieval. The eight tasks in use differentiated according to the kind of expected performance (phrase production, pluralization, etc.), and the lexical material in use (simple nouns, derived forms, compounds). Much of the original material for simple and derived nouns is in use also in the present task in order to allow for a comparison between PADs and agrammatic speakers; I made a few changes though, in order to adjust the task in the light of my own research questions, and to conform the material to the needs of PADs (i.e., reduced length of execution).

First, I intentionally excluded compounds from the present investigation in order to avoid inserting a further source of complexity. Previous findings from Chiarelli, Menichelli & Semenza (2007) pointed out that PADs are impaired at retrieving compound nouns in a picture-naming task and usually compensate their weakness by producing simple words or by applying the most productive structure among those available for Italian compounds, namely Verb-Noun compounds. Moreover, retrieval of the second element in the compound is much more impaired in comparison to the first one: authors explained the data in the view of a problem of information overload. Thus, given the intrinsic difficulties brought up by compounds, I decided to start my investigation on GG by focusing only on simple and derived forms.

The task includes 100 Italian nouns, subdivided across four main classes – regular, opaque, irregular and derived nouns – and a number of subclasses.

I set aside derived nouns for a moment, and start by illustrating the first three classes of simple nouns. The baseline categorization takes into account the transparency of final

word markers, this leads to the distinction between regular, opaque and irregular nouns. A further subcategorization takes into account natural gender (NG) and distinguishes between nouns referring either to animate (with masculine or feminine natural gender) or to inanimate referents. Moreover, words are counterbalanced across masculine and feminine gender. The results of the operation are presented in Table 4.1 for regular nouns, in Table 4.2 for opaque nouns and in Table 4.3 for irregular nouns.

Table 4.1. Regular nouns (1) include four subclasses.

Example	GG	NG	Marker	Class
a. *marito* 'husband'	masculine	masculine	*-o*	I (transparent)
b. *mamma* 'mum'	feminine	feminine	*-a*	II (transparent)
c. *mondo* 'world'	masculine	inanimate	*-o*	I (transparent)
d. *musica* 'music'	feminine	inanimate	*-a*	II (transparent)

Words in class (1a) and (1b) are regular nouns endowed with natural gender, while (1c) and (1d) are regular nouns with inanimate referents. In this latter case, regularity stems from the transparency of the final word markers (*-o* for masculine and *-a* for feminine). In the former case, (1a) and (1b), regularity is reinforced by matching natural gender too (in the sense that grammatical gender and natural gender coincide). Each subclass includes five items.

The nouns in the four subclasses of regular nouns ((1a) to (1d)) were then paired to equivalent opaque nouns, which share similar characteristics with respect to natural gender (masculine, feminine or inanimate) and grammatical gender (masculine and feminine). Table 4.2 provides examples.

Table 4.2. Opaque nouns (2) include four subclasses.

Example	GG	NG	Marker	Class
a. *padre* 'father'	masculine	masculine	*-e*	III (opaque)
b. *madre* 'mother'	feminine	feminine	*-e*	III (opaque)
c. *cuore* 'heart'	masculine	inanimate	*-e*	III (opaque)
d. *voce* 'voice'	feminine	inanimate	*-e*	III (opaque)

Table 4.3 exemplifies irregular nouns included in the experimental material. The irregularity exemplified in (3a) stems from information clash between the masculine GG (matching with masculine NG in (3c)) and the final marker *-a*, otherwise characterizing regular feminine nouns. Subclass (3b) includes nouns characterized by masculine natural gender and final marker *-a*; in this case, GG assignment regularizes according to the final marker: nouns have feminine GG. In this case the information clash is caused by the mismatch between masculine natural gender and feminine grammatical gender. Nouns in (3c) present the pattern already illustrated for (3a) (masculine GG despite *-a* as final marker), except for being inanimate and therefore not specified for natural gender. The reverse pattern is exemplified in (3d): inanimate referents, feminine GG and *-o* final markers.

Table 4.3. Irregular nouns (3) include four subclasses.

Example	GG	NG	Marker	Class
a. *poeta* 'poet'	masculine	masculine	*-a*	V (irregular)
b. *guardia*[29] 'guard'	feminine	masculine	*-a*	II (regular)
c. *problema* 'problem'	masculine	inanimate	*-a*	V (irregular)
d. *mano* 'hand'	feminine	inanimate	*-o*	VII (irregular)

[29] The selected names for this class refer to army-related jobs and roles (guard, recruit, etc.), which traditionally were carried out only by men and have been practised only more recently by women. Based on this consideration, they are classified as masculine for natural gender.

In order to balance factors across the items, each noun referring to an entity endowed with natural gender was paired to a noun sharing the same characteristics for GG, and declension class, but crucially inanimate[30].

In the experiment, the role of derivational morphology has been taken into account too: in order to test the relevance of derivational morphology with respect to GG retrieval, I included nouns from seven further classes. The derivational suffixes in the selected nouns all carry a specific feature for GG and can therefore be considered transparent, despite ending in *-e*. Table 4.4 provides examples.

Table 4.4. Derived nouns (4) included in the experimental material are divided into 7 subclasses.

Example		GG	NG	Marker	Class
a.	*pompiere* 'fireman.M'	masculine	masculine	*-iere*	III
b.	*direttore* 'director.M'	masculine	masculine	*-tore*	III
c.	*scrittrice* 'writer.F'	feminine	feminine	*-trice*	III
d.	*malore* 'stroke'	masculine	inanimate	-(*t*)*ore*	III
e.	*quartiere* 'district'	masculine	inanimate	*-iere*	III
f.	*stupidaggine* 'nonsense'	feminine	inanimate	*-aggine*	III
g.	*solitudine* 'loneliness'	feminine	inanimate	*-udine*	III

Nouns in the classes (4a), (4b) and (4c) have either masculine or feminine GG in accordance with the natural gender; the derivational suffixes in use are *-iere* and *-tore* for masculine items, and *-trice* for feminine nouns. The last four subclasses include nouns with no NG, characterized by four different derivational suffixes: two for masculine inanimate nouns (*-ore* and *-iere*), and two for feminine inanimate nouns (*-aggine* and *-udine*). In order to counterbalance the number of feminine and masculine nouns, class

[30] Nouns in classes (3b) were paired with nouns in class (3d). This represents an exception in the design, in that the two classes do not differ only with respect to natural gender. In this case, the opposition between the two classes is built on the fact that class (3b) is irregular in what concerns the clash between natural gender and final marker on one side, and natural gender on the other. In contrast, nouns in class (3d) are irregular in what concerns the relationship between feminine GG and *-o* final markers, otherwise representative for masculine nouns.

(4c) includes 10 items, rather than 5 (like all other classes do). This allows for an even number of masculine and feminine derivational nouns with either masculine or feminine NG, as well as for an even number of feminine nouns with feminine NG and feminine nouns with no NG. Derived nouns are characterized by lower frequency, with respect to the other classes in the study[31].

Overall, the five items included in each subclass[32] make a total of 100 nouns: 20 regular nouns, 20 opaque nouns, 20 irregular nouns and 40 derived nouns.

No word starting with a vowel was included in the task because, in that case, definite articles neutralize their phonetic difference between masculine and feminine grammatical gender, preventing the possibility to record whether participants retrieved the proper GG or not.

4.2.2 Participants

Forty-one participants in the PAD group were enrolled through two institutions operating in Northern Italy in the field of research on dementia.

All participants enrolled in the study meet the NINCDS-ADRDA criteria (1984) for a diagnosis of probable Alzheimer' Disease; they have received their diagnosis at least six months before the task administration and they take part in a monitoring programme, which implies that they undergo tasks for cognitive assessment twice a year in order to track the development of the disease. No patient with previous history of alcohol abuse, cardiovascular dysfunction or neuropsychological disorders is admitted in the study.

Participants in the PAD group are further subdivided into three groups according to their level of dementia; this is estimated through the MMSE 'Mini Mental State Examination Test' (Folstein et al., 1975), a widespread test in use in neuropsychology in order to assess the level of cognitive impairment in subjects suffering from different kinds of disease. The test includes thirty items that cover different areas, e.g. spatial and temporal orientation, working memory and calculation. Test scores range from 0 to 30, with higher scores (29 and 30) matching with intact cognitive abilities, and lower ones corresponding to severe impairment. In general, patients that score between 28 and 25

[31] Frequency rates were retrieved from the *Corpus e lessico di frequenza dell´Italiano scritto contemporaneo* (Bertinetto, Burani, Laudanna et al., 1996).
[32] Except for class (4c) including ten nouns for the reasons described above.

are considered 'mildly impaired'; 'moderate impairment' corresponds to scores in-between 24 and 19; finally, patients with scores between 12 and 18 are classified as 'severely impaired' (Folstein et al., 1975). Patients with very severe impairment, corresponding to MMSE scores lower than 12, are not admitted in the study because their level of dementia usually does not allow for task comprehension and completion. Thirty-six out of forty-one PADs completed also the Wh-question comprehension task (Chapter 6). All patients (41) participated also in the Relative clause comprehension task (Chapter 7).

Besides PADs, a second group of participants is enrolled. This is composed of healthy elderly speakers, characterized by age and level of education similar to PADs. Their function is to act as controls and provide with a benchmark for the evaluation of patients' performance. The control group includes 21 subjects, who were enrolled through two local associations in Northern Italy[33]. In order to be admitted in the control group, candidates had to perform at ceiling on the MMSE test, which corresponds to test scores equal to 29 or 30 out of 30.

The following table allows for a comparison between PADs and controls (CO), with respect to Age, Level of Education (LoE) and MMSE scores:

Table 4.5. PADs and COs have similar values for Age and LoE. Their MMSE scores differ.

	PAD	CO
n°	41	21
Age (sd)	77;6 (7;3)	75;10 (5;2)
LoE (sd)	6.5 (2.9)	7.3 (2.8)
MMSE (sd)	20 (3.7)	29.7 (0.45)

As for PADs, before the start of the test session families were informed about the study and the characteristics of the tasks in order to obtain written consensus from participants.

4.2.3 Procedure

Each noun is printed on an A4 sheet with Calibri 180 point-size font and centred according to the landscape page-layout. Nouns are randomized according to declension

[33] In order to control for diatopic variation, only controls from the same areas as PADs are admitted in the study.

class and GG. The experimenter displays the items on the table in front of patients and ask them to read the word aloud together with the suitable determinative article, thus providing a DP. A low line precedes each noun, in order to remind participants about how to perform the task.

Three items precede the experimental trials, in order to give participants the opportunity to familiarize with the task. No time constraints are set; a short break (two to three minutes) is allowed after the first half of experimental trials.

The researcher met participants individually. The experimental section took place in a quiet room at the hospital (for PADs) or at a local association (for controls).

4.2.4 Coding

The expected performance targeted the production of proper definite articles for each noun. Nonetheless, indefinite articles were equally accepted as a valid response, as long as gender markers were clearly recognizable.

Moreover, Italian counts on two different singular masculine definite articles, namely *il* and *lo* (both for 'the.M'). The two definite articles are characterized by a complementary distribution, which depends on the first phoneme of the following word. However, some of the participants in the task are speakers of a variety that counts only with one of the two articles, namely *il* (also in the form of *el*). For this reason, items requiring *lo*, but spelled-out with *il*, were classified as correct. In other words, the opposition between *il* and *lo* article forms, being a phonological rather than lexical and morphological phenomenon, was disregarded in the process of data coding and both forms were equally considered as correct, independently of distributional rules in Standard Italian.

Participants' outputs were subdivided into three categories: 'accurate', 'incorrect' and 'failed' retrieval. 'Accurate retrieval' means that a participant retrieves and produces the target definite (or indefinite) article for the noun. Under the label 'incorrect', I classify cases of ineffective gender agreement between determiner and noun, independently of the kind of determiner produced (definite/indefinite and the *il*/*lo* opposition for masculine). 'Failed' retrieval means participants read aloud the experimental items as a bare noun.

In case participants incorrectly perceived the target noun and substituted it with a phonologically or semantically similar one, the item was erased from the list of correct

target words and counted as 'failed', even in case the produced determiner corresponded to the target one. That is the case, for example, of a patient reading *zia* ('aunt.F'), instead of *spia* ('spy.F').

4.3 Results

All elderly speakers enrolled in the control group completed the task effortlessly.

In what follows, I analyse the data provided by 38 PADS out of the 41 subjects initially enrolled in the experiment. Three PADs dropped the experiment: for two of them the decision to interrupt the session was taken by the the experimenter due to their major behavioural problems.The third participant left the experiment for independent reasons.

Table 4.6 presents the results from the two groups and allows for a comparison between elderly controls (COs) and patients affected by Alzheimer´s Disease. The table represents levels of accuracy with respect to the total amount of items presented to each group, namely 3800 nouns for PADs and 2100 for COs.

Table 4.6. Percentages of accurate, failed and incorrect retrievals by elderly controls (CO) and PADs.

	PADs			COs		
	n°	SD	%	n°	SD	%
accurate	3570	7.09	93.94	2095	0.6	99.76
incorrect	134	3.2	3.52	4	0.6	0.19
failed	96	5.2	2.52	1	0.2	0.04

The performance of COs is at ceiling as they manage to retrieve the expected determiner for a 99.77% of given nouns. Overall, their performance is incorrect or fails only for five items out of 2100 (corresponding to the 0.23%): among those, four are cases of incorrect GG retrieval and one is an omission.

The performance of PADs is also very careful, as they all together provide with 3570 correct answers, corresponding to 93.94% of the total number of trials. Still, the number of incorrect or missing answers is much higher in PADs than in COs: they fail on 96 items (2.52%) and retrieve inadequate determiners for 134 nouns (3.52%).

Despite a rather proficient performance by PADs, there is no doubt that the two groups perform statistically differently with respect to the number of correct answers ($\chi^2(1,N=5900)=119.57$, $p<.01$).

Data are further analysed according to the level of impairment of the participants (mild for PAD1, moderate for PAD2 and severe for PAD3) in order to verify whether a pattern of accuracy reduction is at play along with the worsening of the disease. Data for the three groups are reported below:

Table 4.7 Percentages of accurate, failed and incorrect retrievals by PAD1, PAD2 and PAD3.

	PAD1			PAD2			PAD3		
	n°	SD	%	n°	SD	%	n°	SD	%
accurate	383	3.0	95.75	2018	3.0	96.10	1169	10.2	90.00
incorrect	10	2.8	2.5	63	2.8	3.00	61	3.9	4.7
failed	7	1.5	1.75	19	1.5	0.90	70	8.0	5.4

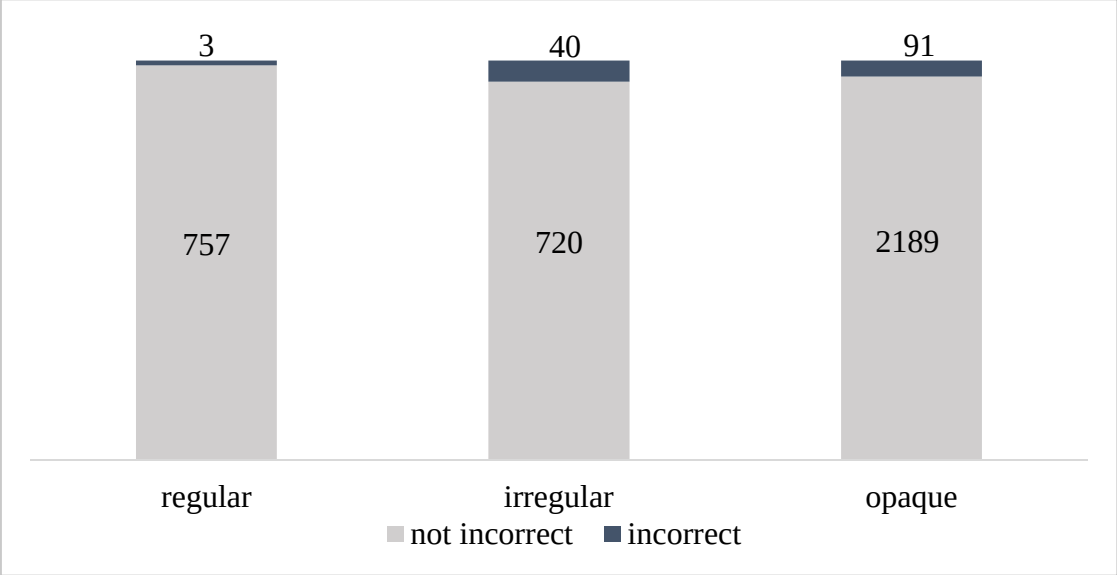
Despite a descriptive difference between the performance of the PAD3 group with respect to the accuracy of PAD1 and PAD2 participants, from the statistical point of view the number of incorrect and failed outputs does not reveal any significant differences among the groups. For this reason, for the purpose of the present Chapter, in what follows I will consider data from PADs as a unitary group and disregard the three aforementioned sub-groups.

The data analysis will now take into consideration what kind of mistakes PADs made. I would like to start the analysis by having a look at how the 134 occurrences of incorrect GG retrieval distribute across the three main noun classes (with respect to the transparency of the final markers). For this purpose, I will collapse simple and derived opaque nouns under the label 'opaque', given that the final marker *-e* characterizes nouns from both classes. Figure 4.1 provides with a first-hand impression of the incidence of incorrect[34] retrievals on the overall number[35] of experimental trials.

[34] In Figure 4.1, I want to focus the attention on incorrect outputs, leaving aside for a moment failed DP production (bare nouns). For this reason, failed outputs and correct productions are collapsed under the label “not incorrect”, as opposed to “incorrect”.

[35] Please recall that each participant saw 20 regular nouns, 20 irregular nouns, 20 simple opaque nouns, and 40 derived noun. In Figure1 simple opaque and derived nouns are collapsed under the “opaque” label. This explains the discrepancies between noun classes with respect to the total amount of items per class: 760 for

Figure 4.1. Number of incorrect and not-incorrect outputs provided by PADs for regular, irregular and opaque (simple and derived) nouns.



PADs are very solid at retrieving GG for regular words, while their accuracy decreases on irregular and opaque nouns. Indeed, the number on incorrect retrievals on regular nouns statistically differs from the number of mistakes on irregular nouns (χ^2(1,N=1520)=32.76, p<.01) and on opaque (simple and derived) nouns (χ^2(1,N=3040)=24.60, p<.01). In contrast, the difference on the amount of mistakes on irregular and on opaque words is not significant.

No error analysis can be run on regular nouns, as these are affected only by three mistakes and all of the same kind, namely cases of incorrect gender retrieval with no apparent explication. The three incorrect production were *la maschio* ('the.F male.M'), *la fosso* ('the.F trench.M') and *il carezza* ('the.M caress.F') instead of *il maschio* ('the.M male.M'), *il fosso* ('the.M trench.M') and *la carezza* ('the.F caress.F'). The mistake on the word *maschio* is particularly surprising, given that the noun represents a regular case in which natural gender, grammatical gender and final marker *-o* coincide.

Out of the 91 mistakes committed on opaque (simple and derived) nouns, in three cases subjects assign the GG on the basis of the word final wovel *-e* and re-interpret it as a marker for plural feminine: **le peste* ('the.F.Pl pest'), rather than *la peste* ('the.F pest'). In one case, a patient succeeds at retrieving the proper GG and article form, but she

regular nouns, 760 for irregular nouns and 2280 for opaque nouns (of which, 760 are simple opaque nouns and 1520 are derived nouns).

regularizes the word form by changing the final vowel into *-a*, typical of feminine nouns: *la *pesta* ('the.F pest.F') rather than *la peste* ('the.F pest'). The wide majority of mistakes (87 out of 91) is represented by cases of incorrect GG retrieval: *il peste* ('the.M pest') rather than *la peste* ('the.F pest'); among these, 17 mistakes occur on the word *nubile* ('unmarried woman') and could be due to a semantic error[36] rather than to incorrect GG retrieval. I then run an analysis on the remaining 70 cases of incorrect retrieval, in other to verify whether mistakes are evenly distributed across subclasses or affect only specific factors. I consider in particular the opposition between simple and derived nouns, the one between masculine and feminine words and the one between nouns endowed with natural gender (either masculine or feminine) and the inanimate ones.

Table 4.8. Characteristics of opaque nouns affected by errors (in raw numbers).

Word form:	Simple	Derived	
	12	58	$p < .01$
Grammatical Gender:	Masculine	Feminine	
	12	58	$p < .01$
Animacy:	Natural Gender	Inanimate	
	20	50	$p < .01$

According to the distribution of mistakes within opaque nouns, errors are more frequent on derived than on simple nouns ($\chi^2(1,N=2280)=8.52$, $p<.01$), on feminine than on masculine nouns ($\chi^2(1,N=2280)=31.19$, $p<.01$), and on inanimate items than on words endowed also with natural gender ($\chi^2(1,N=2280)=13.26$, $p<.01$).

A similar analysis should be run on irregular nouns, which were overall affected by 40 mistakes. However, before starting with the analysis, I would like to cut out the 12 mistakes occurring on the word *la dinamo* ('the.F dynamo'), often retrieved as *il dinamo* ('the.M dynamo'). They could be cases of GG assignment based on the final marker; nonetheless, causes could also be searched within the dialectal variety spoken by the participants, which might differ from standard Italian with respect to GG. For this reason, I set the word *dinamo* apart and focus my attention on the remaining 28 mistakes. A small portion of incorrect retrieval (3 occurrences) is based on natural gender. Namely, subjects

[36] Speakers of Italian often mistake the word *nubile* ('unmarried woman') for *celibe* ('unmarried man').

produce determiners for irregular nouns basing their assignment process on the natural gender of the corresponding referents, as in the case of **il guardia* ('the.M guard') for *la guardia* ('the.F guard'). In one mistake, the regularization process of irregular nouns affects the noun rather than the determiner: in that case, the target phrase *il pianeta* ('the.M planet') is produced as *il *pianeto* ('the.M planet.M'). All remaining mistakes (24 occurrences) are cases of GG retrieval based on the final word marker. In other words, subjects base their GG assignment on the word form as if nouns were all part of declension Class I or II. Again, I looked for regularities or tendencies among affected nouns, with respect to grammatical gender (masculine vs. feminine) and animacy (natural gender or neuter inanimates), but no significant difference could be detected (see Table 4.9).

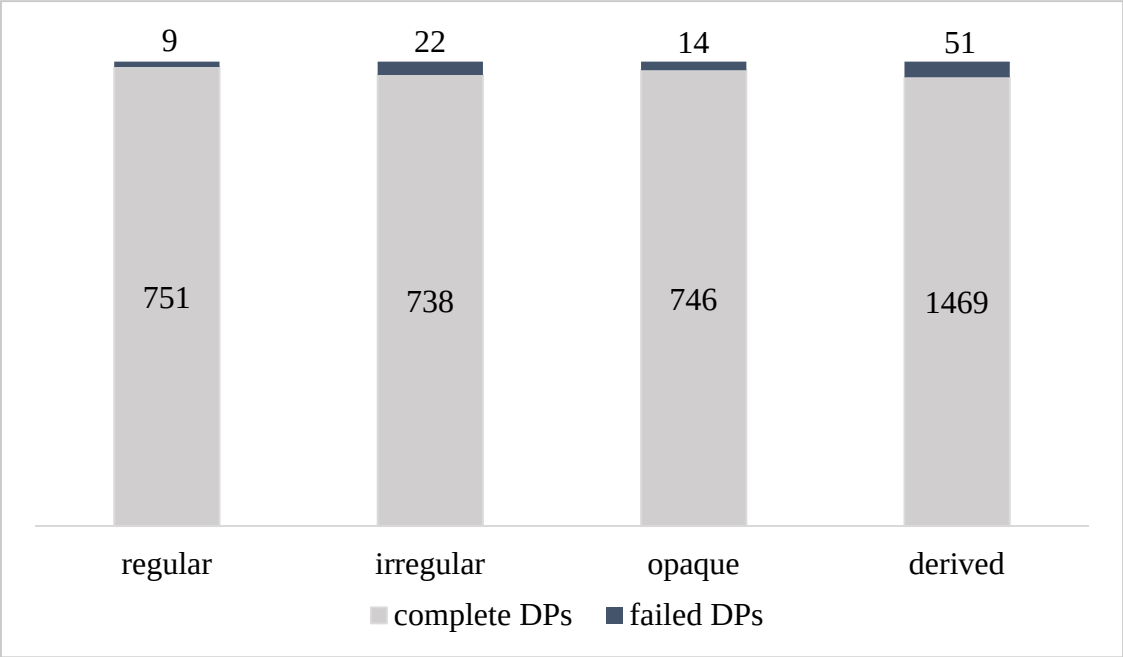
Table 4.9. Characteristics of irregular nouns affected by over-regularization in GG retrieval.

Grammatical Gender:	Masculine	Feminine	
	19	9	$p > .05$
Animacy:	Natural Gender	Inanimate	
	12	16	$p > .05$

In addition to episodes of incorrect GG retrieval, PADs overall also failed at providing with an appropriate output on 96 items. Failed outputs, mainly in the form of bare nouns, distribute across the four noun classes[37], as shown in Figure 4.2.

[37] As for the data in Figure 4.1, please recall that each subject saw 20 regular nouns, 20 irregular nouns, 20 opaque nouns and 40 derived nouns. This is the reason why the overall number of administered derived nouns doubles the number of regular, irregular and opaque ones.

Figure 4.2. Number of successful and failed DP outputs provided by PADs for regular, irregular and opaque and derived nouns.



Unlike the distribution of mistakes across noun classes, failed outputs affect all noun classes at very low rates. The only significant comparison is the one between regular and derived nouns: mistakes are more frequent in the latter than in the former ($\chi^2(1,N=2280)=9.32$, $p<.01$). However, a word of caution is necessary here. Out of the 51 failed outputs on derived nouns, thirteen are produced by the same participant. Thus, the asymmetry is not significant at the group level.

4.4 Discussion

The main and most striking result of the present research is that PADs retain an unimpaired ability to retrieve the proper GG for prompted nouns. Although their performance is different from that of COs (who perform at ceiling) rates of accuracy above 93% do not leave any doubt about the fact that PADs still master grammatical gender retrieval. This claim takes into consideration the overall performance of PADs on all experimental items and is therefore a bit sketchy, but still well-founded. Moreover, results are compatible with the model of lexical storage and retrieval proposed by Levelt (1989), also in the revised version of Caramazza and Miozzo (1997) and much subsequent work (see Friedmann et al., 2013): in these, lexical entries are described as storing

semantic and syntactic information at different levels. The retrieval of semantic information is notoriously impaired in PADs (Chertkow & Bub, 1989; Hodges et al., 1991; Almor et al., 2009), while GG retrieval seems not to be in the present data, thus providing a dissociation between the two informational components attached to the lexical entry. Data from the present study confirm that PADs efficiently retrieve morpho-syntactic information from the lexicon. It is worth underlining that the task did not implicate explicit retrieval of semantic information, thus potentially allowing for independent activation of pure syntactic ones. It might be the case that word activation works better for PADs, when they are relieved from the burden of semantic component retrieval. The hypothesis of a dissociation between syntactic and semantic information in PADs certainly deserves further investigation. Therefore, I leave for future investigation the project to realize a study made of combined tasks in order to check both for semantic and for GG retrieval on a fixed list of experimental trials.

However, a more fine-grained look at incorrect outputs in the present study can reveal a few interesting tendencies in PADs performance: these might point out which factors represent potential sources of difficulties and which strategies subjects adopt in order to cope with the process of GG retrieval. These are discussed in what follows. I will start by focusing on cases of failed retrieval in order to show that it occurs randomly (4.4.1). I will then switch to incorrect answers, in particular those concerning opaque nouns (4.4.2). A particular attention will be devoted to the role played by derivational morphology (4.4.2.1) and animacy (4.4.2.2), and to the opposition between feminine and masculine GG (4.4.2.3). In Section 4.4.3, I will discuss mistakes on irregular nouns. Finally, in 4.4.4, I will draw a comparison between the present data and results from previous studies on Italian-speaking aphasic speakers and PADs.

4.4.1 Failed retrievals

At first sight, the number of failed productions of complete DPs might look quite surprising, but, actually, it is not. PADs experience great difficulties at keeping in mind the requirements of the task they are completing and the kind of behavioural responses they are expected to produce. The experimental material was designed so as to reduce the risk of forgetting how to perform it, a situation that can produce high frustration in the subject. The measures taken in order to help participants remember how to perform the

task were effective, but still they could not completely avoid the risk for non-target performance. The 96 missing productions were mainly due to the spell-out of bare nouns or to phonological errors on the target nouns, in other words participants read aloud the noun, but did not produce any determiners or produced a non-target noun. Overall, the analysis on failed items revealed that occurrences equally distributed across the four main noun classes (regular, irregular, opaque and derived nouns). This confirms the hypothesis that failures were accidental, due to short attention span for the task or to misreading of the prompted nouns; they were in no way determined by specific complexity factors. If any complexity factors were the real cause for failed output, this could be identified thanks to a higher number of failures on a specific class. In other words, if the problem were derivational morphology, the number of failures on derived nouns would significantly exceed the number of failures on simple nouns. However, this is not the case, which means failures were purely accidental.

4.4.2 Incorrect retrievals on opaque nouns

The error analysis offers insights into what happens when GG retrieval fails at activating the expected gender, in particular in the case of opaque words. Mistakes on grammatical gender retrieval were collected and classified and it emerged that they mainly concern irregular and opaque nouns, with a higher incidence on derived ones. PADs produced 134 mistakes; of those, few concern regular nouns (n=3). The very low amount is therefore statistically different from the number of mistakes recorded on irregular (n=40) and on opaque nouns (n=91). In contrast, words in the irregular and opaque classes were equally effected in rates that, according to the statistics, are not significantly different. If we now think of the double-route to GG activation described in paragraph 3.4, namely the direct access to the gender information stored in the lexical entry and the form-based retrieval, we can actually see that the latter is well preserved and guarantees a very accurate performance, while the former shows slight signs of impairment. The two complementary routes, due to their core functional characteristics, certainly differ with respect to the kind of linguistic components they represent. Direct access is an example of declarative knowledge, while form-driven retrieval goes under the procedural system. Roughly speaking, the two components are not balanced in PADs, as the cognitive impairment affects declarative components first and to a more severe extent, while leaving procedural

components relatively spared. This generalization holds for a variety of cognitive abilities, and linguistic competence makes no exception. The data at stake in the present chapter further support the generalization outlined above by showing that procedural retrieval of GG is well preserved and assumes the role of default strategy. Patients indeed produce mistakes on irregular or opaque words by regularizing them, rather than the contrary. We can therefore assume that PADs tend to resort to this procedure when unable to recover the target GG directly form the mental lexicon. The hypothesis is also compatible with the reverse pattern of regularization, namely the one exemplified in *la *pesta* instead of *la peste* ('the.F pest'): in this case, correct retrieval of GG takes place and affects the opaque word by changing the final vowel *-e* into *-a*, the typical final marker for singular feminine nouns. No matter then in which direction the regularization process flows; in any case, the tendency to rely on the procedural mechanism of regularization clearly emerges in PADs.

4.4.2.1 Sensitivity to derivational morphology

If I now take a further step and have a closer look at error analysis, I can actually infer more on how over-regularization takes place and the value assigned to morphological markers. In particular, they are mistakes on opaque nouns to be informative about which factors represent a source of complexity.

First of all, the incidence of mistakes on opaque nouns is higher for derived nouns than for simple nouns, thus attesting that patients are sensitive to derivation. This agrees with previous studies on unimpaired (Meinzner et al. 2009) and on aphasic speakers. Unimpaired speakers are indeed sensitive to the number of derivational steps involved in word formation and they differently process words depending on their level of derivational complexity (one- or two-step derivation). In particular, a higher number of derivational steps mirrors into the activation of wider brain substrates.

Although low frequency of derived nouns might have influenced the results (thus representing a confounding factor), an interesting comparison with agrammatic speakers is possible. Based on similar material, Luzzatti and De Bleser (1996) show that aphasic speakers are also sensitive to derivation and even profit from it. Their participants better retrieve GG for derived nouns than for simple ones, probably as a by-product of their capacity of cutting down words into their component morphemes and of recognizing the

specific GG entailed in each derivational morpheme. PADs show a similar sensitivity, but in the opposite direction, in the sense that GG retrieval is more problematic for derived nouns than for simple opaque nouns, which nonetheless can be taken as a sign of sensitivity to derivation.

Moreover, the impairment at processing derived nouns and retrieving GG from the derivational suffix resembles results from a previous study conducted by Chiarelli, Manichelli and Semenza (2007) about patients' ability to deal with compounds. Chiarelli et al. (2007) report that PADs make more errors in picture-naming with compounds rather than with simple words and, in particular, the impairment affects the second element in the compound, while the first one is more often smoothly retrieved. As a compensating strategy, PADs tend to overcome their difficulties by systematically using the most productive structure for compounds, namely Verb-Noun.

Taken together, results from the present study and from Chiarelli et al. (2007), reveal that PADs suffer from a cumulative effect and their impairment increases along with the level of internal complexity entailed in the word structure. In particular, the difficulty at retrieving GG from the derivational suffix resembles the attested difficulties at activating the second element in compounds.

4.4.2.2 Role of animacy

One more question I addressed at the beginning of the present chapter concerns the role played by natural gender. In other words, I wondered whether natural gender can facilitate GG retrieval and therefore produce an asymmetry between nouns endowed with the [+human] feature and the inanimate ones, namely artefacts or abstract nouns. Indeed, the prediction was met and the number of mistakes on inanimate nouns (n=50) significantly exceeded the number of mistakes on nouns with natural gender (n=20). I assume that this is due to the fact that PADs can resort to semantic information when direct access to GG is impaired[38].

That might appear to contradict the assumption that PADs are highly impaired at retrieving semantic information (Hodges et al., 1991; Almor et al., 2009), but semantics is not all equally affected by disruption. Impairments proceed from fine-grained and

[38] In Chapter 3 I outlined the distinction between natural gender and grammatical gender and underlined that assignment of the latter does not depend on the former. However, NG and GG very often overlap, such that a GG retrieval strategy based on natural gender can be highly successful.

detailed information to information that is more generic (see Chapter 2). In this view, biological gender can be considered as a macro-category from the semantic point of view, which probably holds rather spared through different stages of impairment and is available as a relevant piece of information in case the form-based procedure is not available and direct access to GG fails.

In this line of reasoning, I would also like to add one more observation, namely the fact that the asymmetry between words with natural gender and others with no natural gender might be due to an effect of concreteness, which favour the former over the latter. Indeed, PADs have been reported to be more proficient at retrieving nouns rather than verbs (Almor et al., 2009; Bushell & Martin, 1997; Grossman et al., 1996; Robinson et al., 1996), and researchers ascribed the difference to the fact that the former is more concrete than the latter from the semantic point of view. Here I consider the hypothesis that the same difference might hold between nouns with either natural or no natural gender and that this might represent one further factor playing a role in the results.

4.4.2.3 Masculine as default grammatical gender

The third thing I can notice with respect to mistakes on opaque words is that feminine nouns are more affected by mistakes on GG retrieval than masculine nouns. In other words, cases of feminine nouns incorrectly classified as masculine, are more frequent than the opposite. This might be due to frequency asymmetries in the two classes of nouns, as it was not possible to balance feminine and masculine derived nouns for frequency. However, it could also be the case that masculine grammatical gender is regarded as a sort of default gender, to which PADs resort in case of difficulties at retrieving the proper GG for opaque words. When faced with opaque nouns, speakers can in no way retrieve the proper GG through the word form, as no morphological cue is available; direct access to the syntactic information is indeed the only possible strategy, but when PADs cannot use the latter, they must activate a last-resort strategy, which consists in assigning a default gender, namely masculine.

If the assumption is correct, we should also take into account the fact that (most probably) not all masculine nouns classified as such are the outcome of successful retrieval of GG gender from the mental lexicon. It might be the case that masculine was assigned to some items as a default value, which corresponds to the correct one just by

accident. I intend to assume that masculine and feminine are affected by unsuccessful GG retrieval to the same extent; what changes is the fact that the use of a default gender leads to an apparent successful classification in the case of masculine nouns. In contrast, in the case of feminine nouns, default assignment causes mistakes, which are more salient. Otherwise, we should assume a dissociation between masculine and feminine nouns, with feminine nouns evenly stored separately from masculine ones and for some reasons more affected by retrieval impairment than their counterparts are.

Unfortunately, I cannot further test the assumption with the data at hand, so that the proposal can be considered only at a pure speculative level.

4.4.3 Incorrect retrievals on irregular nouns

Finally, no equivalent effect for grammatical gender (feminine vs masculine) and biological gender (natural vs neuter) could be detected in the list of words included in the irregular class, although for this group of nouns the incidence of mistakes was comparable to the one registered for opaque nouns. Irregularities and information clash between morphological cue, natural gender and grammatical gender represent the major source of difficulty *per se*, and do not leave any space for effects played by other minor factors. Indeed, all mistakes (41) committed on irregular nouns are of the over-regularization kind, with GG assignment based on the words' final marker. Indeed, feminine names with final marker *-o* (*la radio*) were often assigned masculine GG (**il radio*); while masculine names with final marker *-a* (*il pianeta*) were assigned feminine GG (**la pianeta*), thus reinforcing the claim that over-regularization is the main trend to be detected in patients' performance when it comes to coping with impaired linguistic components of the declarative-knowledge type.

4.4.4 A comparison with previous studies on aphasia and PADs

At this point in the discussion, it is interesting to draw a comparison between the collected data and those found in previous studies on aphasia and on PADs. As underlined in Chapter 3, a direct comparison across languages with respect to phenomena concerning GG is highly problematic because GG systems widely differ across languages. The comparison is therefore limited to studies that concern Italian-speaking participants.

With respect to aphasia, the present data confirm the results of previous studies, but also offer the possibility to highlight at least one interesting difference. As in Badecker et al. (1995) and in Luzzatti and De Bleser (1996), I find an overall spared access to GG in regular nouns.

As expected, one further characteristic that PADs share with aphasics is an enhanced difficulty at retrieving the proper GG for opaque words. In both cases, speakers with impaired cognitive abilities struggle at dealing with words characterized by the final marker *-e*. However, an interesting observation emerges from the comparison between the present study and the one by Luzzatti & De Bleser (1996). The two aphasic speakers tested in the latter showed clear signs of performance improvement in the case of derived nouns, meaning that they could analyse words into their morphologic components and recognize derivational morphemes that are specified for gender although ending in *-e*. In contrast, the present study shows the reversed pattern: PADs do not benefit from the presence of derivational morphology marked for GG and even experience more difficulties in correspondence to derived nouns. This means that PADs are sensitive to the internal complexity of derived nouns, but they do not succeed in exploiting the information on GG.

With respect to previous studies on Italian-speaking PADs, the present results seem to contradict those presented by Paganelli et al. (2003) at first. In the cited study, authors do not find signs of gender activation when PADs experience TOT-states, while they found them in normal controls. This might be read as a sign of impaired activation of GG. However, I agree with the author in claiming that the effect might derived from a more basic source, namely a failed activation of the correct conceptual and semantic information. The discrepancy between the results in the present and in the cited study should therefore be ascribed to the difference between the tasks in use. In Paganelli et al. (2003), the performance required the complete retrieval of lexical items, starting from its conceptual components to the phonetic realization. In contrast, the task in use in the present study provided the lexical item through visual input, thus activating the process of lexical retrieval from the opposite direction, which results in a lower involvement of semantic activation for task completion.

Finally, the data can provide further support to the results discussed in Manenti et al. (2004), in which PADs show sensitivity to gender priming in lexical retrieval. The two

studies taken together argue for a (rather) spared status of GG in the linguistic competence of patients with Alzheimer.

4.5 Conclusions

In the present chapter, I discussed data from a grammatical gender retrieval task performed by 38 PADs and 21 COs (healthy elderly speakers). Overall, the performance of PADs was not at ceiling as it was for COs, but still, levels of accuracy above 93% confirmed that the process of GG retrieval is well mastered by PADs. Therefore, PADs are not impaired at retrieving syntactic information attached to lexical entries. Overall, mistakes were as low as 3.5% of the total amount of experimental trials. In particular, PADs were very accurate on regular nouns, which suggests that the form-based route to GG access is preferred over the direct access to the information. This claim is further supported by the fact that mistakes affecting irregular nouns were all of the over-regularization kind, with GG assigned in consideration of the final word-marker. The two observations taken together reveal that PADs crucially rely on the procedural mechanism, and they resort to this when they face difficulties with irregular nouns.

In contrast, despite the low number of registered mistakes on opaque nouns, these were informative about slight tendencies within patients' performance. Indeed, in this category of nouns, mistakes prevailed on derived rather than on simple nouns, on nouns with no natural gender rather than on nouns with natural gender and on feminine rather than on masculine nouns.

The first effect is due to the fact that PADs might perceive the higher complexity of derived nouns and be incapable of retrieving the correct GG from the derivational suffix. Different frequency rates might have played a role in this, however the asymmetry opens to a comparison with the agrammatic speakers in Luzzatti and De Bleser (1996), who actually performed better on derived than on simple nouns (on a similar list of items).

Furthermore, PADs seem to tend to restore to other strategies when direct access to GG fails and final marker *-e* does not offer any cue for form-based retrieval. One of these strategies builds on natural gender as a spared semantic category in PADs, who therefore are more successful on nouns endowed with natural gender than on nouns depicting artefacts or abstract concepts.

When natural gender is not at hand, PADs finally restore to masculine as a default gender, and assign it to opaque nouns, for which no other route to correct retrieval can be followed. For obvious reasons, this last strategy is successful with masculine nouns, but not with feminine ones, which are affected by a higher rate of mistakes.

In conclusion, I would like to underline once more that the described phenomena are minor tendencies, as GG retrieval is actually well preserved in PADs. The low number of mistakes does not allow for any sound claim with respect to the potential presence of forms of impairment. This further confirms that speakers store semantic and syntactic information at separate levels and can suffer from a dissociated impairment, as in the case of PADs, who are spared at the latter but not at the former.

5 RELATIVE CLAUSES AND Wh-QUESTIONS: AN OVERVIEW

5.0 Introduction

In the present chapter I will illustrate some relevant characteristics of Wh-questions and relative clauses, the two kinds of clauses that I will use in order to sample the capacity of PADs to process complex sentence structures. The main reason to include the two clause types in the study is that they both entail a phrasal movement to the higher functional field of the sentence. The movement takes place in order to satisfy requirements of the scope-discourse type, which are determined in turn by the specific semantic of the sentence to be realized, either a question or a modification of a nominal phrase (as in the case of restrictive relative clauses). Previous research on aphasia (Avrutin, 2000; Friedmann, 2002; Garraffa & Grillo, 2008; Grillo, 2008, a.m.o.) and on L1 acquisition (De Vincenzi et al., 1999; Friedmann et al., 2009; a.m.o.) pointed out that syntactic derivations are problematic when they imply A-bar movement to the CP-layer. Therefore, they turn into privileged points of observation for the evaluation of the grammar used by speakers. Precisely, the two structures entail an instantiation of the so-called Wh-movement, which is responsible for the movement of a phrase to the CP-layer and for leaving a gap in the clause. However, the semantic function of the operators that allow the derivations are different. In the case of Wh-questions, the operator represents a variable in the sentence, which must receive a value. In the case of Relative Clauses, the clause contributes to the restriction of a nominal element in the matrix clause. Besides, Wh-questions can be realized both in the form of root questions and in the form of embedded clauses, while relative clauses are subordinates, due to their nature of being attached to an argument in the matrix clause. Therefore, Wh-questions and Relative Clauses share some structural properties while differing on others, a combination which allows for different observations and comparisons.

The present Chapter is organized as follow. In section 5.1 I recall the basic mechanism allowing for the derivation of both Wh-questions and RCs. Sections 5.2 and

5.3 deal with Wh-questions from a structural point of view and review relevant works in the syntax of Wh-questions. The same procedure is followed in section 5.4 for Relative Clauses. The second part of the Chapter is devoted to empirical studies on the processing of both structures and how these contribute to the linguistic debate. In section 5.5 I review relevant works on the processing of RCs both in adults, in children and in aphasic patients. In 5.6 it will be the turn of experimental studies on Wh-questions. The reasons for this order of presentation is that the amount of experimental work run on the processing of RCs not only quantitatively exceeds the one on Wh-questions, but it is also the case that the former often influenced and guided the latter. Finally, section 5.7 will provide with a summary of the relevant information for the design of the experiments on the comprehension of Wh-questions (Chapter 6) and RCs (Chapter 7) in PADs.

5.1 Wh-movement

In a seminal work presented in 1977, Chomsky crucially pointed out that what superficially looks like a variety of structures available in the grammar, actually reduces to different applications and instantiations of a singular mechanism, namely Wh-movement. Until that moment, the syntactic analyses of different structures (for instance direct and indirect Wh-questions, Relative Clauses, Clefts, and Topicalizations) had built on a variety of transformational rules, each specialized for the derivation of only one kind of sentence. This attitude radically changed after Chomsky's (1977) proposal.

The intuition about the presence of a unified account to the different structures started from the observation of the characteristics shared by the sentences at stake. First, the cited structures entail movement of a phrase to a position higher than the one that would correspond to their original site in an unmarked declarative sentence. This is described as an instantiation of extraction, which leaves a gap in the internal structure. The gap also corresponds to the point in which the fronted element is actually interpreted by the parser. Second, the extracted phrase can move from the argument position of an embedded predicate to the highest matrix clause through cyclic movement across the CPs of multiple nested clauses. The operation is allowed by the presence in the nested clauses of verbs like *to think*, *to wonder*, *to order*, etc.. Third, extraction is regulated by few constraints, among which, the presence of complex nominal phrases and of Wh-islands. Complex NPs represent sorts of barriers, which do not allow for sub-extraction of an

element, as well as questions do. Indeed, the presence of a Wh-element in the CP of an indirect question blocks the movement of a second Wh-element outside of the embedded clauses. Chomsky (1977) recognizes these characteristics in all structures cited above and concludes that they must be the result of a unique transformational rule, namely Wh-movement. The label is due to Wh-questions, in which movement is signalled through the presence of a dedicated element of the Wh-kind. However, a close look at relative clauses, clefts and topicalizations reveals that the same characteristics apply to these structures too, despite the absence of a lexicalized operator. Chomsky's (1977) proposal represented a turning point in Generative Grammar because it highly influenced the forthcoming syntactic research by setting a new goal, namely the search for unified mechanism behind structures and phenomena that superficially and cross-linguistically appear unrelated.

The parallelism between Wh-questions and Relative Clauses has been further pursued in subsequent research, however, the present work does not primarily aim at discussing to which extent the parallelism between the two structures is to be pushed; for this reason, the following sections consider Wh-questions and relative clauses in turn, by summarizing the milestones of the research on the former and on the latter.

5.2 Wh-questions: a classification

In the previous section I recalled that Wh-questions are realized through the fronting of an element, but languages actually differ with respect to the nature of this movement. For instance, Italian is characterized by overt movement of the interrogative element, while other languages, like Chinese and Japanese are not (Cheng, 1997; Cole and Hermon, 1998, a.o.). In the latter, the Wh-element is pronounced *in-situ* and, if movement is assumed, this takes place at LF. The asymmetry between in-situ and non in-situ languages has often be interpreted as a parametric one, although it has also been proposed that the difference might actually stem from the characteristics of the Wh-element in use. According to Cole and Hermon (1998), it depends on whether the Wh-element can lexicalize both the operator and the variable. In case these are lexicalized in the same element, overt movement takes place. In contrast, languages like Chinese and Japanese have a null operator directly merged at CP, so that the Wh-element lexicalizes only the

in-situ variable. Italian clearly belongs to the first class of languages, in which overt movement to the CP-layer always takes place.

Apart from the nature of the movement, the classification of Wh-questions distinguishes between Yes/No questions and constituent questions, the latter further subdivide into argument and adjunct questions. In Yes/No questions, the interrogative operator takes scope over the complete sentence (1a), so that the expected answer can have only either a Yes or a No form (as the clause label suggests). In contrast, in constituent questions the interrogation concerns one specific phrase in the sentence, which in turn can correspond either to one of the predicate's arguments (e.g., the subject in 1b) or to a (temporal) adjunct to the sentence (as in 1c).

(1) a. Did Paul arrive early in the morning?
b. Who arrived early in the morning?
c. When did Paul arrive?

Interrogatives can also be realized as indirect questions in embedded clauses (2). In the case of Yes/No questions, the interrogative operator is lexicalized through specialized complementizers (2a); while for constituent questions, the same Wh-element appears in root (1a to 1c) as well as in indirect Wh-questions (2a to 2c).

(2) a. I wonder [*if*/*whether* Paul arrived early in the morning]
b. I wonder [*who* arrived early in the morning]
c. I wonder [*when* Paul arrived]

The following discussion mainly covers previous literature on root argument questions (1b), because this is the kind of interrogative in use in the present study (see Chapter 6). However, occasional references to other forms of questions will be brought up whenever it will be relevant for the discussion.

5.3 The derivation of Wh-questions

A seminal work for the comprehension of how Wh-questions are derived is the one in which Rizzi (1996) argues for the existence of a Wh-criterion, responsible for the specific

configuration of questions. Rizzi proposes that a [+wh][39] feature marks all questions and constrains the way the sentence must be derived. The proposed Wh-criterion (Rizzi, 1996:64) consists of two mutual constraints:

a. A wh-operator must be in a Spec-head configuration with an $X°_{+wh}$
b. An $X°_{+wh}$ must be in a Spec-head configuration with a wh-operator.

The presence of a [+wh] feature on the complementizer layer determines the nature of the sentence and, at the same time, it establishes that the sentence is well-formed only if the operator and a head endowed with the same interrogative feature enter into a Spec-head configuration.

Based on the observation that in English the only element which is allowed to appear directly to the left of a Wh-element is the verb (and no other constituent is allowed to intervene between the Wh-element and the verb), Rizzi claims that the verb is endowed with the [+wh] feature and it must enter the required Spec-Head configuration. At this point, once assumed the presence of the Wh-phrase in SpecCP, a second assumption must be done, namely the movement of the inflected verb to the CP. In its first formulation, the Wh-criterion predicts the presence of the Wh-element in SpecCP and of the Verb in C° in order to satisfy the required Spec-Head configuration. In these terms, the proposal also accounts for the Subject-Auxiliary Inversion (SAI) that characterizes root questions in English. Rizzi's analysis (1996) starts from the observation of English root Yes/No questions. Sentences entail a configuration that the author defines as a form of residual V2, because it consists in the movement of the auxiliary to the C° position (3):

(3) a. (*you) did you arrive this morning?
b. When (*you) did you arrive?

Both in the Yes/No question and in the adjunct question in (3), the subject can neither precede the auxiliary nor intervene between the Wh-element and the auxiliary. It is therefore assumed that the subject stays in a lower position within the IP, while the

[39] Authors alternatively refer to the feature for Wh-questions as [+wh] or [+Q]. I will adopt [+Q] later in the study (see Chapter 6). However, Rizzi's proposal is here represented in its original version with [+wh].

auxiliary moves from I°, the position actually endowed with the [+wh] feature, to C°, the position where the Wh-criterion is to be satisfied.

Building on the fact that Italian shows the same behaviour with respect to subject position in interrogatives, namely the fact that subjects follow the finite verb rather than preceding it, Rizzi (1996) extends his analysis to Romance languages too. However, the proposal has not received unanimous consensus. Rather, it has opened a debate on Romance interrogatives, which covers three different aspect of the derivation of Wh-questions: *i*) are Romance questions instantiations of CPs or of IPs? *ii*) what is the landing site of V in interrogatives?, and *iii*) what is the position of the subject in non-subject interrogatives?

For instance, Barbosa (2001) rejects the idea that Romance questions are CPs and claims that a structure starting with the IP node is sufficient in order to account for the derivation of Romance Wh-questions. Her claim builds on the observation of the positions that subjects can occupy in Romance interrogatives, in comparison to the positions they occupy in Romance main declarative clauses. Barbosa points out that in Italian interrogatives, the subject can intervene neither between the Wh-operator and the verb[40], nor between the Auxiliary and the Past Participle. The latter holds true for declarative sentences too. For instance, the subject of Italian interrogatives can follow the past-participle, can precede the interrogative as a hanging topic or can be marginalized to the left of the interrogative (Antinucci & Cinque, 1977).

The parallelism between interrogative and declarative sentences with respect to the position of the subject is confirmed by other Romance varieties: in Spanish, for instance, the subject is tolerated in between the Auxiliary and the Past Participle only with the auxiliary *estar* and with the imperfect past form of *haber* (have), namely *había*. Whenever the auxiliary is in a different tense form (e.g., the present *ha* or the future *habrá*), the subject is banned from that position and it is allowed to appear only either before the Auxiliary or after the Past Participle. The reasons for this are not relevant here; what actually matters is the absence of asymmetries between declarative and interrogative clauses with respect to subject positions. Building on this observation, Barbosa (2001) rejects the idea of using subjects as a diagnostic tool for I° to C° movement.

[40] In Italian, the subject can intervene between the Wh-operator and the verb only in the case of embedded interrogatives with a verb at the subjunctive. In this case, the [+wh] feature is represented in C° and not in I°, such that verb movement from I° to C° is not implemented (Rizzi, 1996).

In her view, it is not possible to assume verb movement to C° since there is no overt sign of the presence of the verb in the CP. Along this line of reasoning, Barbosa does not see any reasons for assuming that the Wh-criterion is satisfied at the CP level; rather, she assumes that the relevant Spec-head configuration between the Wh-phrase and the verb takes place in IP. Her proposal posits the presence of the Wh-phrase in SpecIP in order for this to enter in a relation with the element carrying a [+wh] feature in I°. Given that Rizzi's (1996) proposal does not set any strict constraint on the exact position where the Spec-Head agreement is to be realized[41], Barbosa actually leaves unchanged the core mechanism of the Wh-criterion. Several consequences arise at the syntactic level though. The main issue concerns SpecIP and the nature of this position: in order to account for the presence of a Wh-phrase in SpecIP, it is necessary to assume that this is a position of the A-bar kind, able to host operators for scope-discourse features. This observation unveils a curious fact: Barbosa (2001) rejects verb movement to C° because of the lack of overt signs of movement, while she assumes Wh-movement to SpecIP, despite clear evidence against it. The main reason why Romance interrogative cannot be assumed to be root IPs is that SpecIP plays a relevant role in subject-verb agreement and therefore it cannot be exploited for other purposes. In particular, SpecIP is the landing position for *pro* in Italian. An example will help illustrate the point at issue:

(4) *Cosa hai portato*?
What Aux_2ps bring_PTCP
'What did you bring?'

If we analyse an object question with a null subject of the kind in (4) and claim that the Wh-phrase is in SpecIP, we automatically have to imply that *pro* occupies a different position, namely a post-verbal one. From this assumption, a clear clash arises, due to the fact that *pro* has been known for a long time to align with preverbal subjects rather than with post-verbal ones (Cardinaletti, 2004; Guasti, 1996a). A variety of observations supports the claim.

[41] Strictly speaking, Rizzi (1996) does not explicitly claim that the Wh-criterion must be satisfied in the left-periphery; however, syntactic evidence points to this conclusion.

Guasti (1996a) in particular remarks the relevance of the following observations: first, floating quantifiers can appear in between the auxiliary and the past participle only in the case of preverbal subjects and of *pro*, but not with post-verbal subjects (Rizzi, 2000). Second, the subject of a main clause can control the subject PRO of a temporal adjunct only if it is null or if it overtly appears in the preverbal position; in contrast, post-verbal subjects cannot control PRO in temporal adjuncts.

The detailed distribution of Italian subjects is certainly beyond the purpose of the present structure. However, for the sake of the present discussion it will be sufficient to follow Guasti (1996a) and assume that a null subject occupies the upper-portion of the IP and that this position cannot correspond to the landing site of Wh-operators. Through a careful analysis of a variety of phenomena, Guasti concludes that wh-operators must appear within the CP layer in order to satisfy the relevant [+wh] feature and they are not allowed to appear in IP. With respect to the position of the Wh-operator, her analysis is therefore identical to the one originally formulated by Rizzi (1996) for English interrogatives. However, the author also admits that no overt sign of verb movement to the CP layer is to be detected in Italian, an observation that very much resembles the main argument brought up by Barbosa (2001).

A relevant contribution to the issue of V movement to C in Romance languages comes from the analysis conducted by Poletto (2000) on Italian varieties. If standard Italian fails to provide the desired information, some indirect proof can be derived from the analysis of varieties of Northern Italian dialects. In particular, the author presents two clear cases in which the verb has moved to the CP layer of interrogative clauses.

The first piece of evidence comes from Fassano dialect, a Rhetoromance variety, characterized by Subject clitic inversion in root questions. The variety also disposes of a focus particle *pa*, which signals the interrogative character of the clause and presumably marks the lower boundary of the CP layer. The interesting piece of information for us is that the verb and the clitic subject can appear (inverted) in a position higher than *pa*. The example in (5) clearly signals that the verb has moved across the focus marker and has landed in the CP layer (Poletto, 2000:46):

(5) *Co l fas-to pa?*
How it do-you interr. Marker
'How do you do it?'

Second, Poletto (2000) finds instances of do-support phenomena in Eastern Lombardian varieties. In the dialect of Monno (Eastern Lombard), interrogatives imply the insertion of an auxiliary in the higher functional portion of the sentence, while the lexical verb is lower in the structure (Poletto, 2000: 49). As can be observed in (6), the choice of the auxiliary *fare* ("do") and its position in the sentence clearly resemble the do-support phenomenon typical of English, a language for which V to C movement is widely assumed (Rizzi, 1996).

(6) *Come fa-l comportas*?
How does-he behave-himself
'How is he behaving? '

The two pieces of evidence for V to C movement in interrogatives in Northern Italian Dialects are only an indirect proof for the variety at stake here, which does not allow for certain conclusion with respect to standard Italian. Nonetheless, along with the observations concerning the impossibility of inserting wh-operators as low as the SpecIP position, their presence clearly points towards an analysis that assumes V to C movement.

Summing up, with respect to the first two questions concerning the analysis of Romance questions (see above), the present work assumes that standard Italian root interrogatives are CPs nodes, in which the higher functional field (the CP) hosts both the wh-operator and the verb. In particular, in the case of root argument questions the Wh-element will target the Focus position in CP[42] (Rizzi, 2001).

As for the position of the subject in non-subject interrogatives, it will be assumed that this, when lexically realized, will stay in SpecvP and check its features against the verb from there (Belletti and Guasti, 2015); while a *pro* is generated in correspondence of SpecIP.

[42] A complete taxonomy of the different positions that Italian Wh-operators can target in CP depending on the kind of interrogative to be realized (root/embedded, Yes/No questions and argument questions; Why questions) can be found in Rizzi (2001).

5.3.1 Bare Wh-elements vs WhichNP

In order to complete the overview on the syntactic characteristics of Wh-questions, it is necessary to consider one further factor, namely the kinds of Wh-element in use. For instance, Italian has two kinds of Wh-elements for argument questions: either bare *Chi* ('who') and *Cosa* ('what'), or *Quale NP* ('which NP'), namely a wh-element endowed with a lexical restriction. The two classes of elements are not equivalent though, as they realize different functions.

Pesetsky (1987) first observed that the English counterparts of the two kinds of elements behave differently. First of all, the author pointed out that in case of English multiple wh-questions, WhichNP phrases do not obey the Superiority Condition, while Who and What strictly do:

(8) Superiority Condition
In a multiple interrogation, where a wh-phrase is in Comp and another is in situ, the S-structure trace of the phrase in Comp must c-command the S-structure position in the Wh-in-situ. (Pesetsky, 1987: 104).

The claim derives from the following examples, in which movement of the interrogative subject phrase to CP is always allowed (in 9a and 10a). By contrast, movement of the interrogative object phrase across an interrogative subject phrase left in-situ is allowed only in case the wh-element in use is of the WhichNP type (10b), otherwise the structure is very marginal (9b):

(9) a. Who did you persuade to read what?
b. ?? What did you persuade who(m) to read?

(10) a. Which man did you persuade to read which book?
b. Which book did you persuade which man to read?

Pesetsky (1987) ascribes the asymmetry to specific properties of the two elements in use. He suggests to distinguish the elements on the base of their relationship with the discourse set. In particular, WhichNP phrases usually refer to a set of elements already present in

the discourse; they are D(iscourse)-linked because they anchor the interrogative to the information shared by participants in the discourse and bias the answer, by suggesting that this should be retrieved among the elements already mentioned. On the other hand, bare Wh-elements of the Who/What type do not operate the same kind of presupposition and leave the answer open.

Rizzi (1990) also points out the relevance of referentiality in the formation of Wh-questions, although his observations substantially diverge from the one presented in Pesetsky (1987). Rizzi specifies that the notion of referentiality should be anchored to thematic roles. He distinguishes between *theta*-roles, which identify referents taking part in the event, and elements that further characterize the event (e.g. manner, measure). The former are proper arguments, while the latter only qualify as quasi-arguments. The distinction generates the dichotomy between arguments and adjuncts. In Rizzi's view, arguments always receive a referential index through *theta*-role assignment, which allows for long-distance movement of the phrase. Quasi-argument do not receive *theta*-roles able to license a referential index. It follows that they obey different movement constraints. Only arguments that receive a referential index are properly bound by the predicate, while quasi-arguments are not and must undergo a more local kind of movement.

Cinque (1990) builds both on Pesetsky (1987) and on Rizzi (1990), and suggests a sharp distinction between WhichNP and Who interrogatives in syntactic terms. The author limits the use of referential indexes to elements present in the discourse; it follows that only WhichNP phrases receive an index, because they refer to members of a set, known both to the speaker and to the hearer. In this sense, quantifiers of the Who/What kind do not bear an index. This distinction in indexes has crucial consequences on the kind of syntactic chains that the elements can enter. WhichNP phrases enter binding relations, while Who/What quantifiers enter in chains of the antecedent-government kind. The syntactic formalization is rather outdated in its terms, but a number of experimental studies has nonetheless resort to it in order to account for speakers' asymmetric performance on questions of the WhichNP and of the Who type. We will therefore go back to this issue in the following sections; for the moment, it should only be considered that the two kinds of Wh-elements can generate different presuppositions, which have to be integrated in the processing of interrogatives. In case of a lexical restriction, the number of represented features on the phrase is higher and the set of alternative answers

is limited to the elements known to the participants in the conversation. A lower number of features is represented on Who/What quantifiers and these do not set requirements on the expected answers (except for the distinction between human and non-human referents).

The distinction between bare and lexically-restricted Wh-elements represents one of the main factors in the empirical studies on Wh-questions; the topic will therefore be further discussed later in the present chapter.

5.4 Relative clauses: syntactic issues and structural analysis

Starting from the '70s, linguists have devoted much work to the investigation of relative clauses. Much has been written and discussed since then, while the interest in the topic increases rather than fading away. A complete overview of all the proposed accounts and observations, as well as answers to still open questions, exceeds the goals of the present section. I will therefore limit myself to touch upon the fundamentals, with the mere purpose of putting down some guidelines, which will help in the understanding of how I designed the task on RC comprehension in use with PADs (see Chapter 7), and how the data collected can be interpreted.

I will start by providing a definition of Relative Clauses and by illustrating basic issues that have provoked and inspired linguists in their research.

Relative clauses play the role of modifiers of a constituent, typically of a noun phrase; they differ from adjectives for being syntactically complex structures. The realization of a RC involves an operation of abstraction (Bianchi, 2002) that connects the modified element in the matrix clause (underlined in 11) to a position internal to the relative clause (the empty category). Previous literature has labeled the former as "head" of the relative, while the latter is the relativization site:

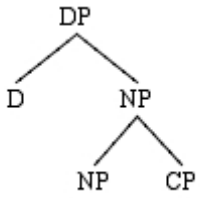
(11) <u>The student</u> that the professor helped *e* comes from Italy

The result of this operation is that the noun phrase *the student* plays a double role, it contributes simultaneously to the meaning of the matrix and of the relative clause: it is the subject of the former, but it also corresponds to the object in the latter. The element is

therefore able to satisfy the selectional requirements of two different predicates at the same time.

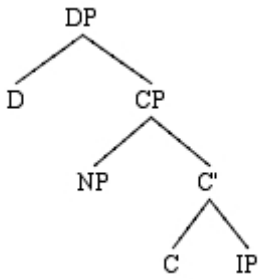
Once recognized this basic system, two fundamental issues arise. The first one concerns the relationship between the modified head in the matrix clause and the modifier, i.e. the relative clause. With respect to this, it is necessary to determine whether the relative clause is an adjunct, a complement or a modifier to the matrix clause. The first analysis defined RCs as adjuncts to the matrix clause, as exemplified below. In this structure, the RC is right adjoined to the NP:

(12)



The alternative account claims that RCs are the complement of the determiner (D°) by which they are selected:

(13)



The second issue concerns the relationship between the relative head and the relativization site and how this operation of abstraction can take place successfully. With respect to this topic, two different accounts have been advanced. On one hand, it is assumed that the relative head is generated as an argument of the matrix clause and bound to the relativization site in the relative clause through a semantic operation. Under this view, we find the proposal known as *matching analysis* (Bianchi, 2002), which implies

that two instantiations of the DP are initially independently generated in the two clauses (the matrix and the relative one). Subsequently, the embedded occurrence of the DP is substituted by a relative pronoun and/or deleted, so that only the DP in the matrix clause (the head) is overtly realized.

On the other hand, after Chomsky demonstrated that dependencies within clauses can be derived through Wh-movement, and that this kind of analysis could be extended to RCs too (Chomsky, 1980), it was assumed that a (silent) Wh-Operator is at work in RCs in order to create a dependency between the gap and the head. The refined proposal predicts that the head noun is generated within the relative clause, extracted from there, and raised to its target position in the matrix clause through Wh-movement. In this perspective, the two positions are bound by being the first and the last instantiations of a chain created through Wh-movement.

The two issues concerning the relationship between the relative head and the RC on the one hand, and the relative head and the gap in the RC on the other hand, influence each other and cannot be completely torn apart. The hypothesis of a RC as adjunct to the DP in the matrix clause requires that the head is generated externally to the RC; in contrast, the complement analysis for the RC clusters well with the idea that the head is generated internally to the RC and raised through Wh-movement.

In their remarkable sketch of the milestones in the history of research on RCs, Alexiadou, Law, Meinunger & Wilder (2000) recall some of the most significant arguments in support of the raising analysis of the relative head from the complement of the DP (namely the relative CP). I will briefly report them in what follows.

First, there exist determiners that require a RC as complement, and do not accept a bare noun in the same place. For instance, that is the case of the German *derjenige* ('the very') in the following example (Alexiadou et al., 2000:8). Thus, the selection of a CP by a D° cannot be pointed out as an *ad hoc* machinery for the derivation of RCs, because it is actually attested in natural languages.

(14) derjenige Mann *(der dort sitz)
'the very man who is sitting there'

Second, the raising analysis allows for the resolution of reconstruction effects. This is visible in the case of pronouns in the relative head that are controlled by the subject of the RC containing the gap (example from Alexiadou et al., 2000:9):

(15) a. the [picture of *himself*$_j$]$_k$ that John$_j$ painted t$_k$
b. the [portrait of him$_j$]$_k$ that John$_j$ thinks that Mary painted t$_k$

As it is evident from the examples above, the (reflexive) pronoun is interpretable and correctly bound only if extracted from the embedded clause, where it can be controlled by the subject of the RC. The alternative analysis with the head generated externally to the RC, is not able to account for such phenomena.

The third argument in support of the raising analysis comes from the so-called 'definiteness effects' on the trace. Within the relative clause, gaps may sometimes correspond to positions in which definite DPs are banned, that's the case, for examples, of presentative structures (16a). However, in case the DP is modified, it acquires a definite reading in the matrix clause, while its interpretation is still indefinite in its original position in the relative clause (namely, the presentative/existential one):

(16) a. There were (*the) men in the garden
b. The men$_j$ that there were t$_j$ in the garden

A similar effect is found also in correspondence to mass nouns, a category that usually refuses strong determiners like, for example, definite articles or universal quantifiers (Alexiadou et al., 2000:11):

(17) Americans exhibit much/some/little/*the/*any/*all courage in such situations.

The interesting observation comes from the insertion of a modifier to the mass noun, in the form of a RC: in that case, judgments about the kind of acceptable determiners are reversed, resulting in strong determiners preferred over the weak ones (again from Alexiadou et al., 2000:11). This is taken as a proof of the fact that the relative head takes its definiteness from the determiner in the matrix clause, but is generated inside the RC.

However, the argument is rather weak and the phenomenon could possibly find an explication also under a matching analysis. The example in (18) indeed shows that heads of relative clauses acquire definiteness, but do not provide irrefutable signs in favor of the raising hypothesis.

(18) American exhibit *much/*some/*little/the/any/all courage *that is required* in such situations

Idioms offer further insights into the raising nature of the relative head. Whenever the nominal part of an idiomatic expression appears in correspondence of the head of the relative clause, the verb that completes the expression must appear in the RC (19a), otherwise the structure results ill-formed (19b). This is a strong proof in favor of the fact that the relative head is first merged in the embedded clause and only subsequently extracted and promoted to the matrix clause (Alexiadou et al., 2000:12).

(19) a. The headway that we made was insufficient
b. *We made the headway that was insufficient

One further argument comes from the observation of languages that differ from English (the language on which much of the research on RC focused) and Italian (the language of interest in the present study): that is the case of languages like Japanese, Quechua and Lakhota, which have head-internal relatives. The (quite self-explanatory) label refers to the fact that the head of the relative appears within the relative clause, rather than external to it (like in Romance and Germanic languages). An example from Lakhota illustrates the structure (Alexiadou et al., 2000:13):

(20) [$_{DP}$[$_{CP}$ Mary [$_{DP}$ owiza wa] kage] ki] he ophewatu
Mary quilt a make the I buy
'I bought the quilt that Mary made'

In light of this phenomenon, any analysis claiming that the head of the relative is generated in the matrix clause is able to account only for some languages, those with

head-external structures. In contrast, the raising analysis offers the possibility for a universal account, able to provide an explanation both for head-external and head-internal relatives, by claiming that the two differentiate only for the extraction part, while they both start the derivation with the same structure, namely with the relative head merged in the argument position of the relative clause.

All this arguments taken together - namely, the existence of determiners that exclusively select CPs, reconstruction effects that allow for the control of reflexive pronouns in the relative head from the subject of the RC, definiteness effects provided by the determiner in the matrix clause to otherwise indefinite DP in the RC, external heads making part of an idiom in the RC and the existence of head-internal RCs – lead researchers in the direction of an account based on the raising, rather than on the matching hypothesis, on the complement account rather than on the adjunct status of RCs.

On top of these arguments, new inputs came from Kayne's (1994) refined proposal in favor of a raising analysis. His contribution represented a decisive turning point in the study of RCs and therefore deserves to be presented extensively.

5.4.1 Kayne´s (1994) raising analysis

Kayne's (1994) analysis of RC builds on his theory of the anti-symmetry of syntax. A fundamental component of the theory is the Linear Correpondence Axiom (LCA), which determines that binary branching is the only way to build acceptable syntactic structures; it follows that right-adjunctions are banned. Because of this *modus operandi*, maximal projections contain no more than three elements: a head, a complement and a specifier. These can combine only according to the order generated by asymmetric c-command, namely the universal specifier-head-complement order.

Building on previous analysis from Schachter (1973) and Vergnaud (1974), Kayne comes to the conclusion that RCs are complements to the determiner in the relative head. In other words, the relative head we find in the matrix clause consists of a D° that does not take an NP as a complement; rather, it takes a CP, the relative clause. The baseline structure of a RC is exemplified in (21):

(21) $[_{DP}$ D° CP]

Determiner and noun in the relative head do not form a phrase *per se*. Rather, the noun is initially generated within the relative clause (22a), merged either in the internal or in the external argument position (depending whether the RC to be derived is an OR or a SR), and it is subsequently extracted and re-merged in SpecCP (22b). The D-N-RC linear order is derived through this mechanism. A covert trace is left in the gap; this is still bound by the moved element.

(22) a. [$_{DP}$ the [$_{CP}$ that [$_{IP}$ the professor helped (the) student]]]
b. [$_{DP}$ the [$_{SpecCP}$ student$_i$] [$_{CP}$ that [$_{IP}$ the professor helped t$_i$]]]]

The analysis above concerns *that*-RCs, namely RCs introduced by the general complementizer *that*, otherwise typical of declarative sentences. However, English counts a second kind of relative clauses, namely those introduced by relative pronouns of the *which*/*who* type. In this alternation, English exemplified the possible strategies found also in other languages. For example, Italian RCs are usually introduced by the general complementizer *che*, while their German counterparts make use of dedicated relative pronouns from the *der*/*die*/*das* paradigm[43].

Turning the attention to RCs introduced by relative pronouns, Kayne (1994) assures that this derivation is compliant with the structure exemplified above. The author claims that relative pronouns are generated within the argument in the relative clause (the internal one in the example in 22), in the position of the determiner, so that they are part of the DP moved to Spec-CP. Finally, the ultimate linear order is achieved through a further sub-extraction of the Noun from the moved DP (23c), while the first part of the derivation takes place according to an identical procedure, independently of the presence of a *that*-complementizer (22) or of a relative pronoun (23):

(23) a. [$_{DP}$ the [$_{CP}$ C° [$_{IP}$ the professor helped [$_{DP}$ who student]]]]
b. [$_{DP}$ the [$_{SpecCP}$ [$_{DP}$who student]$_i$ [$_{CP}$ C° [$_{IP}$ the professor helped t$_i$]]]]
c. [$_{DP}$ the [$_{SpecCP}$ student [$_{DP}$ who t$_k$]$_i$ [$_{CP}$ C° [$_{IP}$ the professor helped t$_i$]]]]

[43] This is just a generalization given that Italian disposes of relative pronouns (*quale*/*cui*) in use in indirect ORs and in non-restrictive RCs, while some varieties of German use the uninflected complementizer *wo* for RCs.

Although Kayne's (1994) proposal can account for a variety of phenomena (among these the derivation of noun-final RC found in languages like Amharic[44]), it is not immune from critiques. Borsley (1997) points out a few issues, which, in his view, are left unexplained by Kayne (1994) and require a certain machinery to be accommodated. In particular, the author focuses on the role of the determiner internal to the relative clause and wonders what its form and role are when it is not realized as a relative pronoun. He also wonders what triggers the movement of the head NP outside of the DP in order to reach the SpecCP position. A third issue arises from the observation of languages endowed with case morphological markers (German, for instance): in these, the relative head is marked for the case assigned in the matrix clause, rather than the one in the RC: a circumstance that speaks in disfavor of the raising analysis (and is actually one of the strong points used by supporters of the matching analysis). Answers to these (and other) questions are provided by Bianchi (1999) and De Vries (2002), who refine their initial hypotheses by building on Kayne (1994) in order to achieve the explanation of a wider range of phenomena. With regards to Borsley`s (1997) first point, Bianchi (1999) assumes the idea that all NPs are actually computed as DPs, in which the determiner can be optionally null (Longobardi, 1994). In other words, bare NPs are not admitted and nouns enter the derivation only in the form of DPs, in which the determiner is eventually present, but silent. That is precisely the case of the derivation exemplified in (22) above. Moreover, it is claimed that the silent D that raises to SpecCP, is further incorporated and licensed by the determiner in the matrix clause that selects the RC. In Bianchi's view, incorporation and deletion of the internal D° are allowed due to the fact that the two Ds, the internal and the external one, share their feature arrays (except for case).

Concerns regarding the actual trigger of the movements required to derive RCs are solved through the assumption of a system of features at work. Bianchi (1999) proposes that the internally-generated DP is endowed with a [+Rel] feature (which stays for "relative clause"), and is therefore attracted to the CP layer of the clause in order to satisfy the Relative Criterion. The final step, namely the sub-extraction of the noun from the raised DP, is also justified by the necessity to satisfy a selectional requirement. The external D° is indeed characterized by an N-selectional feature, which is not satisfied by

[44] Ahmaric has RCs with the following linear order: RC-D-N. According to Kayne (1994), they are derived through movement of the relative clause to SpecDP, above the Determiner that initially selects it.

the merged CP. Thus, it is necessary that the noun raises out of the CP in order to enter the minimal domain of D° and satisfy its selectional requirement for a [+N] feature.

The answer to Borsley's third issue builds on Giusti's (1993) proposal: D° is the element to be marked for case, while N° is not; N° tends to inherit case-marking morphology from the D° it is selected by. Under this assumption, it is therefore easy to explain why the head noun, *per se* not marked for any case, takes the case markers of the external D°, precisely the last determiner it is selected by. Under this view, the idea that head nouns are externally generated based on the observation that they are marked by the case assigned in the matrix clause, does not hold anymore. In contrast, a raising analysis is perfectly suitable.

The opposition between the raising and the matching analyses has been around for a few decades, until a new proposal tried to reconcile the dichotomy between the two positions. In Cinque's (2008, 2014) proposal, the derivation of a relative clause implies the initial merge of both an external and an internal head noun and can imply both a raising or a matching procedure, depending on the language at stake. Details of the proposals are provided in the following section.

5.4.2 Cinque (2008, 2014): head internal/head external relative clauses

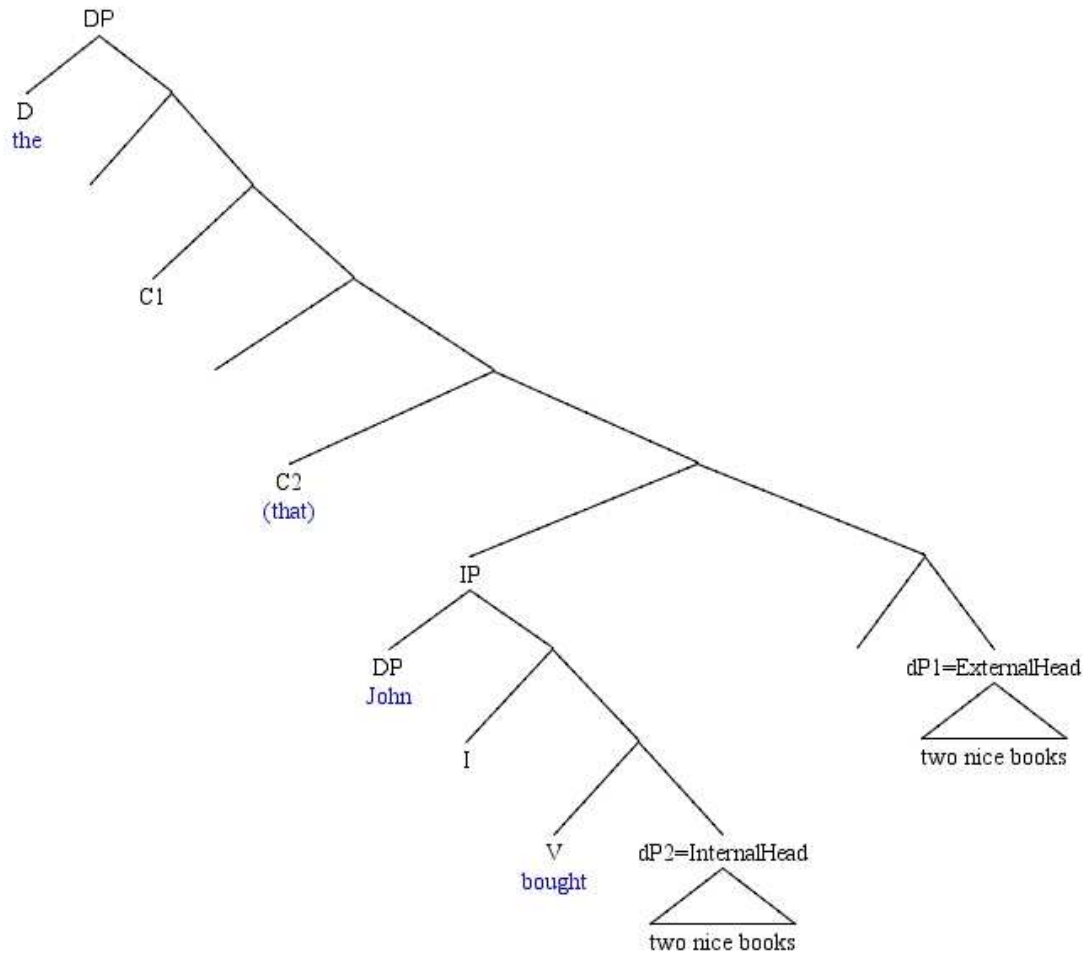
Starting from the observation of languages showing very different (and sometimes contrasting) evidence with respect to the discussed alternative procedures for the derivation of RCs, Cinque (2014) comes to the conclusion that it is not possible to solve the puzzle by eliminating either one of the two alternatives, the matching or the raising analysis. This is due to the fact that the two analyses alternatively account for phenomena that manifest themselves differently in a variety of languages (e.g., reconstruction effects). Therefore, the author shows how the two positions can be reconciled under the assumption that RCs comprehend both an internal and an external head in their structure. Ultimately, his proposal includes a structure that is able to account for all types of RCs: externally headed postnominal, externally headed prenominal, internally headed, headless relatives and correlatives. Crucially, it also succeeds at accounting for cross-linguistic variation, as the structure he proposes is a powerful tool for the derivation of many possible alternative structures. In what follows, I will sketch the basics of Cinque's

proposal. The provided example takes into consideration the case of externally headed post-nominal relatives.

First of all, Cinque (2014) starts by recalling that, because of the universal left-right asymmetry that governs syntactic structures, modifiers that appears to the right of a lexical head (i.e. an adjective with respect to a noun) are not initially merged there. Rather, the final word order is derived by merge of the modifier in a functional projection to the left of the relative head and subsequent raising of the relative head above it. Once applied to the specific case at stake here, it means that the RC is merged in a functional position, inserted above the projections dedicated to Numerals, Adjectives and the NP, but below Universal Quantifiers, demonstratives and definite determiners (Cinque, 2014: 172).

Two fundamental elements characterize Cinque's proposal: first, the presence of an internal and of an external head (as anticipated above); and second, the prenominal position of the embedded clause with respect to the external NP. The basic terms of the proposal are exemplified in the following syntactic structure (adapted from Cinque, 2014:172):

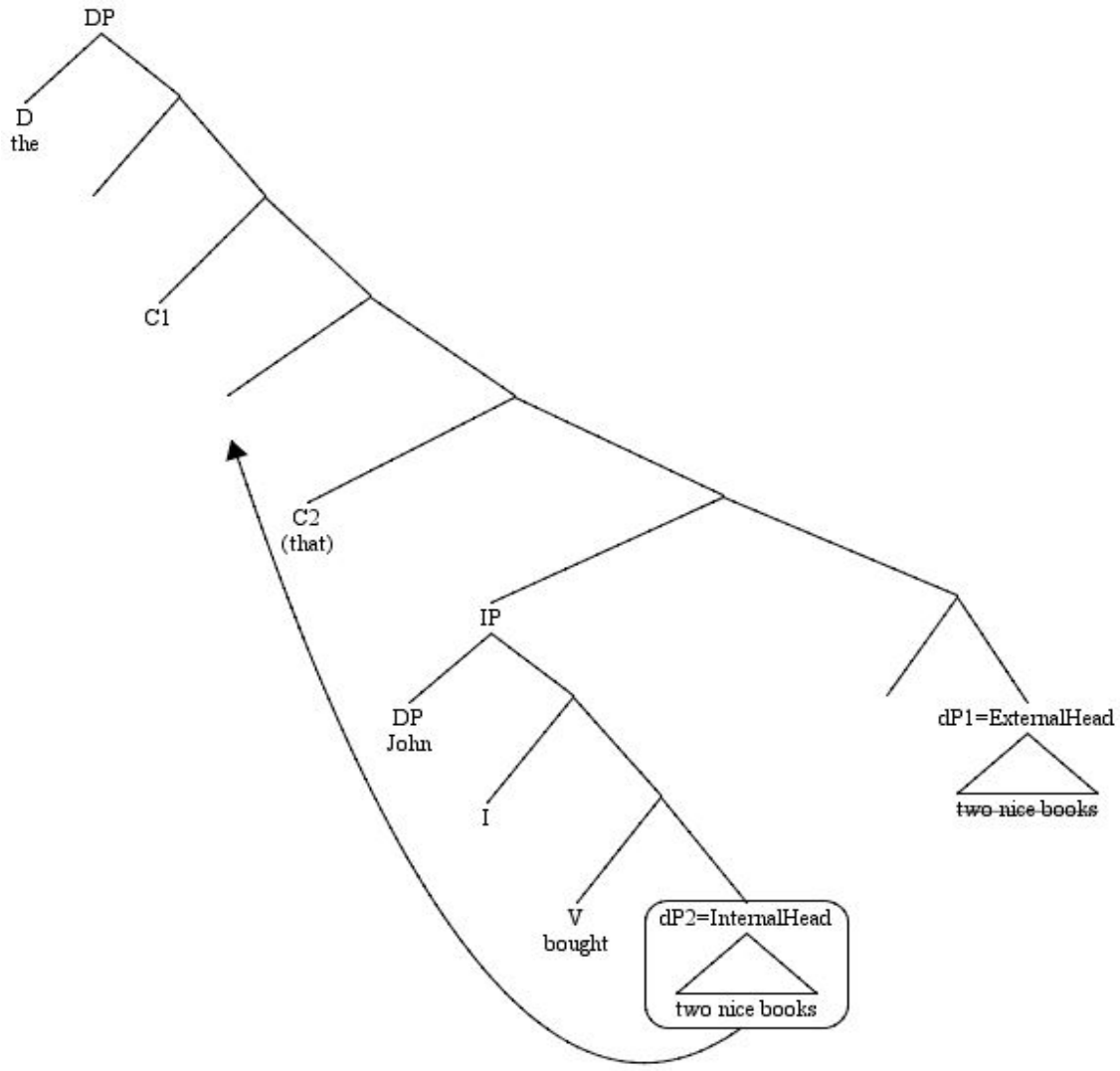
(24) The two nice books John bought



In the structure above, the presence of two functional C projections (C_1 and C_2) is necessary to host the two noun heads, in case they both raise, so that both a raising and a matching derivation are possible. With respect to the presence of two Cs, in Cinque's proposal, the higher node of the relative clause is actually an IP.

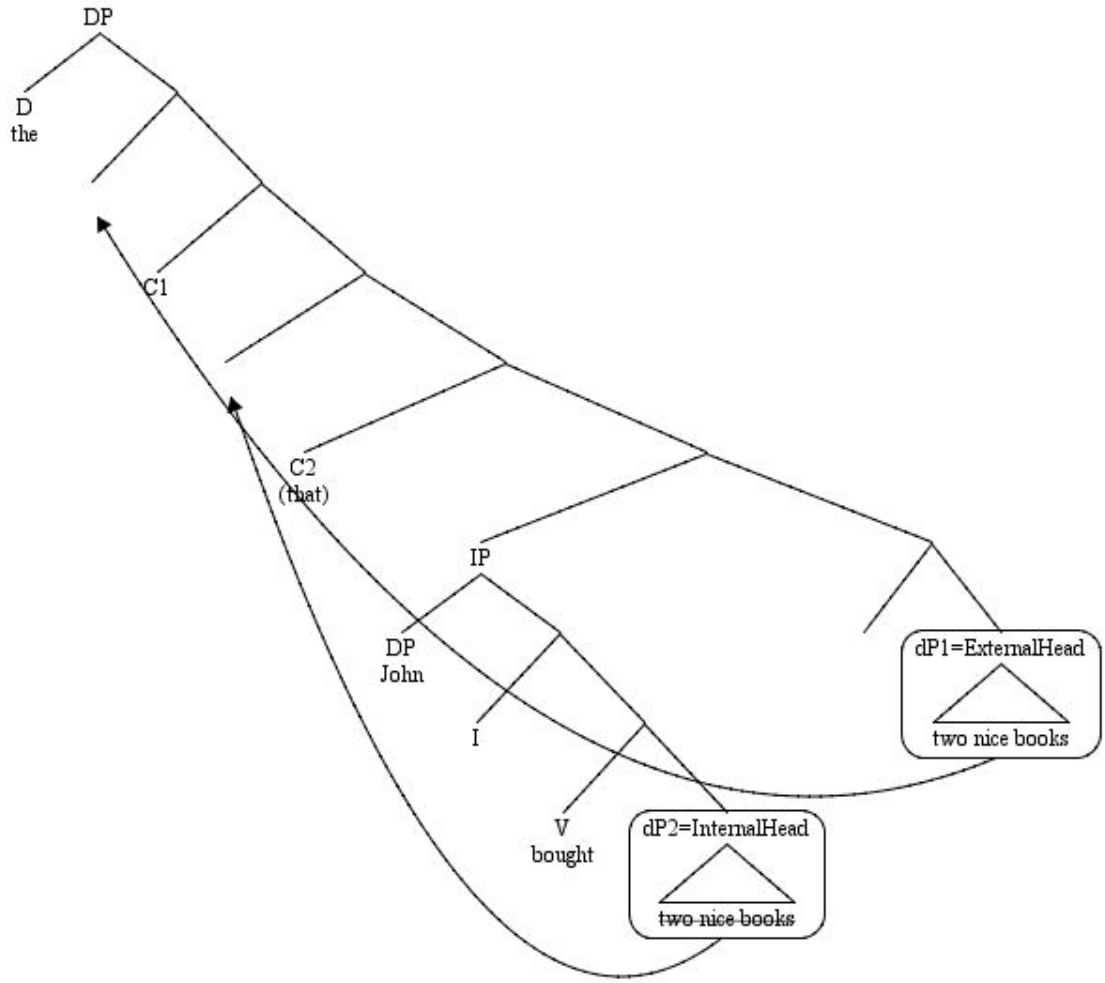
In case of raising, it is the internal head (dP_2) that raises to SpecC_2 and acquires the possibility to c-command the external head from the newly achieved position. It follows that the internal head (dP2) is pronounced, while the external one (dP_1) is deleted at PF. This procedure accounts for all the derivations in which reconstruction and island effects are at hand, precisely for the fact that the pronounced head is the one originally merged in the relative clause and it is therefore able to control for the other elements in its chain.

(25)



The matching analysis differs from the raising analysis for the fact that both heads, the internal (dP_2) and the external one (dP_1) are raised to C functional positions, respectively to SpecCP_2 and SpecCP_1. From this double movement, it derives that the order between the two heads is reversed with respect to the structure corresponding to the raising analysis, namely: the external head in SpecCP_1 precedes and c-commands the internal head in SpecCP_2. It follows that the former is pronounced, while the latter remains covered.

(26)



Moreover, in (26), the fact that the Head to be pronounced is the external one (crucially the one that is not involved in a chain within the RCs) accounts for the cases in which reconstruction and island effects are absent.

The potentials of the universal structure for RCs presented in (24) can be further explored by applying it to different kinds of relatives (prenominal headed relatives, correlatives, etc.), but the operation would exceed the goals of the present section. For the moment, I consider Cinque's (2008, 2014) proposal in its basic lines, and these will be sufficient in order to account for the experimental material in use in the present study.

5.4.3 Complementizers and relative pronouns

As discussed in previous paragraphs, the bulk of the research on the syntactic derivation of RCs has mainly concerned two issues so far: *a)* the relationship between the head and the relative clause; *b)* the relationship between the head and the gap internal to the RC. More recently, a further aspect of the derivation has attracted much attention, namely the nature of the complementizers that introduce RCs.

In the raising analysis proposed by Kayne (1994), the complementizer *that* is generated in C and it takes an overt phonological form in case no relative pronoun enters the derivation. In contrast, whenever a relative pronoun of the *who*/*which* kind is present in the internal DP that raises to SpecCP, *that* is present only in a silent form, i.e. it is covert at PF. More recently, the author has proposed a new analysis for the complementizer *that*, which focuses on the nature of the element. Kayne (2014) claims that *that* derives from the demonstrative pronoun and is equivalent to a relative pronoun. This is not surprising, especially in the light of the fact that other Germanic languages (e.g., German and Dutch) have a similar mechanism, of relative pronouns derived from the demonstrative paradigm. What is particularly interesting is that Kayne (2014) extends its proposal to sentential complements and claims that this are ultimately instances of relative clauses.

With respect to Italian, Kayne observes that the complementizer *che* introducing SR and OR restrictive relative clauses is similar to the English *that* for the lack of overt number (and gender) agreement with the head noun, and for the insensitivity to the animate feature [+/- human][45]. However, contrary to English, he does not trace the origin of *che* back to demonstratives, but to wh-interrogative items. Indeed, *che* appears in questions (27) and exclamations (28) of the following kind (examples from Kayne, 2014):

(27) Che libro/libri hai letto?
'What book/books have you read?'

(28) Che bel libro/bei libri hai letto!
'What beautiful book/beautiful books you read!'

[45] Although Kayne takes this to be not completely true, given that he manages to detect a few contexts in which *that* is sensitive to the [±human] feature.

The fact that the Italian *che* does not agree at least in number is quite surprising because other interrogative items do, that is the case for example of *quale* (also meaning 'which') and *quanto* ('how much'), which also agrees in gender. Therefore, Kayne (2014) claims that *che* is not a modifier of the noun (*libro*), rather, it entails a structure similar to the German expression *Was für ein...?* ('What kind of…?') , very often reduced as in (29):

(29) *Was für Bücher liest du?*
'What for books do you read?'

In (29), Kayne assumes the presence of a silent noun, which stays for 'sort' and actually is the element the Wh-element *was* ('what') agrees with. The analysis is applied also to Italian: *che* accompanies a silent head, represented as SORT. Therefore, Italian relative clauses of the type in (30a) take the structure in (30b), in which *che* precisely modifies the silent element SORT, and not the noun *libri* (Kayne, 2014: 205):

(30) a. *I libri che Gianni ha letto*
'The books that Gianni has read'

b. *I libri*$_i$ [[*che* SORT]t$_i$]$_j$ *Gianni ha letto* t$_j$

In a similar vein, the Italian relativizer *cui*, used in oblique RCs introduced by a preposition (31), is considered to be a form derived from a wh-item modifying the silent head SORT and taking *–ui* as a mark for the oblique case.

(31) *La donna con cui parlo*
'The woman with whom I talk'

More recently, Sanfelici & Poletto (2015) have worked on the third kind of element that can introduce a RC in Italian, namely *il quale*. This differs from the other two complementizers (*che* and *cui*) because it is used both in RCs introduced by a PP and in non-restrictive RCs; moreover, it shows number agreement with the head. A review of

the forms and uses that this complementizer takes in different Italian varieties brought to the conclusion that there must be a grammatical cycle at work with respect to the features that can be specified on the element. The authors assume that relative pronouns are generated within a diachronic relative cycle that concerns Wh-pronouns and demonstratives. The cycle takes the following steps:

(32) 1. wh-pronoun → 2. agreeing complementizer → 3. complementizer → 4. demonstrative → 5. wh-pronoun.

Under this view, it could well be the case that the derivation of the complementizer *that* from the demonstrative *that* (Kayne, 1994) is actually only a narrower portion of a larger cycle at work in languages in a diachronic perspective.

Although the issue has been only more recently addressed in the literature, it offers the possibility to deepen our understanding of Italian RCs, in the hope that this will contribute also to a better understanding of how RC are computed (see Section 5.5).

5.4.4 Relative clauses: a classification

In the preceding sections, while revising the main syntactic accounts to the derivation of RCs, I overlooked an important aspect, namely the fact that a variety of structures goes under the "relative clause" label, although they very often differ from the semantic and syntactic points of view. In other words, different kinds of RCs exist. In what follows, I will limit myself to provide a brief overview of the major kinds of RCs by adopting a descriptive attitude, without entering the syntactic details that account for the manifested differences at the semantic level. For the classification, I follow Grosu & Landman (1998).

The first distinction to be drawn is the one between headed and non-headed relative clauses. The discussion in the previous paragraphs addressed arguments concerning the former kind, in which a noun head appears externally to the RC, i.e. in the matrix clause. Non-headed relative clauses are so called because they do not overtly realize an external lexical head, although expletive functional elements or light nouns are present. Rather, the N endowed with the [+Rel] feature is generated inside the RCs and does not leave it.

Moreover, headed relative clauses can be subdivided in at least three categories: appositive (ARC), restrictive (RRC) and maximalizing (MRC) relative clauses. Restrictive relative clauses define the restriction of the determiner; they contribute at individuating the precise intended referent. Appositive RCs differ from restrictive RCs because they do not contribute to the restriction of the meaning of the noun; i.e., they modify the head noun by providing further information, but do not participate in individuating the exact referent.

Finally, the function of maximalizing RCs is to individuate the unique maximal individual that can satisfy the description introduced by the head noun. For instance, RCs of the degree type belong to this class: they individuate precise amounts rather than general properties. The examples provided below refer to specific amounts of wine (adapted from Alexiadou et al. 2000):

(33) The wine that there was on the table

(34) The three bottles of wine that there were on the table

A baseline classification of RCs is necessary for the sake of clarity and completeness; however, the issue will not be further analyzed in depth here. In what follows, my attention will mainly focus on restrictive relative clauses.

5.5 The processing of relative clauses

Besides their syntactic derivation, relative clauses are very interesting also from a different perspective: the one taking into consideration their computation. In the last decades linguists devoted many efforts to the comprehension of a well-known phenomenon, namely the asymmetry between subject and object relative clauses. By 'asymmetry', it is commonly meant that SRs (35) are processed with less effort in comparison to ORs (36):

(35) SR: The student that helped the professor comes from Italy

(36) OR: The student that the professor helped comes from Italy

Both sentences in (35) and (36) are characterized by the presence in their structure of a gap to be filled with the head of the relative clause. In order to understand the sentences correctly, the parser must identify the gap and link it to the relative head; this means that the head must be interpreted within the relative clause, in the position where it initially merged and where its trace is not spelled-out. The relevant aspects of the structures for (35) and (36) are reproduced respectively in (37) and (38), with traces exemplified within angled brackets.

(37) SR: The student that <the student> helped the professor comes from Italy

(38) OR: The student that the professor helped <the student> comes from Italy

Much debate concerned the role of those traces and their position in determining the asymmetry in processing between SRs and ORs. Before going into an overview of the proposed accounts, I would like to further emphasize the relevance of the phenomenon by recalling the fact that difficulties at OR comprehension and/or production are a well-attested phenomenon. The asymmetry is indeed observed in adults (De Vincenzi 1991a; Gordon, Hendrick & Johnson, 2001; Traxler, Morris & Seely, 2002 among others), as well as in impaired speakers (Caramazza & Zurif, 1976; Friedmann, 2008; Garraffa & Grillo, 2008; Grillo, 2008), and in L1 acquisition (Contemori & Belletti, 2014; Friedmann, Belletti & Rizzi, 2009, Kidd, Brandt, Lieven & Tomasello, 2007, among many others). Only few studies report the mirror pattern, with ORs favoured over SRs; interestingly enough, those studies concern languages characterized by pre-nominal RCs (Hsiao & Gibson, 2003; Lin & Garnsey, 2001; for Chinese) or ergativity (Carreiras et al., 2010; Laka, 2012; for Basque). However, the observation is not robust because Hu presented contrasting data for Mandarin: in her study, both children and adults show a preference for SRs over ORs (Hu, 2014; Hu, Gavarró, Guasti, 2016), in line with the standard asymmetry described above.

As for adults, the canonical asymmetry usually emerges more sharply in on-line studies that imply (multiple) embedded RCs. Indeed, Warren and Gibson (2002) prove how the locus of attachment of the RC, either within the matrix clause or at its right, is

particularly relevant. The authors find a cumulative effect that determines centre-embedded RCs (39) to be more challenging than their right-branching equivalents (40).

(39) The professor that knows the student comes from Italy

(40) The professor knows the student that comes from Italy

The effect is due to the necessity in (39) to keep the first chunk of the matrix clause active in the working memory during the processing of the RC, before meeting its conclusive part. Traxler et al. (2002) sample adult processing of RCs in a study supported by eye-tracking techniques and find out that native English speakers need prolonged fixation times for the comprehension of ORs with respect to SRs; moreover, participants more often tend to regress backwards while reading the former in comparison to the latter. The experimental design also allows for a further interesting observation: stimuli entail centre-embedded RCs, so that the sentence continues after the subordinate clause, with the matrix verb and its object. In this final sentence portion, effects of increased difficulties for ORs manifest in the form of a spill-over effect, which means that prolonged fixation times (to be interpreted as prolonged processing time) also characterize the final portion of the matrix clause in the case of ORs. Crucially, SRs do not produce this effect. Gordon et al. (2001) reach similar results with a self-paced reading task in which they detect longer reading times for ORs, especially in the case these include two full DPs, both designating animate referents[46]. A second interesting result of their study concerns accuracy in sentence comprehension: participants are more correct in answering questions concerning complex sentences including SRs, than complex sentences including ORs.

As for Italian, the processing of A' dependencies by adult native speakers is directly addressed by Marica De Vincenzi (1991a, 1991b) in her work on syntactic parsing strategies. The author finds a sharp asymmetry between sentences entailing subject extraction and the ones entailing object extraction, with the former preferred over the latter. However, given that her experimental material mainly covers different types of

[46] I will examine in more details the role of animacy and lexical restriction in the processing of RCs in Section 5.6.3.3.

Wh-questions, the results of her work will be more extensively illustrated in Paragraphs 5.6.

The bulk of the research on the subject-object asymmetry deals with L1 acquisition and aims at unveiling how children proceed towards the mastery of RCs. In these studies, adults are often enrolled in the control group and provide data that are taken as baseline for the evaluation of child performance. In many cases, task designs and stimuli are tailor-made in order to meet children´s abilities and therefore their level of complexity is kept as low as needed in order to allow young children to provide with meaningful feedback. In those contexts, adult participants usually perform at ceiling in all conditions, failing to provide us with further evidence for the Subject-Object asymmetry. Only few studies on child acquisition also provide interesting results concerning adult proficiency. For instance, Contemori and Belletti (2014) find out that Italian adult speakers tend to avoid the production of ORs by adopting a passive voice.

In what follows, I will illustrate the accounts that have been proposed so far in order to illustrate the asymmetry between subject and object RCs. In this, a particular attention will be dedicated to the proposal advanced by Friedmann, Belletti & Rizzi (2009; and much subsequent work), according to whom, the computation of RCs observes locality constraints. As mentioned, the bulk research on the topic has been carried out with children; the review will therefore present a number of studies covering L1 acquisition, along with studies on adult proficiency and language impairment. Indeed, due to their complexity and their status in the competence of both adult and young speakers, previous studies already used RCs to sample the syntactic and processing abilities of speakers affected by impaired linguistic abilities (Caramazza and Zurif, 1976; Friedmann, 2008; Garraffa & Grillo, 2008; Grodzinsky, 1989). The review of previous studies and accounts plays the role of setting some critical issues that are of fundamental importance for the design of the task in use with PADs, and the interpretation of the results presented in Chapter 7.

5.5.1 Accounts to the subject-object asymmetry in relative clauses

Many different accounts have been proposed in order to explain the observed asymmetry between syntactic derivations entailing argument extraction either from the subject or from the object position. Different classifications of the provided accounts are possible,

depending on the parameters and factors that are taken as guidelines. I classify each study according to the linguistic or the cognitive aspect each account mostly focuses on. I will distinguish between memory-, processing- and syntactic-based accounts: the first group of accounts focuses on the role of working memory in sentence computation, the second one on processing strategies, and the third one on syntactic structures. Most ideas and hypotheses advanced in the past decades broke boundaries of this schematic classification and considered the interplay of different factors as the cause of enhanced difficulties with ORs. I decided to classify the different accounts according to the component that seems to characterize each at most.

Memory-based accounts (Ford, 1983; Frazier & Fodor, 1978) focus their attention on the limitation that working memory sets to our capacity to compute syntactic structures characterized by increasing levels of complexity, i.e. recursive embedding. According to Frazier & Fodor (1978), the parser proceeds by analysing a limited amount of words at the time, cutting the word string into sub-units, in order to try to assign them a structure. In case one (or more elements) in the string at stake cannot receive a proper role in the structure, storing loads highly increase, thus reducing memory resources for further analysis, at least until a proper structure is reconstructed. Along similar lines, Ford (1983) claims that (embedded) ORs bring along a heavier burden than SRs do, because they require the parser to keep in mind the unassigned relative head for a longer time span, before assigning it to the proper gap. When a gap is met, the parser must search backwards for a proper filler. Gaps in the object position automatically require longer inspections in the previously processed sentence portion, because the filler is set further away. Gibson (1998) shifts the attention to integration costs and claims that parsing demands are determined by the number of elements that intervene between the filler and the gap. Building on Bever's idea (1974) according to which some syntactic items (e.g., pronouns) are easier to process than others (lexical referring expression), Gibson claims that lexical referents are particularly demanding because they require higher levels of lexical, semantic and discourse activation for their integration in the structure. Gordon, Hendrick and Johnson (2001) further refine the proposal by hypothesising the existence of a similarity-based interference. In their study, authors find reduced parsing difficulties in ORs when the subject and the object belong to different word classes, for instance a pronoun and a noun. In their view, working memory is particularly challenged by

sentences in which the involved DPs are highly similar, because of the difficulty at keeping both in mind and discriminating between two (almost identical) elements, while assigning them different syntactic functions and discourse roles. Accordingly, clearly different elements, e.g. a noun and a pronoun, facilitate processing. This contrasts the idea that lexical categories are associated to specific levels of complexity *per se*, as proposed by Gibson (1998) for pronouns. In line with Gordon et al. (2001), I assume that the processing demand of an element is not to be determined *a priori*; rather, the context in which the element appears is crucial to determine its processing demand.

As already mentioned, a second group of accounts ascribes the observed asymmetry between subject and object extraction to processing strategies. McWhinney & Pléh (1988; but see also McWhinney, 1987) claim that the parser builds expectations while processing a sentence; in particular it tends to form and maintain perspectives in which elements maintain fixed functions. If the parser meets a DP and analyses it as the subject of the matrix clause, it expects that the item maintains the same function when modified by a RC. If this is not the case and the element is assigned a different syntactic function, the parser is forced to change perspective, resulting in enhanced processing loads. Perspective maintaining is economical and allows for smoother sentence processing. Certainly, the focus of processing-based accounts is the claim that processing takes place always in an economic way, in which predetermined procedures or strategies play a facilitating role. That is precisely the core idea of the *Active Filler Strategy Hypothesis* (Clifton & Frazier, 1989; Frazier & Clifton, 1989): the parser tends to assign a filler the first possible gap in order to build the simplest possible structure. De Vincenzi (1991b, 1996) also specifies that the parser avoids postulating nodes that are not necessary, but, at the same time, does not delay the insertion of required chain members. In other words, the parser is engaged in building just the proper amount of structure needed to represent the linguistic input. This strategy works particularly fine for SRs because the filler (the relative head) is assigned the first possible gap, namely the one corresponding to the subject position in the relative clause, and processing completes successfully. In contrast, the strategy is less proficient when it comes to ORs: in this case, after assuming a gap in the subject position, the analysis must be revised and corrected, in order to allow the assignment of the head to the correct object position. Under this view, prolonged reading times for ORs are the manifestation of the ongoing correction of the initial structure.

However, the described principle does not predict the possibility to reduce the asymmetry between subject and object relatives through a manipulation of the arguments involved in the derivation. In this sense, it cannot account for some of the experimental results I will review below.

Processing-based accounts also include a proposal that dedicates a particular attention to discourse constraints. The claim is formalized in Kidd, Brandt, Lieven & Tomasello (2007) and it states that speakers produce ORs only under specific circumstances. The first condition authors set to the computation of ideally-formed ORs is the presence of inanimate referents in correspondence of the relative head (see also Fox & Thompson, 1990). As animate referents strongly trigger SR representations, in their view inanimate nouns posit looser constraints for the sentence interpretation and therefore favour an OR reading. The second condition for ideally-formed ORs concerns the subject. According to the author, subjects of ORs most often refer back to referents already presented in the discourse and therefore are expressed in the form of pronouns. According to Kidd et al. (2007), these two constraints taken together, namely inanimacy of the head and pronominal subjects, represent basic conditions for the spontaneous production of ORs and indeed characterize the majority of uttered ORs. The authors take the issue to have a certain relevance at the processing level because computation takes place based on statistical data. In other words, frequency of use determines speed and accuracy in processing, assumed that very frequent structures are faster and better processed. In the authors' view, the same principle holds for L1 acquisition too, with children producing earlier and better the structures that they more often find in the linguistic input. Authors also review previous studies from this perspective and claim that the subject/object asymmetry widely found in a variety of tasks is ultimately due to ill-formed ORs in the experimental input. That equals to say that the use of very rare and not-common-at-all ORs is the primary cause of poor experimental performance.

I now go back to the issue of the distance between filler and gap because this is relevant in the first of the syntactic-based accounts I would like to mention here, namely the *Structural Distance Hypothesis* formulated by O'Grady (1997; O'Grady, Lee & Choo, 2003). O'Grady thinks of distance neither in terms of number of referents entering the discourse between the filler and the gap (as in Gibson, 1998; Gordon et al., 2001) nor in linear terms (De Vincenzi, 1991b, 1996), rather in structural terms. The author claims that

complexity is proportional to the degree of embedding and therefore to the gap depth. In his view, what is crucial is the number of syntactic nodes that are necessary to create SRs and ORs. Given that the latter entails more nodes in the structure assumed by O'Grady, effects of increased processing difficulties are automatically explained.

Authors who build on the principle of Relativized Minimality (Rizzi 1990, 2004) in order to explain the asymmetry, take quite a different syntactic approach. Grillo (2009) and Friedmann, Belletti & Rizzi (2009) first observe that locality constraints are at play in ORs. In ORs the internal subject intervenes between the extracted object DP and its target position in the CP layer of the RCs. The syntactic principle also accounts for a variety of phenomena observed in previous studies (as mentioned above). For example, I refer to the issue concerning the asymmetry (brought up by Gordon et al., 2001) between lexical referents and pronouns and the different ways in which they can alternatively improve or not OR comprehension, or to the role played by the number of referents intervening between the filler and the gap (Gibson, 1998). All these issues find an explanation under the syntactic formalization offered by RM, as I will show in the following and in the subsequent sections.

5.5.2 A Relativized Minimality account of relative clauses

Relativized Minimality (Rizzi, 1990) is certainly not new to linguistic research, but its theoretical potential is still under investigation. Only more recently it has been proposed that RM can offer a precise explanation to a variety of phenomena observed in RC computation. The first observation comes from Grillo (2009) in his PhD thesis on the computation of A'-dependencies by agrammatic patients; the hypothesis is further explored by Friedmann, Belletti & Rizzi (2009) in their seminal work on L1 acquisition. I will now start by briefly illustrating the basics of RM in order to show how the principle applies to RCs.

Relativized Minimality is grounded on the fact that syntactic relations must take place locally, where by 'locally' it is meant the smallest structural domain where a relation can be satisfied (Rizzi, 1990). The minimal configuration to take into consideration in order to understand how RM works consists of the three elements in (41): the original position of the extracted element (Z), its target position (X) and an intermediate position (Y).

(41) …X...Y…Z…

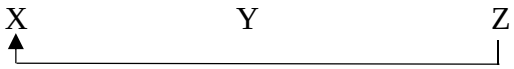
Given the configuration in (41), Z can enter in a minimal configuration with X, namely it can satisfy its syntactic relation in X, only if the two following conditions are met:

a) Y is not of the same structural type as X;
b) Y does not intervene between X and Z.

In the conditions above, being of the 'same type' is equivalent to sharing relevant features and 'intervening' means blocking the movement of Z towards its target position in X. Once applied to relative clauses, the configuration in (41) looks like (42) for SRs and (43) for ORs:

(42) SR: The student that <the student> helped the professor

(43) OR: The student that the professor helped <the student>



In (42) there is no element intervening between the relativization site Y and the target position X, and the two positions can enter into a syntactic relation without further complications. In (43) the configuration involves three positions: the target position X attracts the object from its embedded position Z; in between, there is the subject Y. If the subject were exactly of the same kind of the target position, ORs would result in being illicit constructions; but this is not the case. Rather, the fact that ORs are grammatical but difficult to process, signals that the subject (Y), despite being inserted between the two relevant positions, does not reach the threshold level of activation in order to completely block the derivation, i.e. the movement of the object towards its final position. The intervention of the subject makes the computation more difficult. As stated in condition (1) above, the features shared (or not shared) by the involved elements in the derivation determine whether intervention will take place and block the movement or not. According

to Friedmann, Belletti & Rizzi (2009), the fundamental features to be considered for the derivation in (43) are [+Rel] and [+N]. The first one belongs to the class of discourse-related feature and is represented both in the relative head (the probe) and in the relativization site (the goal) and it is actually the feature that triggers the formation of a RC. The second feature, [+N], stands for lexical restriction and is represented in the three involved positions:

(44) OR: The student that the professor helped <the student>

[+NP, +Rel]	[+NP]	[+NP, +Rel]
X	Y	Z

The subject ('the professor') shares some relevant features with the relative head, at least the fact of being characterized by a lexical restriction; on the other hand, [+Rel] characterizes only the relative head and the relativization site. Thanks to this feature, which makes positions X and Y different enough, intervention is avoided.

After the first proposal formulated in Garraffa & Grillo (2008; Grillo 2009) and in Friedmann, Belletti & Rizzi (2009), much subsequent work focused on the analysis of the possible further implementation of RM into the analysis of RCs. The observation that features like animacy of the relative head and grammatical class of the embedded subject have an influence on speakers' linguistic performance, raised questions concerning whether it is precisely feature manipulation on the relevant items that is able to trigger visible effects on the processing of the corresponding structures.

The Principle of Relativized Minimality is intrinsically syntactic in nature and it determines how processing takes place according to syntactic constraints. Under this view, the dichotomy between processing- and syntactic-based accounts to the subject/object asymmetry must be reconciled.

The next section is dedicated to a review of studies which have contributed to pointing out which features (and which sorts of manipulations) are responsible for reducing the processing load imposed by ORs. Although most recent studies on the issue were conducted on children engaged in L1 acquisition, their results are nonetheless of a crucial importance for the design and the interpretation of the study I will present in Chapter 7.

5.5.3 Factors reducing the complexity of object relatives

While reviewing studies concerning different accounts to the subject/object asymmetry, I came across works that claim that differences between the two involved DPs can improve the processing of ORs (Gibson, 1998; Gordon et al., 2001; Kidd et al., 2007). Different factors (e.g., animacy and word class) are called into cause, but all cited authors fail at individuating the precise role of those asymmetries and the mechanism allowing for better processing in case of mismatch in characteristics between the two arguments involved in the derivation (the internal and the external one). Crucially, Friedmann, Belletti & Rizzi (2009) succeed in pointing out the common principle that regulates how mismatch can improve OR processing and they also formalize it into syntactic terms.

In the preceding section I reported that ORs are made difficult by the fact that the target and the potential intervener in (44) are very similar, where by 'similar', it is meant that they share some features, for sure at least [+NP], which stays for the lexical restriction. Still, the two positions differ thanks to the [+Rel] feature that characterizes the target but not the intervener. The configuration of the relevant positions for the computation of ORs is represented in (44), in which the feature arrays of the target and the intervener are in an inclusion relation, in the sense that the features characterizing the intervener are a sub-set of the features entailed in the target. Grillo (2009) and Friedmann et al. (2009) claim that it is this specific feature array of inclusion that determines the difficulties in parsing ORs. Indeed, children´s very low performance on this kind of structure allows the authors to claim that younger speakers apply a stricter version of RM, by blocking derivations in which the feature arrays of relevant positions are in a configuration of inclusion. Otherwise the inclusion configuration hinders the processing of ORs in adult speakers, but does not block it completely. Belletti et al. (2012) confirm the observation for adults, but refine the analysis for language acquisition and claim that the child grammar allows only feature configurations of non-inclusion, namely disjunction and intersection. Authors summarize their proposal as follows (adapted from Belletti et al., 2012: 1063):

(45)

	Target	Intervener	Trace	Adults	Children
Identity	+A	+A	<+A>	*	*
Inclusion	+A, +B	+A	<+A,+B>	ok	*
Intersection	+A, +B	+B, +C	<+A, +B>	ok	ok
Disjunction	+A, +B	+C, +D	<+A, +B>	ok	ok

With respect to the hypothesis summarized above, Belletti et al. (2012) claim that only features responsible for movement attraction are involved also in the computation of the feature arrays that determine the well-formedness of a configuration.

In the last few years, much work has focused on the search for the features that are relevant for the computation. Authors focused on lexical restriction, number, animacy, gender, case, passive voice, and resumption. I am going to review relevant studies on the role of those features in what follows.

In addition, I would like to remark that I am going to review both comprehension and production studies in the following sections. I will focus on the features taken into consideration by the different studies and disregard asymmetries that might arise between production and comprehension in language acquisition and in adults.

5.5.3.1 Syntactic category and lexical restriction

After claiming that RM regulates RC processing, Friedmann, Belletti & Rizzi (2009) support their hypothesis by showing that feature mismatch improves OR comprehension in Hebrew-speaking children aged 3;7-5;00. For their purpose, authors employ different syntactic categories for the two arguments involved in the derivation. In particular, data from several sentence-to-picture matching tasks report improved comprehension on conditions entailing at least an argument not realized as a noun. In free ORs and in ORs with an impersonal *pro* in the subject position, the absence of a lexical restriction in either one of the two arguments sharply allows for better comprehension, with respect to the baseline condition in which both arguments are endowed with lexical restrictions. Therefore, a mismatch in lexical restriction can be listed among the features that are relevant for the syntactic derivation.

The account put forth by Friedmann et al. (2009) also provides a good explanation for results gathered under very different approaches and accounts. I refer to the results collected by Kidd et al. (2007) with German and English speaking children and by Arnon (2010) on Hebrew. Both studies report that children start producing ORs with subject pronouns, being this the condition favoured by discourse-related constraints and therefore resulting in the most frequent configuration in the input. If we look at the results from the RM perspective now, it is evident how the processing is favoured by a feature mismatch between the internal and the external argument involved.

The reviewed study shed also new light onto the results presented by Gordon, Hendrick & Johnson (2001) and by Warren & Gibson (2002) with adult native speakers of English. In the former study, authors find different patterns of performance in accordance with the nature of the arguments involved: subject pronouns and proper names lead to faster reading paces, as compared to the condition involving two full DPs. Equivalent results are obtained by Warren & Gibson (2002) in a task in which adult native speakers of English are requested to evaluate the level of difficulty of different sentences. In doing so, participants show that they find ORs with two DPs more difficult than sentences in which the embedded subject of the ORs is represented by a proper noun or by a subject pronoun, especially in the case of 1st and 2nd person pronouns. In their questionnaire, third person pronouns also bring along evaluations corresponding to facilitated processing, although to a lower extent. Under the account proposed by Friedmann, Belletti & Rizzi (2009), the two experiments on groups of English native speakers reinforce the idea that sensitivity to grammatical category is visible in adults' performance too and is not a mere characteristic of L1 acquisition.

An open issue with respect to the role of pronouns in OR computation is the relevance that different pronouns can assume; the research question concerns in particular third person pronouns. Haendler, Kliegel & Adani (2015) find that facilitating effects drops when a third person pronoun is elicited, despite excellent results with first and second person pronouns. The results clearly resemble the one reported by Warren & Gibson (2002) for adult English native speakers. Haedler et al. (2015) explain the effect as due to the accessibility of the referents in the discourse. First and second person pronouns unambiguously signal referents directly involved in the discourse, while 3rd person pronouns entail a harder task, i.e., the hearer must search the previous discourse

portion and the broader context in order to individuate the possible referents corresponding to the pronoun, thus overloading and worsening the processing. With respect to this issue, it would be very interesting to verify the effects that the introduction of proper names could generate in the processing of RCs, but unfortunately the study did not include conditions with the mentioned characteristic.

Besides the issue concerning third person pronouns, data overall consistently show that a mismatch in lexical restriction improves RC processing.

5.5.3.2 Number

Adani et al. (2010) report that a mismatch between singular and plural arguments can improve the comprehension of ORs with respect to ORs in which both the subject and the object DPs are characterized by the same value for the number feature, either singular or plural. The comprehension study consists in a sentence-to-picture matching task, in which 50 Italian speaking children (between 5 and 9 year old) are enrolled and sampled on the comprehension of centre-embedded RCs. Results from the study clearly speak in favour of a facilitating effect emerging when the Object-DP of ORs moves across the subject-DP and is characterized by a different number feature. Adani (2012) employs the same design also with English-speaking children in the same age range as the Italian ones (except for the absence of children at age 5) and finds equivalent results, with the mismatching condition allowing for better accuracy than the matching condition.

Results are confirmed also by an on-line task performed by English-speaking children (Contemori & Marinis, 2014). The self-paced reading task in use does not reveal any relevant difference among conditions with respect to how they are read; the asymmetry arises only on the subsequent comprehension task.

Guasti, Stavrakaki & Arosio (2008) report results from a task that manipulates the position of the subject (either pre-verbal or post-verbal) as a relevant factor and uses number mismatch as a disambiguating strategy for ORs. Authors report that the effect of number mismatch is amplified in co-occurrence with pre-verbal subjects. Namely, center-embedded ORs are better processed when they are disambiguated both by number mismatch and by the pre-verbal position of the subject, with respect to the condition in which only the former factor is at play and the subject follows the verb. The claim is based on data provided both by Italian and by Greek speaking pre-school children (age

range for Italian: 4;5-5;9; for Greek: 4;5-5;6). Therefore, pre-verbal subjects play the role of providing the hearer with a clear cue of the fact that the filler cannot be assigned to the subject position as there is no gap to be filled in that precise position. This interplay between number mismatch and subject position emerges in a production task reported in Belletti & Contemori (2010) too. The two authors try to elicit ORs with matching and mismatching DPs (with respect to number features) and find that children are not automatically facilitated by this strategy in their production. Children very often fall back into ambiguous RCs by producing ORs with post-verbal subjects and matching number features. The authors therefore interpret the results as biased by a side effect, namely agreement attraction from the relative head. In ORs disambiguated via number mismatch, attraction effects emerge: the verb changes its agreement pattern by taking the features of the object in the sentence. Interestingly enough, pre-verbal subjects contain the phenomenon and hinder agreement changes; in other words, when the subject is produced pre-verbally, it agrees with the verb in a consistent way; when the subject is produced post-verbally, agreement attraction from the head of the OR is at play.

5.5.3.3 Animacy

Studies carried-on outside the RM account have independently proved that animacy plays a relevant role in the processing of ORs. In particular, Kidd et al. (2007) claim that an inanimate head is the prerequisite that allows[47] for spontaneous production of ORs. This assumption is based on discourse-related constraints and is taken to characterize adult spontaneous production. ORs with inanimate heads are therefore the most frequently-produced kind of ORs, a factor able to influence child acquisition. From the authors' point of view, children start their production by the sentence structures that they most often find in the input, namely ORs with inanimate heads.

The assumption is confirmed by on-line studies with adult English speakers (Traxler et al., 2002; Gennari & McDonald, 2009) and with adult Dutch speakers (Mak, Vonk & Schriefers, 2002). In Traxler et al. (2002) English native speakers undergo an eye-tracking task sampling the reading of embedded SRs and ORs. In ORs they perform faster on conditions entailing an inanimate head and an animate subject. The same holds

[47] The second prerequisite Kidd et al. (2007) point out is the presence in the relative subject of a pronoun rather than of a noun.

true for the word-by-word self-paced reading task completed by Dutch native speakers (Mak et al., 2002), who also show an effect of animacy.

The assumption that a mismatch in animacy plays a role in the processing of ORs is in line with the principle of RM (Rizzi, 1990), as illustrated in the previous section (Friedmann, Belletti & Rizzi, 2009; Grillo, 2009). However, studies conducted under this approach (i.e., RM) failed at proving any facilitating effect triggered by animacy. For what concerns children, the hypothesis is unsuccessfully tested both with European Portuguese (Costa, Lobo & Silva, 2011) and with German (Adani, 2012). For instance, only Adani (2012) finds an effect of animacy for the comprehension of ORs, but this is weak and limited to 4-year old children; the later effect disappears in older subjects (5-year old) and in adults.

As for adults, Belletti & Chesi (2011) directly address the issue with Italian native speakers in an elicited production task. Again, no effect of animacy emerges, in the sense that Italian adults reacts to the stimuli with a consistent strategy, independently of whether the head of the OR is animate or inanimate.

In conclusion, in the light of contradictory results, the issue of the role of animacy is still open.

5.5.3.4 Gender

Adani et al. (2010) first address the role of gender in enhancing the asymmetry in features between the two arguments involved in the derivation of ORs. Their study enrols Italian-speaking children from age five to nine and implies a number of stimuli in which different genders characterize the characters depicted in the sentence-to-picture matching task. The notion of gender in use in their work corresponds to grammatical gender; indeed, the characters are all taken from the animal realm. Italian nouns referring to animals are usually characterized by a fixed grammatical gender, which does not necessarily corresponds to the natural gender of the animal. For instance, *coniglio* ('rabbit') is masculine, while *capra* ('goat') is feminine[48]. According to the results, Italian children are not sensitive to the gender feature and their difficulties with OR persist unaltered, also in

[48] Some animals, usually the ones humans are traditionally more familiar with, have both masculine and feminine form, depending on the actual natural gender of the referent: for example *gatto* ('cat.M') and *gatta* ('cat.F'). See Chapter 3 for a more detailed review on the topic.

the case of gender mismatch between the relative head and the subject of the relative clause.

Belletti, Friedmann, Brunato & Rizzi (2012) confirm the results with a different group of Italian-speaking children, while the same does not hold for Hebrew speaking children. Indeed, the latter group shows sensitivity to gender features. A mismatch in gender between the involved arguments allows for enhanced comprehension of Hebrew ORs. The asymmetry between the two groups of children (Italian versus Hebrew speaking subjects) is not surprising if we consider the different status of gender in the two languages. In Hebrew verbs have a rich agreement morphology, which includes a morpheme expressing gender agreement with the subject; while this is not the case in Italian. The authors therefore deduce that gender mismatch does not have an effect *per se*, rather it manifests only in those languages in which gender plays a relevant morphosyntactic role and contributes at triggering the movement of the extracted phrase towards its target position. Based on these results, the authors also get the chance to refine the RM account they propose and they come to the conclusion that only features contributing in the process of triggering movement towards the target position are relevant in the computation of intervention. It derives that gender is relevant in Hebrew but not in Italian.

Moreover, the presence of gender mismatch creates a configuration of feature arrays that cannot be classified under inclusion (the potential intervener corresponds to a subset of the features characterizing the target), rather, it is a case of intersection, in which the potential intervener and the target position share only some of their features and distinguish themselves for others.

5.5.3.5 Case

Despite its relevance in syntactic computation, case has received a limited amount of attention within the study of RCs. Guasti, Stavrakaki & Arosio (2008) showed that in Greek case helps disambiguating the reading between SRs and ORs and ameliorates the comprehension of the latter. In their study, authors also compare the effect of case to the effect of disambiguation via number agreement[49] (combined with pre-verbal and post-verbal subjects), and find that the former (case disambiguation) allows for more effective

[49] Greek allows for case neutralization in nouns marked with neuter gender.

comprehension, especially when compared to sentences entailing number mismatch and post-verbal subjects[50].

5.5.3.6 Passive voice

Studies based on the elicitation of RCs offered the opportunity to unveil more details on how children and adults react when a ORs is prompted. In particular, speakers tend to put in practice a variety of strategies that allow for OR avoidance.

In what follows, I focus on a strategy that has emerged in a significant number of studies, namely the use of object relatives with a passive voice (POR henceforth, following Belletti, 2014; and Belletti & Rizzi, 2013) in order to convey the exact meaning of the attended ORs. This kind of structure (46) consistently emerges in the production of adult Italian speakers.

(46) *Lo studente che è aiutato dal professore*
The student that is helped by_the professor
'The student that is helped by the professor'

The phenomenon is observed in Contemori & Belletti (2014) through an adaptation of a preference task first proposed by Novogrodsky & Friedmann (2006). In this, Italian adults perform very consistently, namely all participants produce PORs to a variable extent that altogether covers the 88% of the total amount of ORs produced in the study.

In the same study, authors also find out that children progressively tend to emulate adults' performance by gradually increasing the number of PORs produced in place of ORs. Children younger than 4 do not produce any POR; while the first occurrences appears at age 4 (although they are limited only to two children). From that age on, the number of children producing PORs visibly increases, until reaching the total of 16 out of the 20 young participants within the group of 8 year-old children. On the same task, adults equally produce copular or reduced passive, while children start out by producing causative passive structures, and turn to copular passives only later (34.6% at age 8).

A subsequent study (Belletti & Chesi, 2011) tries to investigate whether the presence of two animate DPs in the stimuli prompting ORs can be the source of difficulty

and the reason for adopting PORs. Therefore, the same preference task is reproduced with stimuli designed in order to elicit ORs with an inanimate head; but the manipulated factor does not trigger any effects and PORs consistently remain the preferred strategy for adult Italian speakers. On these premises, it is legitimate to deduce that PORs are a simplification strategy that speakers adopt when requested to produce a ORs.

In order to verify the actual level of complexity of PORs with respect to ORs, Contemori & Belletti (2014) decide to test the comprehension of the former with respect to the latter. A sentence-to-picture matching task is therefore administered to children from age 6;5 to 8;11. Beside copular passive, stimuli also include reduced passive (47) and causative passive forms (48).

(47) *Lo studente aiutato dal professore*
The student helpbed by_the professor

(48) *Lo studente che si fa aiutare dal professore*
The student that cl make help by_the professor

As opposed to ORs (for which accuracy is steadily attested around 64% in the present task), Contemori & Belletti (2014) find that children comprehend the three different kinds of PORs with levels of accuracy that reach obove 75% already at the age of 6, and even reach 92% at the age of 8.

Studies on other languages make the results even more robust, as the strategy also emerges in European Portuguese (Costa, Lobo, Silva, 2011), in German (Adani, Sehm & Zukowski, 2013; Sanfelici, Trabandt, Schultz & Thiel, 2014) and in Danish (Jensen de López, Sundahl Olsen & Chondrogianni, 2014). In particular, Costa et al. (2011) find PORs in adult production in the forms both of copular and reduced passives. For German (Adani et al., 2013), PORs are attested both in children and in adult controls, although with different percentages. German children (aged 5 to 9) produce PORs only up to the 15% of the produced sentences, quite far from the 40% attested in adults. However, percentages are higher in the study run by Sanfelici et al. (2014), who find that adult German speakers produce as many PORs as their Italian peers (85.3%). Moreover, their children consistently produce PORs from age 5 on (32.6%). Finally, Jensen de López and

colleagues (2014) show that Danish speaking children have a very similar behaviour to the ones attested in Italian, in Portuguese and in German children, with PORs emerging as a frequent alternative to ORs.

In the light of the reported data, the question concerning the reasons why PORs represent a valid alternative to ORs automatically arises. Belletti (2014; but see also Belletti & Rizzi, 2013) offers a valid explication, which builds on the *smuggling* analysis of passive structures by Collins (2005) and offers the advantage of preserving the RM approach to RCs in the spirit of Grillo (2009) and Friedmann, Belletti & Rizzi (2009).

Under a *smuggling* account of passive (Collins, 2005), linguists assume the following derivation (49): first, the V and its internal argument are sub-extracted and moved to SpecVoiceP, the functional projection dedicated to the passive derivation. Subsequently, the internal argument is further extracted and moved to the subject position in SpecIP. The derivation therefore takes place through two fundamental steps:

(49) The student$_i$ [$_{IP}$ [helped <the student>$_i$]$_j$ [$_{VoiceP}$ *by* [$_{vP}$ the professor <helped the student>$_j$]]]

The derivation proposed by Collins (2005) is particularly effective in a locality perspective because it allows the movement of the internal argument across the external one, thanks to the fact that the former does not move alone. The internal DP moves within a wider chunk that includes the verb. Under this assumption, the two items to compare in order to compute locality are not two DPs, rather, they are a DP and a VP. In these terms, locality is perfectly observed and the subject cannot intervene in the movement of a VP-chunk. Once the *smuggling* operation is performed, PORs are completed by the movement of the internal argument towards its final position corresponding to the head of the relative, while the subject position in SpecIP is filled by a *pro*. At this point in the derivation, no element can intervene between the internal argument and its target because the former happens to be the higher argument in the derivation and therefore the computation is completed smoothly.

With respect to the experimental results presented above, Belletti & Rizzi (2013) argue that the use of a passive voice entails a lower level of complexity in comparison to ORs, although it requires the movement of a larger (and heavier?) portion of the sentence.

The proposal further supports the hypothesis that the computation of RCs is regulated by locality constraints, given that the derivation proceeds smoother when these are preserved.

5.5.3.7 Resumption

Resumption of the relative head within the RC is a phenomenon that some languages have as an option in their standard variety (e.g., Hebrew and Irish), while others accept it only at a colloquial/sub-standard level (Italian, French). Moreover, at least two criteria seem to have an influence on their distribution: on one side, the kind of RC to be realized with respect to its semantic value, with resumption appearing more often in non-restrictive rather that in restrictive RCs in an implicational order (Bianchi, 2004). The second factor concerns the relativization site within the RCs and predicts that resumption is more frequent in the lower positions of the NP Accessibility Hierarchy (c.f., Keenan & Comrie, 1977), i.e. in correspondence of indirect object extraction. At the other extreme of the scale, resumption is only rarely expected in subject position (Bianchi, 2004). Still, resumption is often only an option and not a condition for grammaticality. In Hebrew, for instance, ORs with a resumptive pronoun (ClOR henceforth) in the embedded clause are equivalent to ORs with gap. In other words, resumption is optional in Hebrew and both forms of ORs are attested and accepted (Friedmann 2008; Friedmann, Belletti & Rizzi, 2009).

Researchers do not agree on the structure of ORs with resumptive clitics and argue on the nature of the clause itself. On one side, it was proposed (McCloesky, 2001, 2002) that ORs with a resumptive clitic (ClOR henceforth) are intrinsically different from ORs, in that they do not entail any instantiation of Wh-movement. McCloesky (2002) starts from the observation that in Irish, the complementizer *aL* characterizes all clauses derived through Wh-movement and even appears in all the Comp positions the extracted element touches upon between the relativization site and its target position. That is to say that *aL* overtly marks the cyclic movement of the Wh-operator. Instances of the phenomenon are found for example in RCs with gap, but crucially not in structures entailing a resumptive pronoun. In contrast, the complementizer *aN* marks ClORs. Based on this observation, McCloesky argues that ClORs are not genuine forms of RCs because they do not show any overt sign of Wh-movement.

As for Romance languages, Guasti & Cardinaletti (2003) adopt the same line of analysis, although Romance languages fail to provide with contrasts of the kind exemplified in Irish.

However, Romance languages also present phenomena in line with the alternative analysis to ClORs, namely CLLD. Belletti (2005, 2006) claims that ORs and ClORs are deeply similar in nature and varies only with respect to minor phenomena, while their basic derivation through Wh-movement is not questioned. The crucial difference between ClORs and ORs is that, in the former, doubling takes place and the moved constituent is a big DP (Kayne, 1991; Uriagereka, 1995) which includes the clitic and the argument; while in the latter, no such doubling takes place and the moved DP is of the basic kind. At this point in the derivation, the asymmetry between ClORs and ORs is self-explanatory: in both cases, the DPs are sub-extracted and moved to their target position, but only in the former (ClOR) a clitic pronoun is stranded in the RC (see also Boeckx, 2003).

Looking at experimental data on Hebrew, resumption does not contribute at improving the comprehension of ORs in children (Friedmann et al., 2009; Costa, Friedmann, Silva & Yachini, 2014); and not even agrammatic speakers benefit from the presence of resumptive clitics (Friedmann, 2008): in both experimental populations no significant difference is attested between ORs with gap and those with resumptive clitics, as both are understood with similar (low) levels of accuracy. Nonetheless, children produce ClORs when they undergo a production task that elicits ORs (Friedmann et al., 2009). Taken together, the data from the two tasks, the one for comprehension and the one for production, suggest that children are familiar with ClORs (they produce it sometimes), but they do not prefer resumption over ORs with a gap. In other words, children do not significantly benefit from the presence of a clitic in the RC.

Resumption is even more interesting in languages that do not envisage this possibility, at least not in their standard variety: that is precisely the case of Italian. ClORs do not belong to the standard variety of Italian: Guasti & Cardinaletti (2003) and Contemori & Belletti (2014) consider the structure to be typical of substandard varieties, although it is not clear with which frequency they are attested in adults. Still, both cited study addressed the issue of resumption. Guasti & Cardinaletti (2003) sample the production abilities of 30 Italian children ranging from 5;1 to 10;0 and find that clitics

distribute accordingly to the kind of prompted RCs. No resumptive clitic appears in SRs, some are attested in ORs and many appear in indirect ORs, to the point that the majority of produced indirect ORs contain some form of resumption. Contemori & Belletti (2014) report relevant percentages of resumption in ORs[51], too. Their children produce two forms of resumption that distribute across all age groups: resumptive DPs and resumptive clitics. In the first case, a full DP is present in the RCs, meaning that the copy is pronounced in its original merging position. In the second case, a direct clitic is present in the pre-verbal position. One more interesting detail, which is left unexplained in the study, concerns the fact that resumption is attested in all age groups, starting from 3 until 8, in percentages that do not significantly evolve in time. Indeed, even DP resumption is still produced by 8-year-old children. Data on adults, on the other hand, do not report even a single occurrence of resumption[52], which means that this strategy is abandoned later in acquisition, although we do not know at which age the strategy is abandoned.

The robustness of the phenomenon in production leads to presume that children might find support from resumption in their computation of ORs. In order to verify the hypothesis, Contemori & Belletti (2014) include ClORs in a sentence-to-picture matching task. If the assumption is correct, children should comprehend ClORs more successfully than ORs. Actually, the hypothesis is weakly confirmed, because a slight advantage for the comprehension of ClORs appears only in older children (6;5 – 8;10 years) (Contemori & Belletti, 2014: 1041).

Taken together, data from comprehension and from production suggest that Italian children use resumption as a form of support in the process of ORs computation. In particular, Contemori & Belletti (2014) ascribe this facilitating effect to the fact that the clitic in the big DP enriches the feature array of the moved DP with respect to the potential intervener (i.e., the subject DP), thus reducing locality effects.

[51] Contemori & Belletti (2014) do not elicit indirect ORs, so that we have no data to compare to Guasti & Cardinaletti (2003).

[52] For the sake of clarity, I must also clarify that the production of resumption (either via DP or clitic pronouns) was highly unlikely due to the low amount of authentic ORs produced in favour of PORs (in which no form of resumption is possible). Moreover, being resumption a substandard option in Italian, adults might intentionally retain from producing any in the experiment.

5.5.4 Interim summary

In section 5.5 I reviewed studies that focus on the processing of relative clauses. The majority of these deals with the asymmetry between SRs and ORs, which is found in many languages, both in production and in comprehension. Starting from the results on studies on adult speakers, L1 acquisition and aphasia, authors have proposed accounts to the asymmetry, which are either syntactic or processing-based. So far, Relativized Minimality is the account that captures the wider variety of patterns of performance.

Given the syntactic similarity between RCs and interrogatives, the debate on the two kinds of structures cannot be completely teared apart. The next section will continue the discussion on the topic by switching the focus to Wh-questions.

5.6 The processing of Wh-Questions

The number of studies on the processing of Italian Wh-questions is rather limited. Nonetheless, background literature touches upon the most interesting issues and allows for a comprehensive overview of the relevant empirical fields: i.e., adult processing (De Vincenzi, 1991a, 1991b, 1996), L1 acquisition (De Vincenzi, Arduino, Ceccarelli, Job, 1999; Guasti, 1996b; Guasti, Branchini, Arosio, 2012) and aphasia (Garraffa & Grillo, 2008). The discussion on methods and results from studies on Italian is enriched by a cross-linguistic comparison with experimental data that cover a variety of languages (mainly English, German and Hebrew). The studies I am going to present differ in the approaches they adopt in order to account for the detected asymmetries (in adult processing and in L1 acquisition) and for the linguistic impairment in aphasia. An overview of the collected data will also provide the opportunity to discuss the accounts proposed for the processing of Wh-questions, among these: the Minimal Chain Principle (De Vincenzi, 1991b, 1996; De Vincenzi et al., 1999), agreement interferences (Belletti & Guasti, 2015; Guasti et al., 2012) and Relativized/Generalized Minimality (Garraffa & Grillo, 2008; Grillo, 2008).

Given the fact that both in the case of RCs and in the case of Wh-questions the discussion mainly concerns the A-bar movement of phrases extracted either from the subject or from the object position, studies and accounts which have already been mentioned above with respect to the computation of Relative Clauses will be mentioned

again. The present section will therefore illustrate to which extent those proposals succeed in providing satisfactory accounts for data concerning Wh-questions too.

5.6.1 Wh-question processing in Italian-speaking adults

The processing of Wh-questions in Italian-speaking adults has been investigated by De Vincenzi (1991a, 1991b, 1996) with the purpose of showing how the computation of Wh-questions can follow processing principles, whose validity can extend cross-linguistically. The author's attention focuses on two asymmetries, i.e. the one between subject and object Wh-questions and the one between Who and WhichNP Wh-elements. Once the two factors are combined, four different kinds of Wh-questions are obtained: Who Subject questions, Who object questions, WhichNP subject questions and WhichNP object questions. In all experimental sentences, the post-verbal argument (either the object in subject Wh-Questions or the subject in object Wh-Questions) is realized as a full lexical DP. Data are collected by using a self-paced reading task. Results reveal the following general tendency: participants read Who-questions faster than WhichNP-questions and subject questions faster than object questions. Moreover, the two factors interact so that the shortest reading times correspond to Who subject questions. Therefore, the asymmetry between subject and object extraction is particularly evident in the case of Who questions. In order to interpret the data, De Vincenzi (1991b) primarily evaluates processing loads; in doing so, the researcher comes to the formulation of the Minimal Chain Principle (De Vincenzi, 1991b):

(50) Avoid postulating unnecessary chain members at S-structure, but do not delay required chain members. (De Vincenzi, 1991b:199)

The principle is composed of two constraints that correspond to economical guidelines to be followed during the computation, in order to create the simplest and most economic representation of the sentence at stake. The first part of the principle requires building only the exact amount of structure that is necessary to represent the encountered linguistic material, thus avoiding the postulation of any redundant element. It follows that the parser resorts to movement only when this is unavoidable and it prefers assuming singleton chains over multiple chains. The counterpart to this first economic guideline warns

against procrastinating the insertion of chain members required in the structure. Holding up unassigned elements in the working memory entails a considerable burden; thus, the computation load can be significantly reduced by structuring the material. It follows that the parser should assign each element to a position in the structure as soon as this is possible. In the case of Who/What questions, the parser starts its computation by encountering an element that clearly signals that movement already took place. At this point, the parser discards the hypothesis of a singleton chain and assumes the existence of at least one other element in the chain (the position where the Wh-element was first merged and subsequently extracted). Afterwards, the parser constructs the simplest possible chain, namely the shortest in terms of distance and number of nodes: this corresponds to a chain entailing subject extraction. In other words, the parser assigns the filler (the Wh-element) the first possible gap in the structure, which happens to be the one corresponding to the subject position.

This strategy reminds of the Active Filler Hypothesis proposed by Frazier & Flores D'Arcais (1989), according to whom, parsing is filler-driven and not gap-driven. In other words, authors disregard the possibility that the parser starts the search for a filler only once an unexpected gap is found. Their proposal actually claims that the parser works exactly the other way around: as soon as an element is identified as a potential filler, the search for a corresponding gap starts. In the case of subject questions, the outcome is successful: the filler is immediately assigned to a gap and the parser can continue the computation with a reduced burden. The lack of unassigned elements in the working memory keeps processing loads low and the parser does not encounter any contra-evidence to its structure, thus completing the sentence processing quickly and accurately. In the case of object questions, the structure contains evidence against the shortest possible chain. In particular, the appearance of a second element endowed with a subject syntactic role and an agent *theta*-role forces the parser to abandon the first analysis and proceed through a reparation. At that point, it is necessary to revise the all structure and to repair it in order to account for the syntactic evidence at hand. In the case of object questions, it derives that the Wh-element must be assigned to a new gap, namely to the one corresponding to the internal argument in the VP. According to De Vincenzi (1991b, 1996), the well-known asymmetry between subject and object extraction stems exactly from this parsing procedure based on economical guidelines which impose to start the

derivation by assuming the simplest structure. In the case of subject extraction, the strategy is successful, while it is not in the case of object extraction. In the latter case, this translates into longer reading times: an experimental measure that reveals the necessity to revise and repair the sentence structure.

The discussed experiments take into account a second factor: the influence of the type of Wh-element in use, namely Who versus WhichNP questions. De Vincenzi bases her discussion on properties of Wh-elements that Cinque (1990) and Rizzi (1990) had previously pointed out in their works. In particular, the author is interested in the different kinds of chains that the two elements require and their consequences for the computation of Wh-questions. Wh-elements of the Who/What type are non-referential in nature and enter chains of the antecedent-government type. Thus, once the parser meets an element of this type, it immediately starts the search for the completion of the chain within the sentence. Moreover, according to Pesetsky (1987) the interpretation of such elements can start only once these fit into the argumental structure of the sentence predicate. In De Vincenzi's view, this enhances the pressure for finding a suitable gap for the element as soon as possible. In conclusion, Wh-elements of the Who/What type prototypically fall under the Minimal Chain Principle described above.

In contrast, Wh-elements of the WhichNP type are referential in the sense that they have the capacity to refer to an element already present in the discourse. This property subtracts the elements from being subordinated to strict locality constraints and allows them to enter binding chains, similar to those applied in the relation between a pronoun and its preceding referent. Furthermore, the parser can start interpreting WhichNP operators as soon as one is met, precisely thanks to their referentiality. Due to the combination of these two characteristics, namely referentiality and presence of a binding chain, the Minimal Chain Principle (De Vincenzi, 1996:120) constrains the processing of WhichNP questions to a lower extent. Indeed, once a WhichNP element is met, the parser assumes that it is in a singleton chain and looks for an element this can be bound to. In particular, it looks for an anaphoric pronoun. If it does not encounter any, as in the case of Wh-questions, the WhichNP element must be integrated in the sentence derivation and assigned to an argument position (the gap). The procedure is more demanding in terms of time and resources if compared to the speed and accuracy with which Who subject questions are processed. On the other hand, though, the interpretation of WhichNP

questions proceeds in a very cautious way, that avoids the assignment of incorrect gaps and therefore prevents the necessity to revise and repair the derivation along the way. In these terms, De Vincenzi (1996) accounts for the better performance of her speakers on Which subject and object questions in comparison to Who object questions.

Summing up, De Vincenzi's (1991a, 1991b, 1996) analysis of how adult speakers process Wh-questions builds on processing procedures that, in turn, are determined by strict economic principles. The assumption of the Minimal Chain Principles also implies that processing asymmetries (like the one between subject and object extraction) are to be attributed not only to strictly syntactic reasons. In other words, syntax is not the primary cause of the observed patterns of Wh-questions comprehension. In her view, the primary cause of processing asymmetries is the application of economical principles to specific linguistic systems. Therefore, De Vincenzi stretches the limits of her analysis beyond Italian, and claims that the principle is universally valid across languages, although its manifestation can differ, depending on the syntactic structures in use. Not only does De Vincenzi consider the Principle universally valid for adults, but she also extends it to monolingual young children. The following section provides some insights into L1 acquisition.

5.6.2 Wh-questions in Italian L1 acquisition

As I anticipated in the previous section, De Vincenzi (1996) elaborates a theory of sentence processing that obeys economic principles. The hypothesis also allows for precise predictions for L1 acquisition: the processing strategies should be present from the first stages of linguistic acquisition on and should not require to undergo maturation over an extended period of time.

In order to prove the hypothesis, De Vincenzi and colleagues (De Vincenzi et al., 1999) sample the comprehension of Wh-questions in 352 children distributed across different age groups between 3 and 11 years. The materials in use imply a list of questions subdivided across four different conditions: Who subject (51a) and object (51b) questions and WhichNP subject (51c) and object (51d) questions.

(51)

a. *Chi sta inseguendo le tartarughe?*
   'Who is following the turtles?'

b. *Chi stanno inseguendo le tartarughe?*
   'Who are the turtles following?

c. *Quale gallina sta inseguendo le tartarughe?*
   'Which hen is following the turtles?'

d. *Quale gallina stanno inseguendo le tartarughe?*
   'Which hen are the chicken following?'

All experimental questions are unambiguous thanks to number agreement between the verb[53] and either one of the two nominal elements in the sentence. Each question is presented in association to an image that includes the characters named in the oral stimuli; characters are represented in different action patterns. Children's task is to listen to each question and to answer by pointing to the appropriate character. The percentage of correct answers produced in correspondence of the four experimental conditions reveals a developmental path characterized by two clear tendencies: Italian-speaking children comprehend subject questions better than object questions and Who questions better than WhichNP questions. Therefore, children are sensitive both to the syntactic function of the extracted element and to the Wh-element in use. A closer look at the data reveals that the asymmetry concerning the Wh-element in use is actually particularly relevant in the case of object questions, while the performance on subject questions does not produce a significant difference from the statistical point of view. This means that subject questions are equally comprehended, independently of the kind of Wh-element in use. However, the two factors interact and participants perform the worst on Which object questions. Overall, the clear subject/object asymmetry in participants' comprehension allows the authors to claim that speakers actively use the Minimal Chain Principle since early childhood and do not need to acquire it: the principle determines the correct comprehension of subject questions and causes mistakes on object questions during acquisition too, as revised for adult performance in the previous section. Concerning the

[53] The disambiguation strategy adopted will be further illustrated and discussed in Chapter 6. For the moment, it will be enough to keep in mind that in De Vincenzi et al. (1999) children deal with unambiguous stimuli.

asymmetry between Wh-elements, authors also illustrate how the disadvantage for the WhichNP condition might derive from abilities that undergo a maturational process. In particular, they refer to the phonological loop[54], which could play a decisive role in the individuation of the discourse referents.

The ability to produce questions has been investigated by Guasti (1996b; Guasti et al., 2012) in Italian-speaking preschool children. In one study, Guasti (1996b) elicits the production of a variety of structures that include Wh-questions, Yes/No questions, argument/adjunct questions and questions with a negation. Results from the study are quite robust and show that the 11 enrolled children (aged 3;1 to 4;8) correctly produce adult-like structures for all the sampled questions.

These results and the ones from De Vincenzi et al. (1999) seem to be in contradiction, given that the former (Guasti, 1996b) does not unveil any difficulty on object questions, while the latter does (De Vincenzi et al., 1999), especially in the case of WhichNP object questions. However, as precisely pointed out in Belletti & Guasti (2015), this contradiction could actually be due to the different modalities adopted in the two studies: either comprehension or production. In comprehension (De Vincenzi et al., 1999), children are forced to compute precisely the experimental sentence proposed, such that processing crashes and comprehension fails in case of difficulties; in production tests (Guasti, 1996b), children have the possibility to resort to alternative strategies when the elicited question is characterized by a high level of complexity. Indeed, children enrolled in Guasti (1996b) produce a wider variety of structures in correspondence to non-subject Wh-questions: for instance, they omit the subject (52) or they dislocate it to the left, before the Wh-operator (53)[55]:

(52) *Che cosa compa?*
What buys
'What does he buy?'

[54] Braddeley (1986) describes the phonological loop as a device of the working memory, which is responsible for preventing the decay of relevant verbal information over the execution of a task. Braddeley, Gathercole and Papagno (1998) also explores its role in language acquisition, in particular with respect to the need to keep new speech input active while processing it.
[55] Examples are taken from Guasti (1996b: 252-253).

(53) *E lui, cosa fa là da solo?*
And he, what makes there alone?
'And he, what does he make there alone?'

Taken together, the two studies unveil difficulties on the processing of object questions, but at the same time they also show that children have the instruments to bypass difficulties and to produce structures that are licit in the adult grammar. In other words, children only rarely produce non adult-like constituent orders.

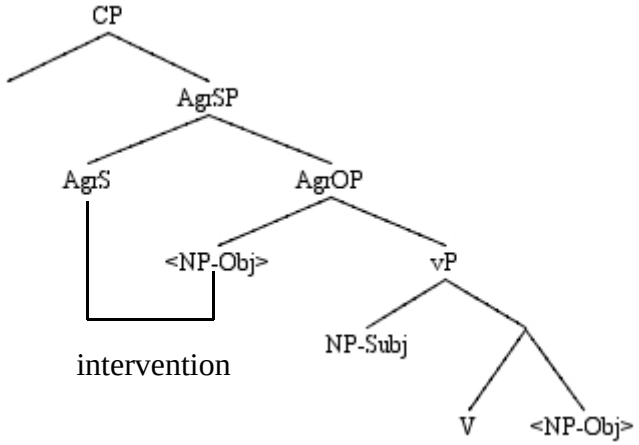
A second study on the production of Italian-speaking preschool children (Guasti et al., 2012) focuses on the two relevant asymmetries in the present discussion, namely the one between subject and object questions and the one between Who and WhichNP questions. The experimental setting uses a puppet and the child is invited to ask questions to it. Overall, collected data confirm the general tendencies already outlined in the data by De Vincenzi et al. (1999) for comprehension: also in production, children produce Who questions better than WhichNP questions and subject questions better than object questions. In order to account for the data, authors build on the interference played by the moved object in the agreement relation that must be established between the subject and the verb. Guasti and colleagues observe a specific kind of mistake, which often appears among the non-target object questions produced by their Italian preschool children, namely verb attraction. Authors notice that children tend to change verb agreement and ultimately sentence meaning when elicited to produce object questions. As a consequence, target object questions of the kind in (54a) are realized as in (54b), namely as subject questions:

(54) a. *Quale bambina baciano le nonne?*
Which girl-SG kiss-3SG the grannies-PL
'Which girl do the grannies kiss?'

b. *Quale bambina bacia le nonne?*
Which girl-SG kiss-3PL the grannies-PL
'Which girl kisses the grandmothers?'

In the target sentence, the object Wh-operator is singular, while the subject is plural and occupies a post-verbal position[56]. Children produce the structure as a subject question: the verb agrees in singular with *Quale bambina*, and the logical subject (*le nonne*) turns into the object. In the vein of Franck et al. (2006) and Guasti & Rizzi (2002), Guasti et al. (2012) claim that the mistake is caused by the attraction played by the object in AgrOP on verb agreement by copying its features on AgrSP. The attraction mechanism is represented below (adapted from Belletti & Guasti, 2015):

(55)



The effect is probably given by the fact that, in order to reach the peripheral position, the object first moves to AgrOP, a dedicated functional position in the IP, which is used to check features on the object. According to the authors, this movement has the capacity to interfere with the agreement process that takes place between subject and verb. This last process, namely subject-verb agreement, takes place in two different steps. The first step is called AGREE and consists of the evaluation of the relevant features (person, number, etc.) between the subject DP (initially merged in SpecvP) and the head of AgrSP. This evaluation takes place under c-command and in a local configuration. The second step usually takes place after movement of the verb to AgrS° and requires, in turn, the movement of the subject to SpecAgrSP. At this point in the derivation, the subject and the verb are in the ideal configuration in order to verify whether they are sharing the

[56] As discussed in section 5.3, in Italian subjects must always occupay a post-verbal position in non-subject questons.

correct features and they check it via Spec-Head movement. AGREE is always completed, while Spec-Head agreement takes place only in those cases in which the subject moves in a pre-verbal position. Although this second step is not essential to the derivation, in Guasti et al.'s (2012) view, Spec-Head plays the crucial role of re-enforcing the agreement. In the case of object questions, two conditions can interfere with the correct execution of the described process. The first one is the presence of the moved object in AgrOP, namely in between AgrS and the verb. As a consequence, the object results in being the closest argument to AgrSP and its features might be incorrectly copied to AgrSP, thus causing an agreement error. This phenomenon is precisely called attraction (Franck et al., 2006), because verb agreement is attracted by the closest argument. The second condition that allows for mistakes in object questions is due to the fact that no element can intervene between the Wh-element and the verb in the final configuration. This prevents the subject (in non-subject questions) from moving higher in the structure; indeed, a post-verbal subject position is obligatory. The subject DP does not move to SpecAgrS and Spec-Head agreement cannot be performed, thus precluding the possibility of re-inforcing verb agreement. The consequences of this configuration, with a preverbal object and a postverbal subject, is that the object attracts verb agreement and endangers the computation of target object questions.

Apart from the intervention of the object in verb agreement, cross-linguistic data further support the role of Relativized Minimality (Rizzi, 1990) for the processing of Wh-questions in children (Friedman et al., 2009 and much subsequent work). These are illustrated in what follows.

5.6.2.1 Wh-questions in L1: a cross-linguistic comparison

Data on the acquisition of Wh-questions in Italian are confirmed by similar findings in other languages. For instance, subject Wh-questions are correctly comprehended at an earlier age with respect to object Wh-questions, which are mastered only later in acquisition: that is the case in German (Biran & Ruigendijk, 2015; Roesch & Chondrogianni, 2015; Schlewensky, Fanselow, Kliegl & Krems, 2000; Schulz, 2013); in English (Avrutin, 2000) and in Hebrew (Biran & Ruigendijk; 2015; Friedmann, Belletti, Rizzi, 2009; see also Friedmann & Szterman, 2011 for children with hearing impairment).

On the other hand, an analysis of production in the spontaneous speech of twelve English-speaking children (Stromswold, 1995) speaks against a strong asymmetry between subject and object extraction and even shows a (weak) advantage of the latter on the former. Stromswold (1995) finds out that children spontaneously start producing object questions at the same time as subject questions and, in a few cases, even earlier; while no child shows the reversed pattern, with a clear temporal advantage for subject questions over object questions. This reminds of the first study conducted by Guasti (1996b) on Wh-question production in Italian children, in which a good performance is registered, independently of the argument role of the Wh-element (i.e., production of adult-like interrogative structures). In particular, the author does not report on specific difficulties and/or agrammatical outputs in correspondence of object wh-questions. Children therefore can cope with object questions, although they tend to realize the subject of ORs in the form of a null subject (*pro*). Through this strategy they reduce the risk for subject intervention in the derivation (as the account based on RM predicts).

If we now look at the asymmetry between Who and WhichNP questions, an advantage for the former emerges. Who questions are generally produced better than WhichNP questions, especially in the case of object questions. Indeed, no particular asymmetry is to be detected within the subject conditions. The combination of the two factors (namely syntactic role and Wh-element type) drives to a peak difficulty in correspondence of WhichNP object questions. Several studies share this observation: De Vincenzi et al., (1999) as well as Guasti et al. (2012) in Italian; for Hebrew, it is attested in Biran & Ruigendijk (2015) and in Friedmann, Belletti & Rizzi (2009); Avrutin (2000) finds the same effect in English-speaking children too. The reversed aspect of this phenomenon is that no major advantage emerges either for subject or object questions within the Who conditions, which signals that the asymmetry neutralizes under specific syntactic conditions, namely the use of bare Wh-elements.

As anticipated above, agreement attraction is not the only threat to the correct realization of an object question. Indeed, the derivation also implies that the object Wh-element moves across the subject position: as already exemplified for relative clauses (see Section 5.5.2), the movement can be blocked on the basis of a strict application of the Principle of Relativized Minimality (Friedmann, Belletti, Rizzi, 2009; Rizzi, 1990). The phenomenon has been already illustrated in what precedes. At this point in the discussion,

it will be sufficient to see how this applies to Wh-questions. For instance, Belletti & Guasti (2015) observe how subject intervention can be responsible for the difficult comprehension and production of object questions in languages like English and Hebrew, especially in the case of feature configurations of inclusion between the two arguments. When the subject is lexically restricted, Who object questions (56) enter in a configuration of disjunction because they are endowed with different features. In contrast, WhichNP object questions enter in a configuration of inclusion (57):

(56) Who do the aunts kiss?
+Q +NP

(57) Which child do the aunts kiss?
+Q, +NP +NP

This hypothesis further predicts that the comprehension of object questions could significantly improve in presence of relevant differences in feature configuration between the two elements. Previous research already proved the hypothesis to be correct with respect to object relatives by manipulating the feature arrays either of the subject or of the object in order to enhance their dissimilarities. Researchers should now proceed with Wh-questions along the same lines of investigation in use for RCs, namely by manipulating the subject features in object questions. This can be done, for example, by introducing a pronominal subject, either in an overt or in a null (*pro*) form or by introducing a mismatch between the two elements with respect to animacy.

Summing up, cross-linguistic evidence confirms the attended subject/object asymmetry but also reveals that this neutralizes under specific conditions. The neutralization is certainly the consequence of reduced difficulty for object extraction, rather than a specific impairment in subject wh-questions. Data shows that whenever no subject/object asymmetry is detected, this is due to an improvement of the performance on the object condition, rather than to a specific deficit on subject questions. In other words, the comprehension of the object condition improves to the level of subject questions, and not the other way around. Cross-linguistic studies reviewed in the present sections show that the processing of object questions can improve whenever the

intervention of the subject is neutralized, either via the presence of a null subject (see children's production in Guasti, 1996b) or, in the case this is realized as a full lexical DP, by the use of bare Wh-elements, so that the two involved elements (subject and object) mismatch with respect to lexical restriction. The data are overall compatible with the account that builds on Relativized Minimality (Friedmann, Belletti & Rizzi, 2009 and much subsequent work). The Principle represents a central node also in an Italian study on Wh-question production in aphasia (Garraffa & Grillo, 2008; Grillo, 2008); section 5.6.3 presents the bulk results of the study.

5.6.3 Wh-questions in aphasia

Garraffa & Grillo (2008) investigate the production of Wh-question in an aphasic speaker of Italian. Their patient's speech is characterized by agrammatism of the Broca's type, as a consequence of a focal lesion in the frontal and parietal areas of the left hemisphere. The study covers both comprehension and production tests on relative clauses, cleft structures and Wh-questions. The performance on the production of Wh-questions is interesting for the present discussion. Authors use an elicitation test based on the model of the studies described above for children (Guasti, 1996b; Thornton, 1990), with the difference that the puppet is substituted by an imaginary person, not represented in the experimental setting. Two factors are introduced in the study: subject versus object extraction and animacy of the Wh-element in use. Bare interrogative pronouns of the Who/What type are in use, so that four experimental conditions are created: Who subject questions, Who Object questions, What subject questions and What object questions. Results clearly reveal a specific impairment on Who object questions (0%), despite a good production of What object questions (75%), What subject questions (83.3%) and Who subject questions (75%). In other words, Who object questions is the only condition in which no target output is produced.

In order to explain the results, the authors focus on the feature sets of the involved elements. In particular, they assume that the [+wh] feature is disregarded in the computation because of the reduced processing resources agrammatic speakers can count on. In their view, impaired speakers find particularly challenging to activate and maintain the activation of discourse-features, which, therefore, do not activate at all or fade away quickly. The assumption is known as Generalized Minimality and it predicts the

application of the Principle of Relativized Minimality (Rizzi, 1990) on a reduced (and therefore generalized) set of features. Once [+wh], the feature with a crucial role in the attraction of the moved element to its target position is omitted from the computation, the possibility for the subject to function as an intervener in the expected movement increases. The enhanced risk for intervention is indeed due to the fact that the feature arrays of the target, the intervener and the moved element are now very similar and the subject results in a suitable candidate for the target position. This condition corresponds to Who object questions and is particularly problematic for agrammatic speakers in Garraffa and Grillo's (2008) view. Nonetheless, intervention is avoided with more ease whenever at least one feature (not of the discourse type) contributes in differentiating the target position and the moved constituent (the object Wh-element) from the potential intervener (the subject). This last described condition is precisely the one corresponding to the What object condition: in this, a major role is played by animacy. The [-animate] feature on What suffices in making the Wh-element different enough from the subject. The asymmetry between the inanimate subject and the animate object therefore allows for the movement of the object across the subject, and intervention is neutralized.

Summing up, Garraffa and Grillo (2008; see also Grillo, 2008) claim that reduced processing resources are the primary cause of the deficit in agrammatic speakers (i.e. failed production of Who object questions). Indeed, reduced resources do not allow for the activation of all features needed in the derivation and syntactic underspecification follows. This underspecification corresponds in their view to the omission of discourse features (i.e. [+wh]).

This proposal is particularly tempting in the case of PADs, who also suffer from a generalized impairment of their cognitive abilities and are therefore likely to have difficulties at maintaining the activation of complex feature arrays over an extended timespan. The comprehension task on Wh-questions presented in Chapter 6 will help in verifying whether the proposal of a Generalized Minimality at work in sentence processing is adequate also for the description of comprehension abilities of speakers characterized by Alzheimer's disease.

However, a criticism to Generalized Minimality can already be formulated at the theoretical level. The proposal builds on the idea that discourse features are the specific target of the impairment as they are omitted from syntactic computation. The difficulty

on object questions, object relatives and object clefts arises from the absence of this crucial feature, which, in principle, should play the role of distinguishing the element to be moved, in this case the object, from the subject. Actually, the basic function of this feature is not the creation of a mismatch between the elements involved; rather, it triggers A-bar movement of the constituent endowed with the same feature. If this feature is omitted, the derivation of a Wh-question or of a relative clause should be equally blocked, independently of the target element (the subject or the object). The case of What object questions can exemplify the observation. In Garraffa and Grillo's (2008) view, agrammatic patients succeed in producing What object questions thanks to the animacy features which is represented in a configuration of mismatch in the relevant positions. It is hard to imagine though, how the object phrase could move towards an A-bar position although it does not entail any discourse feature, only on the basis of the [-animate] feature it includes. The same reasoning could be applied to subject questions too: following the same line proposed by the authors, it follows that subject questions are realized despite the absence of the relevant discourse feature. The authors fail to explain how such structure can be derived in absence of scope-discourse features.

In order to consider alternative explanations to the deficit in aphasic patients, it is therefore necessary to open the overview to studies concerning languages different from Italian. This is precisely the topic of the next section.

5.6.3.1 Wh-questions in aphasia: a cross-linguistic perspective

There exists few studies that analyse the agrammatic speech of patients affected by aphasia of the Broca's type and find results that resemble data from language acquisition. For instance, Hickok & Avrutin (1996) and Tait, Thompson & Balard (1995) find a specific impairment on WhichNP object questions, despite a rather well-preserved processing of the remaining three experimental conditions, namely WhichNP subject, Who subject and Who object questions. Hickok & Avrutin (1996) first interpret the data by building on the Trace Deletion Hypothesis (Grodzinsky 1986, 1990), a structural model of agrammatic speech: according to this, aphasic speakers fail at maintaining the activation of traces at the deep structure level and therefore, in case of moved elements as in interrogatives or in relative clauses, they are unable to form the syntactic chain that allows for the realization of the proper sentence structure. In other words, according to

this account, patients cannot retrieve the gap from where the fronted element was moved. The account also predicts that patients resort to oversimplified strategies and, for instance, assign *theta*-roles to arguments according to a canonical linear order. The strategy accounts for the subject/object asymmetry, as *theta*-role assignment works fine in subject-first structures (i.e. subject questions), but it crucially fails in object-first structures (i.e. object questions). Moreover, the specific impairment on WhichNP questions is initially interpreted as a specific deficit at building binding chains (WhichNP), contra relations of the antecedent-government type (Who), which are parsed with higher accuracy. After focusing on the syntactic properties that distinguish WhichNP and Who questions, Avrutin also revises his interpretation by taking into account the semantic differences brought up by the two Wh-elements.

In a subsequent work, the author compares the performance of English-speaking children to the one of six aphasic patients (Avrutin, 2000); he observes that both groups produce a significant higher rate of mistakes on WhichNP object questions and concludes that there must be a common account for the difficulties suffered both by young children and by agrammatic speakers. The deficit, according to Avrutin's (2000) hypothesis, is of the cognitive type: children and patients count on reduced processing resources with respect to healthy adult speakers. The shortage of processing resources becomes problematic in the case of WhichNP questions because the processing of these specific structures requires the integration of information coming both from the syntactic and from the semantic systems, while for all other conditions (Who questions and WhichNP subject questions), it suffices to deal with the syntactic structures. The integration of semantic information in the case of WhichNP questions derives from the fact that the Wh-element is referential in nature and therefore brings up discourse presuppositions in the processing. This factor is particularly relevant in the case of WhichNP object questions, while WhichNP subject questions can simply be analysed by proceeding with a canonical assignment of *theta*-roles.

However, in the light of previous discussion, it is quite evident that the results discussed so far are perfectly compatible with a processing strategy based on a stricter version of Relativized Minimality (Rizzi, 1990), which blocks the derivation of WhichNP object questions but allows the one of Who object questions on the basis of a mismatch in features (in this case, a mismatch in lexical restriction) between the intervening subject

(lexically restricted) and the bare Wh-element extracted from the object position. In this view, the results presented in Tait et al. (1995), Hickok & Avrutin (1996) and in Avrutin (2000) are perfectly compatible with those collected by Garraffa and Grillo (2008) with respect to the effects of aphasia of the Broca's type in Italian. Despite this, the discussion on aphasic speakers is never exempted from the issue of variability (Caplan et al. 2001; Caramazza et al. 2001; Drai et al. 2001); indeed, other studies fail at replicating the same results. For instance, Thompson, Tait, Ballard and Fix (1999) run similar tests on four aphasic patients, who actually perform very differently. Only one patient performs as predicted by Hickock and Avrutin (1996), while the other three do not. The answers provided are so different altogether that the individuation of a common alternative pattern is not possible, thus redirecting the discussion towards the impossibility of drawing clear conclusions with respect to forms of agrammatism that arise from similar (but still different) lesions to the left hemisphere.

There is one further hypothesis on the modalities of impairment of agrammatic speakers that has received attention and could also be discussed with respect to the performance of PADs. The proposal is known as the Tree Pruning Hypothesis and was advanced by Friedmann (2002), while discussing data on the production of questions in Hebrew- and Palestinian Arabic-speaking agrammatics. First in spontaneous speech and afterwards trough an elicitation task, Friedmann observes that agrammatic speakers deal well with Yes/no questions, while they are badly impaired on Wh-questions. In particular, within this second type, patients show a slight advantage for adjunct questions (especially Why-questions) in comparison to argument questions. Moreover, interestingly enough, no asymmetry between subject and object is detected, as both are impaired to the same extent. Friedmann (2002) interprets the data building on a specific syntactic characteristic of the Semitic languages at stake. Both Hebrew and Palestinian Arabic require movement of the Wh-element and of the verb to the CP-layer in case of Wh-questions, as many other languages do (for instance, English and Italian). In contrast, Yes/No questions do not require such movement and the Verb stays in TP. Therefore, the author hypothesizes that agrammatic speakers manage to project a reduced syntactic tree, which only projects as far as TP. Everything above that node can neither be projected nor be targeted by movement (from here the reference to tree pruning). With respect to Italian-speaking PADs, the hypothesis therefore predicts very poor performance of all kinds of Relative

Clauses and Wh-questions, considered that all sampled structures (see Chapter 6 and 7) entail phrase movement to the CP layer in Italian. The hypothesis will be further discussed in Chapters 6 and 7, in light of the data from the experimental tasks with PADs.

5.7 Conclusions

In the present chapter, I reported an analysis of Italian Wh-questions according the following lines: the Wh-operator is extracted from its argument position and raised to the upper portion of the sentence (Chomsky, 1977) in order to satisfy the Wh-criterion (Rizzi, 1996). The operation leaves a gap in the sentence, which corresponds to the point in which the element is interpreted. The satisfaction of the Wh-criterion requires a further step, namely the movement of the inflected verb from I to C. Once the two elements endowed with [+wh], the wh-operator and the verb, are both in CP, they can enter the Spec-Head relation required by the Wh-criterion. The operation blocks the subject DP from raising, so that this is left in a post-verbal position or realized as a phonetically null *pro*.

Empirical studies on the processing of Wh-questions have addressed mainly two questions: the role of the argument position from where the Wh-element is extracted and the type of Wh-element in use. The first question deals with the asymmetry between subject and object Wh-questions, while the second one concerns the differences between bare and lexically-restricted quantifiers, i.e. between operators of the Who and of the WhichNP type. Italian studies on comprehension in adult speakers (De Vincenzi, 1991a, 1991b, 1996), in L1 acquisition (De Vincenzi et al., 1999; Guasti, 1996b; Guasti et al., 2012) and in aphasia (Garraffa & Grillo, 2008; Grillo, 2008) all found a specific low performance on WhichNP object questions. The robust result is in line with the Principle of Relativized Minimality, which predicts an intervention effect of the subject on the movement of the object towards its target position in CP. This is particularly true in the case of object questions in which both the subject and the object share relevant features, i.e. both phrases are lexically restricted and take animate referents.

Alternative accounts have also been proposed in order to describe the performance of aphasic patients: the Trace Deletion Hypothesis (Grodzinsky, 1986, 1990), the Tree Pruning Hypothesis (Friedmann, 2002) and Generalized Minimality (Grillo, 2008). According to Grodzinsky, aphasic speakers experience specific difficulties at maintaining active traces and therefore fail at individuating the gap from which the Wh-element was

extracted. Grillo (2008) hypothesizes that aphasics are unable to maintain the activation of complex feature sets over an extended time span and therefore tend to omit features of the scope-discourse type from the computation. The omission has decisive consequences on the computation of object Wh-questions whenever the object and the subject only distinguish for the (omitted) feature. Finally, Friedmann (2002) suggests that aphasic speakers might count only on a reduced portion of the syntactic structure, which does not comprehend the sentence left periphery. The hypotheses discussed here will be further analysed in Chapter 6, in light of the collected data, in order to see how they can account for the performance of PADs.

In the present Chapter I have also dealt with the syntactic structure of RCs. In particular, I have reviewed studies that speak in favour of an analysis that sees RCs as CPs that occupy a functional projection within the DP corresponding to the relative head. The position that hosts the relative CP is either the complement of the DP (as proposed by Kayne, 1994) or a functional projection high in the DP (as claimed by Cinque, 2008, 2014). In both views, the internal argument raises to its target position in the CP-layer of the relative clause. In case the movement starts from the internal argument of the RC, locality effects arise. Indeed, a RM approach to relative clauses (Friedmann et al., 2009; Grillo, 2009) accounts for the asymmetry between subject and object RCs with a precision that memory- and processing-based accounts lack (De Vincenzi, 1991a; Frazier & Fodor, 1978; Kidd et al., 2007; Mc Whinney & Pléh, 1988; amo). The well-attested phenomenon is indeed ascribed to the fact that the object DP moves across the subject DP and shares with it part of its feature array. The intervening subject makes the object movement difficult.

With respect to this, I have offered and overview of the factors that improve the comprehension and/or production of ORs. Of course, they are not all equal. As pointed out by Sanfelici et al. (2014), it might well be the case that some of the strategies adopted by children in order to avoid the production of the elicited ORs are universal (i.e. roles reversing), while other strategies might be language specific. I focused my attention on factors that, in my opinion, are relevant for Italian, either because they have already been proved to play a role, or because there are signs that indicate that this could be the case. Summarizing, number mismatch, and passive voice (PORs) improve the comprehension of ORs. Both factors enhance the differences between the moved constituent and the

intervener, thus favouring extraction. As for resumption, the way it operates is not completely clear yet, but Contemori & Belletti (2014) proposed that it might be the case that the presence of a stranded clitic is relevant for the evaluation of feature arrays.

Two distinct studies (Adani et al., 2010; Belletti et al., 2012) robustly confirm that gender does not play any role in Italian at enhancing the difference between the feature arrays of the involved position, while an effect is visible in a language like Hebrew, which has gender agreement on the verb. This asymmetry is interesting in a cross-linguistic perspective, because it makes clear that not all features always enter the locality computation, rather only those that contribute to triggering movement do (Belletti et al., 2012).

Finally, data concerning the role of animacy in Italian is still very poor, as the feature was included only in one elicitation study with adults (Belletti & Chesi, 2011) and actually did not produce any remarkable effect.

The body of information now available on the derivation and the processing of Wh-questions and Relative Clauses is quite wide and rich. The information provided and summarized here are indeed the premises for the design of the experimental studies I am going to present in chapters 6 and 7.

6 Wh-QUESTION COMPREHENSION IN PATIENTS WITH ALZHEIMER'S DISEASE

6.0 Introduction

In the present Chapter, I will present data from a Wh-question comprehension task performed by a group of Italian PADs, whose participants are characterized by different levels of dementia. The collected data will represent the starting point for the discussion on the syntactic deficit in PADs.

A number of reasons motivate the decision to include Wh-questions in the study. First, the selected structures entail an operation of extraction, which moves a constituent out of its argument position and takes it to the clause left periphery through Wh-movement. Previous studies on L1 acquisition and aphasia (see Chapter 5 for an overview) used Wh-questions in order to sample speakers' competence. Overall, they revealed that the operation of extraction increases the level of complexity in computation with respect to unmarked declarative clauses. The claim stems from the observation of increased difficulties in children and aphasic speakers in the comprehension and production of different kinds of Wh-questions, according to patterns that depend on the asymmetry between subject and object extraction, the kind of Wh-element in use (Who vs WhichNP), as well as the experimental task. Therefore, previous research shows that interesting and meaningful data can be collected through studies that include Wh-questions among their experimental material. Wh-questions have already proved to be a fruitful point of observation for the syntactic competence of different groups of speakers. For this reason, I will now sample PADs on this kind of structures.

Previous studies on Wh-question processing proposed different accounts in order to interpret the collected data. This leads to the second reason for introducing Wh-questions in the study, namely the opportunity to compare PADs to other experimental groups and verify whether their performance can be accounted for by any of the previously formulated proposals.

One further reason for starting investigating the syntactic competence of PADs by considering direct Wh-questions, is that these allow for the observation of Wh-movement in root clauses. I start by sampling the comprehension of sentences entailing Wh-movement with a test on Wh-questions, based on the assumption that the level of computational difficulty should be lower in root clauses than in embedded clauses. In Chapter 7, I will address the issue of the computation of Wh-movement in embedded clauses with a test on the comprehension of Relative Clauses.

As for the results of the present task, I can anticipate that patients have different levels of accuracy according to their level of dementia. Moreover, they are sensitive to the manipulation of different syntactic features, thus allowing Wh-movement but only under certain conditions.

The present chapter is organized as follows. The first section addresses the research questions that motivate and guide the study. After that, the experimental task in use will be illustrated in details in section 6.2. The section describes the design, the material, the participants, the administration and the coding procedures. Section 6.3 presents the data collected. Finally, I will discuss the results obtained in section 6.4. The last section contains an outlook and some open questions for further research.

6.1 Research questions

In Chapter 2 I reviewed previous studies on the linguistic deficit in PADs. Results from studies on sentence processing pointed out the role of Working Memory (Almor et al., 1999; Small, Kemper & Lyons, 2000; Kempler et al., 1998), but little data on the syntactic competence were presented. For this reason, I address the following questions:

i. Do PADs suffer from syntactic impairment?
ii. How can the impairment in PADs be accounted for?
iii. At which stage of the disease are PADs affected by syntactic impairment?

The first question is devoted to the search for reliable evidence about the status of syntactic computation in PADs. Given the overall cognitive impairment that appears in conjunction with the disease, an impairment is expected in the performance of the experimental group. However, the important aspect to be determined is whether PADs

perform poorly in a generalized fashion, or whether their poor performance reflects signs of sensitivity to different interrogative structures, in particular with respect to two factors: the extraction site (subject vs object questions) and the type of Wh-element in use (bare vs lexically restricted). The discussion will take into consideration the participants' reactions in correspondence to different levels of complexity[57]. Only in case of clear signs of increased difficulties in correspondence to specific syntactic manipulations, data can show that there exists a specific syntactic impairment in PADs; while in the case of a generalized poor performance in all conditions, the discussion should concern only the cognitive impairment.

Moreover, I will verify whether previous accounts regarding the computation of Wh-questions can appropriately describe the performance pattern found in PADs. In particular, I will measure to which extent the Principle of Relativized Minimality succeeds in correctly predicting the participants' behaviour. As reviewed in Chapter 5, RM can correctly account for the results from a variety of studies concerning both very young and adult participants; this is the first model to which data from PADs should be compared.

In the case patients' performance deviates from that of controls, the discussion will also take into consideration accounts presented in studies on aphasic patients: among these, Generalized Minimality, the Trace Deletion Hypothesis and the Tree Pruning Hypothesis. According to Grillo (2008, 2009) aphasic patients base their syntactic computation on a reduced number of features as a consequence of their impaired cognitive abilities. Therefore, their computation considers underspecified sets of features, a fact that inspires the author to rephrase Relativized Minimality into Generalized Minimality. Second, I will evaluate whether participants' difficulties are determined either by a specific impairment at retrieving the copy of the moved element (i.e., at reconstructing the chain and individuating from where the element was moved out first) as suggested by Grodzinsky (1986, 1990); or by a structural disruption that makes the CP unavailable as landing site for A' movements, as Friedmann (2002) proposed in the so called Tree Pruning Hypothesis.

[57] The level of syntactic complexity for each condition will be judged based on the previous literature on the topic, as reviewed in Chapter 5.

Being Alzheimer a degenerative disease, talking about deficit in PADs can be rough and approximate if observations are not accompanied by precise references to patients' level of dementia. Although most linguistic studies ignored the level of dementia as a factor for performance evaluation, the issue is highly crucial, due to the fact that the set of symptoms that appears in conjunction with the disease can vary along with the worsening of the disease: not all symptoms appear together and some can vary in time. I will address the issue by including in the study patients that subdivide into three different groups according to the level of severity of their dementia. In doing so, I would like to observe and measure whether the syntactic competence decreases and how it eventually changes along with the worsening of the disease.

Finally, results will also offer the opportunity to share some theoretical observations on the syntactic structures in use.

6.2 The experimental task

The present study on PADs aims at sampling their comprehension of different kinds of Wh-questions through a sentence-to-picture matching task. I will illustrate in details the characteristics of the task in what follows. Before that, it is useful to spend a few words on the main constraints to the task design. There are some difficulties when it comes to testing the comprehension and/or production of Italian Wh-questions. The problem is the ambiguity that characterizes sentences of the kind in (1).

(1) *Chi bacia la ragazza?*
Who kiss-2SG the girl
'Who kisses the girl?' or 'Who does the girl kiss?'

The structure in (1) is ambiguous because it can correspond both to a subject and to an object Wh-question. In other words, one possibility is that the interrogative pronoun *Chi* first merges in the external argument position as subject of the sentence and then moves to the interrogative criterial position in CP. The DP following the predicate (*la ragazza*, 'the girl') should therefore be interpreted as the object. Alternatively, *Chi* first merges in the position corresponding to the internal argument and then moves to the CP, thus producing an object Wh-question. In the second case, the lexical DP (*la ragazza*) is to be

interpreted as the subject of the sentence. It is precisely the obligation for the subject to occupy a post-verbal position (which linearly overlaps with the unmarked object position) that gives rise to the ambiguity of the structure in (1). Subject and object questions are characterized by the same linear order of constituents, despite very different underlying structures.

The context disambiguates the meaning in spontaneous speech. However, in an experimental setting with out-of-the-blue sentences the ambiguity is highly problematic. No morpho-syntactic evidence can bias the interpretation towards either a subject or an object question. Tasks with questions concerning pairs of single characters imply the use (or the elicitation) of ambiguous sentences.

Previous studies on Italian already faced the problem exemplified in (1) and implemented different solutions. I will briefly review three ways of avoiding ambiguity in Italian Wh-questions in what follows.

The first strategy builds on semantic disambiguation, as exemplified in (2):

(2) *Quale segretaria licenzia il capo?*
Which secretary fire-3SG the boss?

For instance, De Vincenzi (1991b, 1996) uses this technique in order to test the processing of Wh-questions in adult speakers of Italian. The sentence is still ambiguous at the syntactic level though, because it can express both meanings in (3):

(3) a.Which secretary does the boss fire?
b. Which secretary fires the boss?

The semantic value of the two nouns in use and world-knowledge allow for a disambiguation in favour of the reading in (3a), rather than the one in (3b). However, this solution is not eligible for a study with PADs, given the severe impairment that patients suffer with respect to semantics, pragmatics and world-knowledge. Running a syntactic test that crucially relies on patients' semantic and common knowledge would represent a serious confounding factor. This brought to the discard of the first disambiguating strategy.

A second attempt of testing Wh-questions with unambiguous questions builds on the use of number mismatch: in case the two DPs mismatch in number, verb agreement with the subject clearly biases the reading towards the element that should be analysed as subject in the Wh-question, independently of its position in the structure (pre- or post-verbal). The examples in (4) and (5) clarify the strategy:

(4) *Quale bambino bacia i nonni?*
Which child-SG kiss-3SG the-PL grandparent-PL
'Which child kisses the grandparents?'

(5) *Quale bambino baciano i nonni?*
Which child-SG kiss-3PL the-PL grandparent-PL
'Which child do the grandparents kiss?'

In (4) the verb agrees in singular number with the interrogative phrase, so that the sentence can be read only as a subject question. While in (5), the verb agrees in plural so that the only eligible subject in the clause is the post-verbal plural DP (*i nonni*), and the sentence automatically acquires an object question reading. In this case, disambiguation clearly takes place at the syntactic level, as from the semantic and pragmatic point of view the sentence is perfectly reversible. For this reason, number mismatch could represent a valid solution for the present study. The efficacy of the technique has been proved in studies with Italian-speaking children (De Vincenzi et al., 1999; Guasti et al., 2012); however, this possibility was discarded, because including number mismatch as a factor in the task would make the visual material in use much more complex. I will address the issue again in Chapter 7, but for the sake of the present discussion, it is enough to mention that number mismatch multiples the number of drawings and characters that have to be inserted in the visual material in order to carefully control for the influence of the factor in a sentence-to-picture matching task. That is because syntactic foils must represent all possible alternatives (singular subject and plural object, plural subject and singular object, and *vice versa* with reversed roles). This consequence is undesirable with participants who might suffer also from an impairment at the visual-perceptive level and have post-interpretative difficulties (Almor et al., 2009; Hodges et al., 1991).

The third kind of technique for disambiguation relies on verb agreement too and it consists in introducing a first or a second person referent in the discourse. For instance, Garraffa & Grillo (2008) use a second person singular in their elicitation task of subject and object Wh-questions. The target sentences of their task corresponde either to the example in (6) for subject questions or to the example in (7) for object questions:

(6) *Quale ragazza ti abbraccia?*
Which girl you-ACC-SG hug-3SG
'Which girl hugs you?'

(7) *Quale ragazza abbracci?*
Which girl hug-2PS
'Which girl do you hug?'

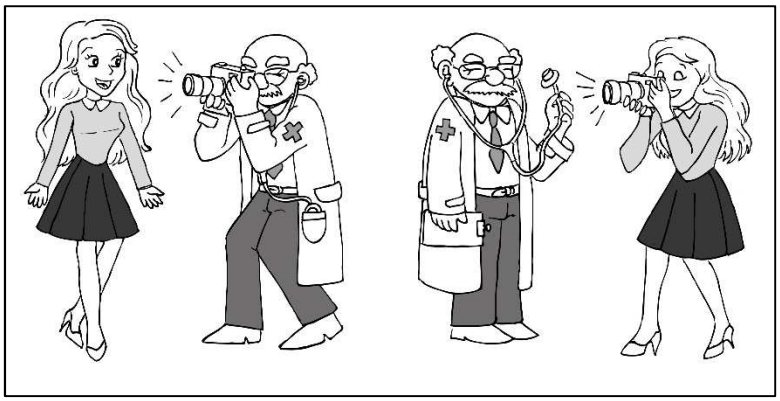
In (6) the interpretation is biased by two elements: the direct object clitic pronoun at the second person singular and the verb agreement form at the third person singular. Based on these two elements, the only possible interpretation is the one corresponding to the Wh-element as subject of the sentence (in name of the person and number agreement with the verb form) and therefore to a subject Wh-question. In (7), the second person singular agreement form of the verb straightforwardly determines the interpretation of the sentence as an object question. As the overt DP and the verb do not agree in person, the only eligible subject for the sentence is the second person singular, syntactically represented in the structure by a silent *pro*. The object question interpretation unambiguously derives from this.

Finally, I selected this last disambiguation technique for the present study: it seems particularly appropriate for a study on PADs, as it overcomes both confounding factors I previously pointed out (i.e., relevance of semantics and visual complexity). However, it poses some difficulties for the implementation. I will discuss the issue in the following section (6.2.1), in which I describe the task design.

6.2.1 Design

The task in use includes both visual and oral materials: for each experimental trial, participants listen to a question and simultaneously see a drawing. The image presents two pairs of characters, who are both engaged in the same action, but with reversed roles (see Figure 6.1). Participants perform the task by listening to the question and answering to it: this has to be done by pointing to the target character in the drawing.

Figure 6.1: Example for drawings in use in the task.



In order to reduce the ambiguity of Wh-questions, all clauses include a singular first person referent, in the role either of the subject or of the object. In the former, this is made explicit by a singular first person agreement on the verb form (and a null *pro*). In the latter, the object appears in the form of a singular first person clitic pronoun (examples follow). Each question is designed as if the character itself asked it.

In order to favor the interpretation, a female character that appears in all drawings throughout the experiment embodies the first person singular referent. Participants are familiarized with the character at the beginning of the session. The experimenter takes care of presenting it to the participants with the introduction in (8):

(8) "We are now going to see the picture album of this young girl. Can you see her? She has long blonde hair and wears a grey skirt. She will appear in all of the drawings we are going to see. She will ask you questions. Please answer by pointing to the correct character."

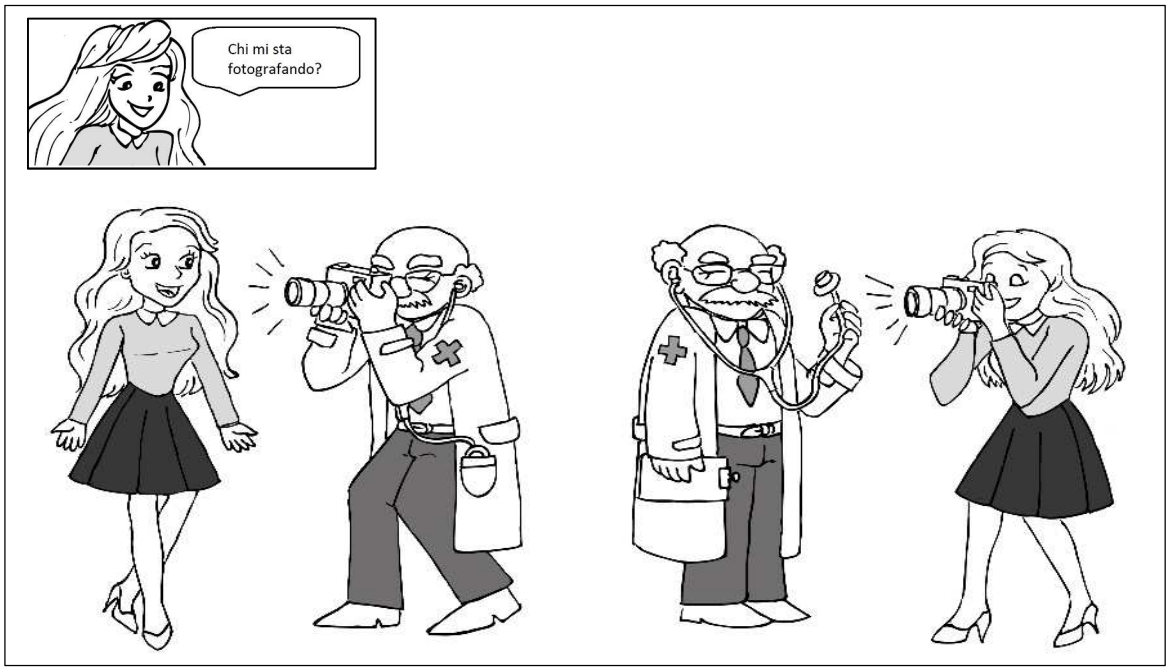
In order to keep the information active about who is asking questions (and therefore who should be identified with the first person singular), the female character is reproduced also in the upper left corner of each visual stimulus with a speech bubble including the experimental sentence.

Figure 6.2: Add-on icon, which appeared in all visual stimuli.

The visual stimuli are presented as a Power-Point Presentation on a laptop (screen size 15.4").

Images are presented in a landscape-oriented format on the full screen and appear as in Figure 6.3:

Figure 6.3: Example of visual stimulus in use in the task.



As for the oral part in each stimulus, the experimenter reads aloud the questions for the participants as soon as the images appear on the screen. The speech bubble turns up in the visual stimulus with the mere purpose of reminding the participants that it is the

female character who asks the questions, according to the task design. However, participants are not expected to read questions by themselves. The decision not to let patients and controls read by themselves is determined by the desire to avoid the insertion of a potential confounding factor in the task. It is known from previous research (Luzzatti, Laiacona & Agazzi, 2003 for Italian) that PADs might suffer from a reduced ability at reading. Therefore, a syntactic study that crucially relies on patients' reading ability could generate unclear data and leave open the question on the actual source of impaired performance. Poor executions could indeed be attributed both to defective reading abilities and/or to impaired syntactic competence.

In each pair of images, the girl performs the action on the other character in one of the two images; in the other image, roles are reversed and it is the second character who performs the action on the girl. According to this design, only either one of the two images corresponds to the situation assumed in the Wh-question and provides participants with the correct answer. In contrast, the second image represents a syntactic foil, in the sense that it corresponds to a sentence interpretation characterized by reversed roles: it signals whether the subject question has been interpreted as an object question and *vice versa*.

The test includes eight pairs of images, which the experimenter shows four times in correspondence to the four different conditions included in the study (see Section 6.2.2). Each experimental session starts with a warming-up phase that includes up to four trials[58] and helps the participants to familiarize with the task and the materials. It results that each participant completes up to 36 trials: 32 experimental trials plus a number of warming-up trials, which varies between two and four, depending on the participant's needs.

Further measures were taken in order to avoid confounding factors. In the visual material, the order of the two pairs of images alternates, so that in half of the trials the target character appears on the left side and in the other half to the right. No pair of images is ever presented twice in a row. Moreover, trials are randomized with respect to the syntactic conditions: Wh-questions characterized by the same structure are never repeated more than twice in a row.

[58] The four trials in the warming-up phase are not counted among the experimental trials, which are 32 (8 trials X 4 conditions) for all participants.

In half of the pictures the second character is a female, in the other half a male. The balanced repartition is actually not relevant for the present task, but it was adopted nonetheless in order to leave open the possibility to test the role of gender mismatch with the same material in feature research. Drawings only include figures with a clear characterization, in order to distinguish the secondary characters from the main one (namely the woman that embodies the first person singular). Concerning this issue, the following characters appear in the eight pairs of images in use: a little boy, a doctor, a priest, a traffic police officer, a little girl, a bride, a nun and a female nurse.

As for the verbs in use, these are: *fotografare* ('to take a picture')[59], *asciugare* ('to dry'), *salutare* ('to greet'), *baciare* ('to kiss'), *spingere* ('to push'), *bagnare* ('to wet'), *coprire* ('to cover'), *pettinare* ('to comb') and *accarezzare* '(to caress'). They are all transitive verbs, which take a direct object. Moreover, they are reversible also from a semantic point of view. In order words, with respect to Figure 1, taking a picture is an action that can be performed by a woman or by a doctor with the same level of probability at the semantic and pragmatic levels. In correspondence of the very same picture, the use of a verb like *to examine* would strongly bias the interpretation towards the doctor in the agent *theta*-role and the woman in the patient one, while this is not the case with the action of taking a picture. None of the verbs in use posits clear expectations on which of the two characters is more likely to perform the action.

The experiment does not include any filler. The main reason for this is the precise will to keep the test administration as short as possible, in order to avoid overloading patients.

Finally, the task crucially relies on the ability to interpret the female character in the picture as the first person singular in sentences. This means that a certain ability to infer someone else's mental states is required. Previous studies showed that PADs usually have spared Theory of Mind, but also that impairments can concern complex tasks with high demands on Working Memory (Castelli et al., 2011; Gregory et al., 2002; Laisney et al., 2011). Based on these studies, I expect PADs to be able to comprehend and perform the task, although it cannot be completely excluded that the operation of abstraction represents a confounding factor in the task.

[59] The Italian verb *fotografare* takes a direct DP as object, not a PP as in the case of the English 'to take a picture of someone'.

6.2.2 Material

As I anticipated in the previous sections, the task includes four different conditions. These are obtained by combining two different factors: the extraction site (subject vs object questions) and the kind of Wh-element in use (bare *Chi* vs lexically restricted *Quale* Wh-phrases). As a result, the experiment includes Who subject questions (WhoS), Who object questions (WhoO), Which subject questions (WhichS) and Which object questions (WhichO). Questions in (9) exemplify the four conditions:

(9) a. WhoS
Chi mi sta fotografando?
who I-ACC.1SG AUX-3PS photograph-PROG
'Who is taking a picture of me?'

b. WhoO
Chi sto fotografando?
Who AUX-1PS photograph-PROG
'Of whom am I taking a picture?'

c. WhichS
Quale dottore mi sta fotografando?
Which doctor I-ACC.1SG AUX-3PS photograph-PROG
'Which doctor is taking a picture of me?'

d. WhichO
Quale dottore sto fotografando?
Which doctor AUX-1PS photograph-PROG
'Of which doctor am I taking a picture?'

In (9a) the reading is biased by two elements: first, the direct object clitic *mi* in first person singular as object of the clause; and second, a third person singular agreement on the verb. The combination of the two restricts the possible interpretations to one: a subject reading, so that the question concerns who is performing the action on the female protagonist (that

embodies the first person singular). The same mechanism repeats in (9c), which differs only for the Wh-element in use, namely a WhichNP phrase. A word of caution is necessary here. Unfortunately, we do not know whether clitics represent a source of difficulty for PADs because the issue has never been investigated so far. Therefore, I am aware of the confounding factor they potentially represent.

As for the object questions in (9b) and (9d), the reading is determined by a mismatch in person between the Wh-interrogative *Chi* or *Quale*NP, which both correspond to the third person, and the verb form, which agrees at the first person singular. Thus, the verb form makes clear that the subject is the first person, corresponding to the female character in the pictures, and this biases the interpretation towards an object reading.

6.2.3 Participants

PAD participants in this experiment are a subgroup of the participants in the experiment presented in Chapter 4. Out of the 41 patients initially enrolled, only 36 could take part in this second task. Two patients could not complete the task because of independent reasons; three participants could not perform it either because of behavioural disorders or of the incapacity to comprehend the instructions. The performance of PADs was compared to the one of 21 controls, with matching values for age and level of education, but spared cognitive abilities[60]. The control group for this experiment includes the same participants enrolled in the task on grammatical gender retrieval in Chapter 4.

Table 6.1 allows for a comparison between the control group (CO) and the 36 PADs, with respect to Age, Level of Education (LoE) and MMSE (Mini Mental-State Examination):

Table 6.1: PADs and CO have similar values for Age and LoE. As expected, they differ on MMSE.

	PAD	**CO**
n°	36	21
Age (sd)	77;6 (7;6)	75;10 (5;2)
LoE (sd)	6.5 (2.9)	7.3 (2.8)
MMSE (sd)	20.1 (3.7)	29.7 (0.45)

[60] Criteria for participants' recruitment and selection are described in Chapter 4.

As illustrated in Chapter 4, MMSE scores provide a first-hand indication of the severity of the disease and the overall cognitive abilities of patients. Four patients score in between 25 and 28 in the MMSE, such that they are classified as mildly impaired; these form the PAD1 group. Twenty-one PADs enter the PAD2 group for moderate impairment with MMSE scores ranging from 19 to 24. Last, eleven PADs show severe impairment as they score in-between 12 and 18: they form the PAD3 group. Unfortunately, it was not possible to balance the number of participants across groups; in particular, the number of subjects who are mildly impaired and therefore fit into the PAD1 group is very narrow. I blame the difficulties at finding subjects with a low level of dementia for the unbalanced distribution of participants. This, in turn, is due to the fact that subjects tend to look for medical support only long after the onset of the first symptoms and therefore receive a diagnosis when the disease is already far beyond the initial mild impairment. Nonetheless, the data collected are informative.

6.2.4 Procedure

Participants complete the task individually, at the presence of the experimenter. Sessions take place either at the research/health care institutions for PADs, or at local associations for controls.

The experimenter illustrates the task material and introduces the protagonist (namely the main female character) to participants as exemplified in (8). After that, the warming-up phase starts with two trials. Depending on the participant's needs, the experimenter encourages participants to perform two to four trials, which are not counted among the results. Once the experiment starts, in case the patient gives signs of being unable to comprehend and perform the task within the first three trials, the administration is interrupted.

Wh-questions are read aloud by the experimenter and repeated upon participants' request for a maximum of two times. However, provided answers are recorded only within the first repetition. A further repetition is allowed in order to avoid frustration in participants, but answers provided after the last repetition are not counted among the results.

Moreover, the experimenter does not provide any feedback to the provided answer. In case participants explicitly asked for one, the experimenter answers in an elusive and positive way.

Task administration takes approximately 5 minutes. No break is planned in between, but one (or more) is allowed, in case participants ask for it or show signs of fatigue.

6.2.5 Coding

For each experimental trial, the participant's answer is coded as (+1) if it corresponds to the target answer. In case the participant does not point to the correct character, the trial is scored with a zero.

Only correct answers provided within the first oral stimulus repetition are coded as (+1). Answers provided after the second repetition are automatically coded as failed (zero), independently of their accuracy.

If patients change their answer or point to one character but say that the other image is also correct, the answer is coded as a mistake (0).

The experimenter assigns scores during the test administration.

6.3 Results

In what follows, I am going to present the results from the 21 controls and the 36 PADs that completed the task. As for the latter group, data are treated according to the level of dementia of each patient. I will now start by providing a group analysis in 6.3.1. An individual analysis will follow in section 6.3.2.

6.3.1 Group analysis

Table 6.2 shows the performance of the participants in the four different experimental conditions. Moreover, it allows for a comparison across the three groups of PADs (PAD1, PAD2 and PAD3) and the control group (CO).

In consideration of the uneven distribution of subjects across groups and therefore of the uneven number of total trials completed by each group, raw scores (number of correct answers) are transformed into percentages, which represent the overall level of accuracy reached by each group on each condition.

Table 6.2: percentages of accuracy (sd) per group and condition.

	CO	PAD1	PAD2	PAD3
n° (MMSE)	21 (29-30)	4 (25-28)	21 (19-24)	11 (12-18)
WhoS	92.85% (11.5)	75.00% (10.2)	61.90% (18.7)	62.50% (12.5)
WhoO	93.45% (10.1)	96.88% (6.2)	57.14% (18.3)	70.45% (21.1)
WhichS	95.23% (7.3)	93.75% (7.2)	77.98% (15.2)	67.05% (19.5)
WhichO	97.61% (5.03)	93.75% (7.2)	70.24% (22.5)	65.91% (25)

A comparison between groups reveals significant differences with respect to COs. PAD1s perform well, so that their performance differs from that of COs only on one condition, namely WhoS (Mann-Whitney, WhoS: U=88.500, Z=2.570, p=.015). PAD2s perform statistically different from controls on all conditions (Mann-Whitney, WhoS: U=391.000, Z=4.848, p=.000; WhoO: U=386.500, Z=4.753, p=.000; WhichS: U=349.000, Z=3.839, p=.000; WhichO: U=385.000, Z=4.845, p=.000). The same pattern is replicated in PAD3s, with significantly reduced comprehension accuracy on all conditions with respect to COs (Mann-Whitney, WhoS: U=218.500, Z=4.294, p=0.000; WhoO: U=194.500, Z=3.331, p=.001; Mann-Whitney, WhichS: U=208.000, Z=3.912, p=.000; WhichO: U=204.000, Z=3.962, p=.000).

Differences in performance do not emerge only in the comparison between COs and PADS. Relevant differences are to be observed also in the comparison among the three sub-groups of PADs. For instance, PAD2s perform differently from PAD1s on WhoO (Mann-Whitney, U=11.000, Z=-2.722, p=.006) and on WhichO (Mann-Whitney,U=12.000, Z=-2.636, p=.007), thus revealing that major changes might take place at that level of dementia with respect to the comprehension of object questions. A comparison between PAD1s and PAD3s reveals that the latter perform poorer with respect to the former on WhichS (Mann-Whitney, U=7.000, Z=-2.375, p=.019) and on

WhichO (Mann-Whitney, U=8.000, Z=-2.268, p=.027), thus signalling a disruption in the comprehension of Wh-questions entailing a complex interrogative phrase.

A second round of analyses compares the accuracy on the four experimental conditions in-between groups. As for COs and PAD1, no significant difference emerged among conditions. The same observation holds true for PAD3 too, although the overall level of accuracy is poorer in this group. This might suggest the idea of a generalized impairment, which gradually reduced comprehension accuracy, independently of the conditions in use. However, data from PAD2 argue against this hypothesis.

A comparison among conditions in the group of patients characterized by moderate impairment (PAD2) detects the following relevant differences on accuracy: the lower performance is registered in correspondence to WhoO, in particular when compared to WhichS (Wilcoxon related samples, Group 2, WhichS vs WhoO: Z=-3.428, p=.001) and WhichO (Wilcoxon related samples, Group 2, WhichO vs WhoO: Z=-2.366, p=.018). Finally, PAD2 participants are significantly less accurate on WhoS than on WhichS (Wilcoxon related samples, Group 2, WhichS vs WhoS: Z=-3.318, p=.001). In sum, PAD2s perform better on WhichNP questions than on Who questions both in the subject and in the object condition.

6.3.2 Individual analysis

The group analysis presented above allows for the individuation of general tendencies within groups, while overlooking individual performances. In what follows, I will now present some data based on an individual analysis in order to deepen the understanding of participants' performance.

Table 6.3 provides an overview of the number of participants in each group who perform above chance-level in the different conditions. Participants complete eight items per condition; accordingly, their performance is classified as 'above-chance' whenever the number of correct answers is equal to six or more.

Table 6.3: Number of participants who perform above chance-level (per group and per condition).

	CO (=21)	PAD1 (=4)	PAD2 (=21)	PAD3 (=11)
WhoS	21	3	8	3
WhoO	21	4	3	6
WhichS	21	4	13	4
WhichO	21	4	13	3

As expected, all participants in the control group perform above chance-level in all conditions. The same observation holds true also for PAD1, except for one participant who has a poor comprehension of WhoS questions. The behaviour has already been detected at the group analysis level, as a comparison between COs and PAD1 actually points out a significant difference between the two groups with respect to the comprehension of WhoS questions.

Also the individual analysis for participants in the PAD2 group confirms data found at the group level. Only three patients (out of twenty-one) perform above chance-level in correspondence to WhoO questions. The specific impairment on the object condition with bare Wh-element is therefore singled out at both levels of analysis. In contrast, thirteen patients perform well on both conditions characterized by a lexically-restricted Wh-element, namely WhichS and WhichO.

Finally, the individual data from patients in the PAD3 group offer interesting insights. Although at the statistical level no significant difference is to be found across conditions in PAD3, the distribution of above-chance performances is uneven. Half of the participants (six out of eleven) performs above chance on WhoO questions, while the same result is not obtained in the other three conditions: three participants perform above chance on WhoS, four perform well on WhichS and three are accurate on WhichO. The image is rather fragmented, with an unequal distribution of above-chance performances across conditions.

The following sections draw further observations from the individual analysis on PAD2 and PAD3 participants.

6.3.2.1 Individual analysis: PAD2

In what follows, contingency tables will now help us to reach a better understanding of the behaviour of participants in the PAD2 group, the only one in which asymmetries

among conditions were detected. In particular, I will take into consideration comparisons between conditions that already emerged as significant at the group level: WhichS vs WhoO (Wilcoxon related samples, Group 2, WhichS vs WhoO: Z=-3.428, p=.001), WhichO vs WhoO (Wilcoxon related samples, Group 2, WhichO vs WhoO: Z=-2.366, p=.018), and WhichS vs WhoS (Wilcoxon related samples, Group 2, WhichS vs WhoS: Z=-3.318, p=.001).

Table 6.4 presents results for the first relevant comparison, namely the one between WhichS and WhoO. The contingency table compares the performance of each patient in the two conditions. In the first row it shows the number of participants who perform well both on WhichS(+) and on WhoO(+), and then the number of participants who perform well on WhichS(+) but not on WhoO(-). In the second row we can see the number of participants that perform well on WhoO(+) but not on WhichS(-) and the number of participants who do not perform well in either of the two conditions (WhichS(-) and WhoO(-)).

Table 6.4: Target comprehension of WhichS and WhoO in PAD2.

	WhoO +	WhoO -
WhichS +	2	11
WhichS -	1	7

Participants mainly distribute across two conditions: eleven patients retain a good understanding of WhichS questions in combination with a poor performance on WhoO; alternatively, seven participants do not reach above-chance comprehension in either of the two conditions. The converse patterns are very rare, as only two participants perform well on both conditions and only one performs well on WhoO but not on WhichS. Altogether, data in table 6.4 confirm that participants perform better on WhichS questions than on WhoO questions.

Table 6.5 shows a very similar pattern in correspondence to the comparison between WhichO and WhoO questions. The specific impairment on WhoO questions is confirmed again by the fact that half of the participants (ten out of twenty-one) perform accurately on WhichO but not on WhoO. Eight patients perform at chance-level in both conditions, while three perform well on both. None of the participants in the PAD2 group has a good comprehension of WhoO in combination with a poor comprehension of

WhichO. Individual data therefore comfirm that PAD2 participants are overall more accurate on WhichO questions than on WhoO questions.

Table 6.5: Target comprehension of WhichO and WhoO in PAD2.

	WhoO +	WhoO -
WhichO +	3	10
WhichO -	0	8

Finally, Table 6.6 offers a comparison within the subject conditions. The group analysis revealed that WhichS questions are statistically better understood than WhoS questions. The claim holds true at the individual level too because the reversed pattern (namely good comprehension of WhoS and poor comprehension of WhichS) emerges only in one participant. However, it should also be pointed out that the other twenty participants evenly distribute across the three remaining possible patterns of performance. Seven patients have an accurate comprehension of both conditions; seven comprehend poorly both, and six have a good comprehension of WhichS questions but not of WhoS questions.

Table 6.6: Target comprehension of WhichS and WhoS in PAD2.

	WhoS +	WhoS -
WhichS +	7	6
WhichS -	1	7

Altogether, data from the contingency tables confirm the observations already drawn at the group level, in particular with respect to the comparison between WhichS and WhoO questions and the one between WhichO and WhoO questions. For what concerns the last comparison, a strong preference for WhichS over WhoS is not completely borne out from the individual analysis because PAD2 participants evenly distribute across different patterns of performance.

6.3.2.2 Individual analysis: PAD3

Table 6.3 shows the number of PAD3 participants who perform above chance in the different conditions: three out of eleven PAD3 participants perform above chance on WhoS questions, six perform above chance on WhoO questions. As for WhichNP

questions, four PAD3 participants perform well on WhichS and three perform well on WhichO.

It is now interesting to observe how above-chance performances distribute among participants. Table 6.7 shows how each participant performs in the four conditions by signalling (✓) whether the number of correct answers is at least equal to six out of eight. MMSE scores for each participant are also provided.

Table 6.7: Distribution of above-chance performances (✓) for each PAD3 participant.

	MMSE	WhoS	WhoO	WhichS	WhichO
PAD3	12	✓	-	-	-
PAD3	12	-	✓	-	-
PAD3	14	-	-	-	-
PAD3	15	-	✓	-	-
PAD3	15	-	-	-	-
PAD3	15	-	-	-	-
PAD3	17	-	✓	✓	✓
PAD3	18	✓	-	-	-
PAD3	18	-	✓	✓	-
PAD3	18	-	✓	✓	✓
PAD3	18	✓	✓	✓	✓

Table 6.7 shows that 3 participants do not perform above chance in any condition. Four patients perform above chance in one condition (either WhoS or WhoO). One participant performs well on two conditions (WhoO and WhichS). Three participants perform well on at least three conditions; these are all characterized by higher MMSE scores (with respect to the defined PAD3 range). Thus, two different profiles emerge from the individual analysis: on one hand, 8 patients perform poorly in the majority of conditions[61]; on the other, three patients master the comprehension of the majority of conditions.

Overall, the PAD3 group is characterized by variability in performance, which does not allow for the observation of relevant asymmetries across conditions at the group level.

6.4 Discussion

The task presented in this Chapter aims at testing the ability of PADs and healthy elderly speakers at comprehending Wh-questions. The data will now be discussed in order to

[61] Only one patient performs above chance in half of the conditions.

provide a first answer to the research questions addressed in Section 6.1: these deal with (*i*) the impairment of the syntactic computation capacities in PADs; (*ii*) the description of how the impairment manifests itself, through a discussion that takes into account previously formulated hypotheses with respect to aphasic speakers; and (*iii*) the individuation of the level of dementia at which major changes take place.

In what follows, I will discuss answers to each research question in turn.

6.4.1 Question (i): Do PADs suffer from syntactic impairment?

The experimental task in use allows the gathering of data that can shed some light into the syntactic competence of patients affected by Alzheimer's disease. The first thing to remark is that the four groups (namely CO, PAD1, PAD2 and PAD3) show different patterns of performance. Accuracy rates drop along with the worsening of the disease. The first two groups enrol either healthy elderly speakers (CO) or PADs with mild impairment (PAD1), who perform very alike, with equivalent rates of accuracy, except for one condition, namely WhoS. Poor comprehension of WhoS questions in PAD1 characterizes in particular one patient (who actually performs at chance-level on the condition). However, the other three participants in the group also do not meet ceiling accuracy, thus resulting in a generalized poorer comprehension of the WhoS condition at the group level (75%).

A comparison between controls and patients either with moderate (PAD2) or severe (PAD3) impairment provides us with more evident signs of impairment in the form of statistically different levels of accuracy in both PAD groups and in all conditions with respect to controls. These data alone reveal a substantial decrease in patients' capacity to comprehend the Wh-questions in use. However, a comparison among the three PAD groups is even more informative, as it reveals asymmetries across conditions. In particular, when compared to PAD1, PAD2 show a specific reduction of comprehension in correspondence to Wh-questions entailing object extraction: they comprehend WhoO and WhichO with a significant lower degree of accuracy. PAD3, on the other hand, stick out for a further worsening of the comprehension of conditions characterized by lexically-restricted Wh-elements (*Quale NP* vs *Chi*).

A within group analysis is also highly informative: it does not register any asymmetry across conditions for CO, PAD1 and PAD3, while it does for PAD2. I will

discuss these asymmetries in the next section. Here, I intend to limit myself to comment on their informational value.

If we consider the asymmetries across groups cited above and the asymmetries across conditions within the PAD2 group, we can conclude that PADs perform differently from Controls from the quantitative point of view and also from the qualitative point of view. Patterns of accuracy do not reflect a mere generalized decrease in comprehension abilities, but also reveal a certain degree of sensitivity to the syntactic factors that characterize the experimental conditions in use. This sensitivity emerges in particular in PAD2 participants. If patients had shown only a gradual decrease of comprehension in all conditions, I could have ascribed the results to a generalized reduction of the cognitive resources that allow for the correct completion of the task. However, this is not the case, as significant asymmetries across conditions emerge. These can be taken as overt manifestations of an impairment at the syntactic level, so that syntactic computation is impaired on specific conditions. The observation opens to further speculations on the precise characteristics of this impairment, which will be the topic of the following section.

6.4.2 Question (ii): How can the impairment in PADs be accounted for?

In order to achieve a better understanding of the syntactic deficit in PADs, I will start by taking into consideration the results that emerge from a comparison across groups. As a second step, I will consider the asymmetries across conditions registered within the PAD2 group and I will get into the details of the syntactic analysis able to account for them. Finally, I will compare results to previously formulated accounts, in order to verify whether any of these can account for the pattern of performance in PADs.

6.4.2.1 Comparison across groups

A comparison across groups reveals that the experimental conditions in use are not equally affected by impoverished comprehension. As presented in section 6.3.1 and further summarized in 6.4.1, the comparison across groups allows for observations on how the impairment develops.

As previously remarked, PAD1s differ from COs on the comprehension of WhoS questions, thus revealing a specific difficulty for PADs in correspondence of this condition. The decrease in comprehension is not dramatic; however, data suffice at

pointing out the emergence of perceivable difficulties in the computation of questions of the type *Chi mi sta fotografando?* ('Who is taking a picture of me? ').

In PAD2 participants the worsening of the disease brings along a decrease in the comprehension of object questions, both in the Who and in the Which condition. In this case the impairment is sharper, as it emerges in most participants. The observation is not surprising as it replicates results from previous studies, which succeeded in pointing out processing asymmetries between structures entailing either subject or object extraction (see Chapter 5 for an overview).

With respect to controls, the performance of the last group (PAD3) differs for a sizable reduction of the comprehension of questions containing the Wh-element *Quale* followed by a lexical restriction, both in the subject and in the object condition.

Based on these data, we can therefore deduce that the syntactic processing competence of PADs reduces along with the worsening of the disease according to a specific pattern of impairment. With respect to the PAD1 group, patients with moderate dementia show a specific poor comprehension of object questions, while patients with severe dementia are further impaired at comprehending WhichNP questions. The ultimate outcome is a considerable variability in performance, with generalized at-chance comprehension in patients with MMSE scores lower than 19 (although some exception to this pattern were also found).

In what follows, I will focus my attention on the performance of PAD2 patients in order to describe how the impairment manifests itself with respect to the different conditions in use. In particular, I will try to point out the underlying mechanism that determines the successful or unsuccessful comprehension of the experimental conditions.

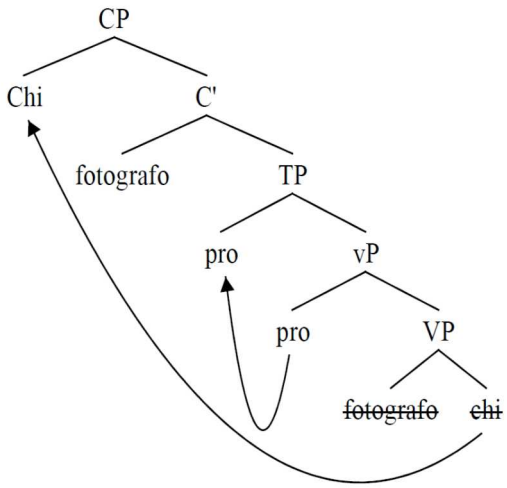
6.4.2.2 PAD2 and their impairment on specific syntactic conditions

As discussed in the previous sections, there is a trend of disruption in PADs, which clearly emerges in PAD2 in the form of a specific impairment on the comprehension of WhoO questions with respect to WhichS and WhichO questions. Moreover, the comprehension of WhoS is rather poor too. In what follows, I will illustrate in turn the syntactic factors that can be held responsible for the observed pattern of performance by analysing the experimental conditions.

The poor comprehension of WhoO questions with respect to WhichS questions was rather expected on the basis of the rich number of studies which previously detected difficulties in correspondence of structures entailing object extraction (see Chapter 5 for an overview). The interesting piece of information in the data concerns the asymmetry registered within the object conditions, namely the poor comprehension of WhoO despite a relatively spared comprehension of WhichO questions. The reasons for this cannot concern only the extraction site from which movement towards the CP-layer starts. I claim that it is precisely the comparison between the features of the two elements to determine whether the computation is successful or not. I recall here the Principle of Relativized Minimality (Rizzi, 1990; Friedmann et al., 2009; see Chapter 5 for more on this) and hold it as responsible for the poor comprehension of WhoO questions.

A representation of the crucial steps of the derivation of WhoO questions can help us to understand how the Principle of RM constrains the structure:

(10) *Chi fotografo?*
Who photograph-1PS



Whenever the Wh-element is extracted from the object position and moved upwards to the CP layer, it has to move across the subject position: this triggers locality effects (Rizzi, 1990). In this configuration the subject occupies an intermediate position between the object and the target position. In other words, the subject intervenes between the two positions which should build a dependency. Based on this principle, all object conditions

should be problematic for speakers. However, the strength of the intervention effect depends on the characteristics of the involved elements. As pointed out by Friedmann et al. (2009) and Belletti et al. (2012), the kind of features that the target and the intervener share can determine whether or not the intervener disrupts the possibility of building the dependency properly. When the two elements share relevant features, the intervener can disrupt the derivation. In contrast, the specification of different features on the two positions can reduce the risk for intervention. In what follows, I consider the feature arrays of the arguments involved in the derivation of WhoO and WhichO questions, in order to explain why the latter is more successfully comprehended than the former.

Table 6.8 presents the sets of relevant features for the arguments involved in the derivation of WhoO questions. In WhoO questions, the subject corresponds to a *pro*, a phonetically null pronoun, while the Wh-element consists of a phonetically realized pronoun. Therefore, both elements are pronominal in nature, although only one is phonetically realized. Both elements lack a lexical restriction ([-NP]), a feature which has been proven to be highly relevant for the mismatch between feature arrays (Friedmann et al., 2009; Bentea et al., 2016). As Table 6.8 specifies, the two positions (target Wh-element in CP and subject) differ in one crucial feature, namely the presence of an interrogative [+Q] on the moved element.

Table 6.8: feature configuration in WhoO questions.

WH-ELEMENT IN CP		SUBJECT	<OBJECT>
Chi	*fotografo*	*pro*	<*chi*> ?
+ Q		- Q	+ Q
+ ProN		+ ProN	+ ProN

The two elements also differ for a second feature, namely person. The subject corresponds to a first person singular (visible on verb agreement), while *Chi* corresponds to a third person singular. The claim for *chi* is based on the fact that the interrogative pronoun triggers agreement with the verb at the third person singular when it is in the subject position (although it is not completely clear whether this can be considered a person feature as normally intended for referents, or should rather be treated as a "non-person"). Thus, in WhoO questions the two elements mismatch with respect to person feature. Now the question is whether person should be included in the set of features relevant for the

computation. As mentioned in Chapter 5, Belletti et al. (2012) assume that only features involved in movement attraction are relevant for feature configurations. Previous studies pointed out that Person occupies a high position in the Hierarchy of *phi*-features (Harley & Ritter, 2002; but see also Benincà & Poletto, 2005; Carminati, 2005). Based on its prominence, I assume that Person is relevant also for the computation of the arguments' feature arrays, at least in normal controls. However, background literature does not provide much information on the role of person feature in Italian sentences of the kind at stake[62]. Moreover, the data collected in the present study do not suffice in clarifying whether or not PADs are sensitive to the person feature and they can insert it in the computation. For all these reasons, I do not include person in the set of relevant features analysed and I assume that WhoO questions have the feature configuration of inclusion presented above (in the terms of Belletti et al., 2012), in the sense that the moved object differs from the potential intervener only for one feature.

As for WhichO questions, I assume the same derivation exemplified in (10). However, the feature configuration of the arguments in WhichO questions is rather different:

Table 6.9: feature configuration in WhichO questions.

WH-ELEMENT IN CP		SUBJECT	<OBJECT>
Quale dottore	*fotografo*	*pro*	<*quale dottore*> ?
+ Q		- Q	+ Q
+ NP		+ ProN	+ NP

As in the case of WhoO questions, the arguments involved in the derivation of WhichO differ with respect to the presence of a [+Q] feature on the moved element. On top of this, one more thing enhances the difference between the two elements, namely lexical restriction. The subject is still represented by a *pro*, namely a pronominal element; while

[62] Belletti & Contemori (2012) present tasks on OR production with Italian-speaking children. In one of the presented studies, the lead-in sentences include either first or third person pronominal subjects, thus eliciting their production in target sentences. The use of pronominal subjects improves ORs production with respect to the production of ORs characterized by lexical subjects, thanks to the mismatch in lexical restriction between subject and object. However, the potential effect of a mismatch in Person feature (i.e., either 1st or 3rd person pronominal subjects for, respectively, mismatch or match with the object head) is not among the goals of the study and authors do not analyse it in depth.
Haendler, Kliegel & Adani (2015) test person feature with 5-year-old German-speaking children and find that thay tend to understand ORs with 1st or 2nd subject pronouns better than ORs with 3rd person subject pronouns.

the interrogative element consists of an interrogative operator followed by a noun phrase, i.e. a lexical restriction. In light of the gathered results, which show a better comprehension of WhichO (70.2%) in comparison to WhoO (57.1%), I must conclude that the mismatch in lexical restriction is crucial for the successful computation of WhichO questions. The mismatch in lexical restriction determines a configuration of disjunction between the two feature arrays. It follows that the subject cannot intervene between the Wh-element and its copy in the VP because it does not share any of the relevant features for the attraction, such that the dependency can be successfully built.

Results also offer the opportunity to observe that the mismatch in phonological realization is not relevant for the participants in the study. If this were the case, the mere difference between an overtly realized pronoun and a null one would have sufficed in ameliorating the comprehension of WhoO questions. This does not seem to be the case in the present study. Nonetheless, the actual role played by *pro* could be further evaluated through a comparison between languages that allow null subjects and others that do not.

The other interesting piece of information revealed by data concerns WhoS questions: patients reach lower than attended levels of accuracy on this condition. In particular, they do not show the attended subject/object asymmetry between the two Who conditions. In contrast, an asymmetry emerges within subject conditions, as participants in the PAD2 group comprehend WhichS questions more accurately than WhoS questions. This piece of data certainly deserves some attention.

At first sight we might assume that there is no movement of the object across the subject and therefore no locality effect can block the derivation, since the latter is extracted from the external argument position and raised to CP, thus landing in a position higher than the one occupied by the object. However, the object in the sentence is the pronoun *mi*, corresponding to an object clitic at the first person singular. The element first raises from its internal argument position to the specifier of the AgrO phrase, the functional projection in use for checking object features. The clitic nature of the element forces a further step though, namely its incorporation into the tensed verb form, so that the clitic raises to C with the verb.

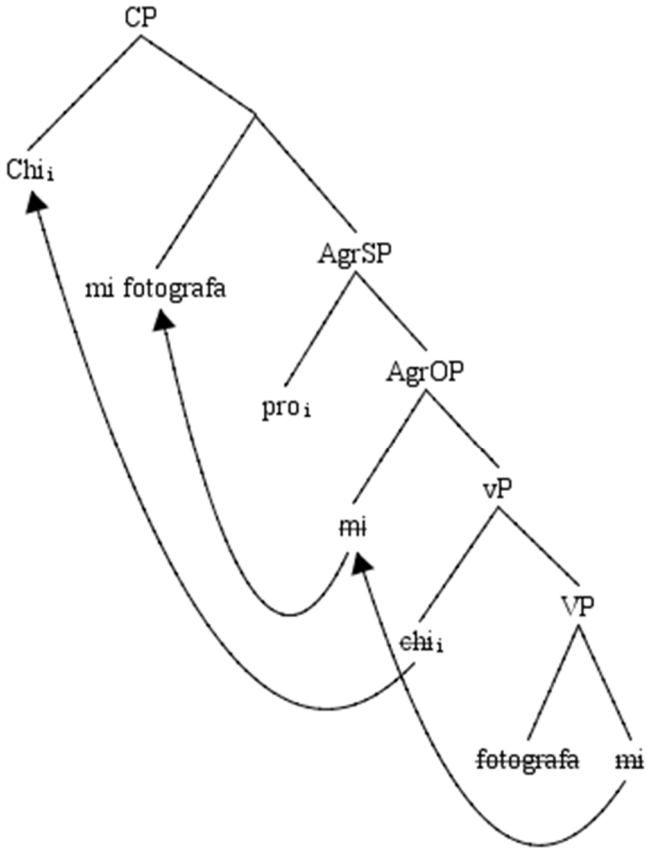
The first movement of the clitic might already be problematic, because it must move across the subject position. Moreover, its intermediate position might posit some further complications. Guasti et al. (2012) and Belletti & Guasti (2015) previously showed that

the derivation of Wh-questions can be disrupted by the object in AgrOP, especially in the case the subject cannot raise to the functional projection responsible for Subject-verb agreement. As discussed in section 5.6.2, Guasti et al. (2012) build on Franck et al. (2006) and Guasti & Rizzi (2002) to point out the specific attraction effect with respect to the processing of Object question. However, a similar configuration can be observed also for WhoS questions of the type in use in the present study.

In the vein of the quoted studies, I represent the relevant steps for the derivation of WhoS questions in (11):

(11) *Chi mi fotografa*?

Who I-ACC.1SG phoptograph-3PS

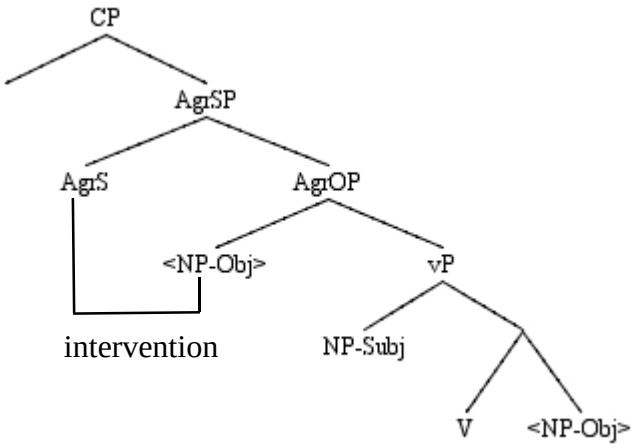


In the vein of Franck et al. (2006) and Guasti et al. (2012), and given the configuration in (11), the presence of the clitic in AgrOP can be problematic under at least three different aspects.

The first one is that *mi* has an effect of disruption with respect to the subject-verb agreement in the sentence structure. The clitic could attract agreement, thus reversing

theta-roles and changing the sentence meaning. The speculation is based on previous studies in Italian and French (Franck et al., 2006; Garraffa et al. 2010). In particular, Franck et al. (2006) found that clitic objects have attraction effects on the verb to the point of changing the agreement, thus altering the sentence meaning. Moreover, Guasti et al. (2012; but see also Belletti & Guasti, 2015) previously built on object attraction in order to account for data on the production of questions in language acquisition. Authors observed that children tend to change verb agreement and ultimately sentence meaning when elicited to produce object questions. As discussed in Chapter 5, Guasti et al. (2012) claim that the mistake is caused by the attraction played by the object in AgrOP on verb agreement by copying its features on AgrSP. The attraction mechanism was already presented in section 5.6.2 and is repeated below (adapted from Belletti & Guasti, 2015):

(12)



Of course, the discussions is only a speculation because the present study only tests comprehension, thus preventing the observation of overt attraction effects (as in the case of production). A further word of caution is necessary because so far studies on attraction only considered number attraction with third person elements. To my knowledge, cases of person attraction (as the one in WhoS would be) have not been discussed yet. In order to verify the hypothesis, further production tests are therefore needed, as these might favour the overt emergence of attraction effects, in the form of verb agreement with the clitic object.

The second hypothesis is that *mi* intervenes in the relation between the attracting feature in CP and the *Chi*, such that the dependency cannot be established. In this case there is no attraction effect on verb agreement; the disruption rather concerns the movement of the Wh-element towards its final position.

The third hypothesis is that *mi* intervenes in the relation that must be established for coreference between the expletive *pro* in SpecAgrSP and the *Chi* in SpecvP. Due to the subject criterion, *Chi* cannot move to the subject position, otherwise it would freeze and it could not move further to the interrogative criterial position. Therefore, an expletive *pro* must occupy the subject position (Rizzi & Shlonsky, 2006). The two elements must establish a relation though. With respect to this, the object clitic in SpecAgrOP might disrupt the probe operation of *pro*.

In all the mentioned hypotheses the feature arrays of the elements involved are relevant in order to determine whether *mi* can represent a potential intervener. Table 6.10[63] presents the feature sets involved in the derivation of WhoS questions. Again, I take into consideration [+Q] and [+NP] as relevant features:

Table 6.10: feature configuration in WhoS questions.

WH-ELEMENT IN CP	OBJECT	
Chi	*mi*	*fotografa?*
+ Q	- Q	
+ProN	+ProN	

Chi and *mi*, the two relevant arguments in the sentence in (11), are both pronominal and have no lexical restriction ([+ProN]). They differ only in one feature, namely [+Q] on the target position and on the moved element. Therefore the two elements are in a relation of inclusion. As observed above with respect to the comparison between WhichO and WhoO, PAD2 participants perform poorly whenever the two elements are in a configuration of inclusion and differ only for one feature.

Finally, one important thing has to be remarked with respect to the preceding discussion on WhoS questions. All the cited cases of potential intervention of *mi* concern

[63] For reason of simplicity, the feature array of the potential intervener is represented under the proclitic pronoun, which corresponds to the final position the element occupies after linearization. However, as represented in (11) and further illustrated in the discussion, I assume that intervention effects can be triggered also by the clitic copy in the lower position.

the relations that *Chi* (the subject) must establish with other elements in the sentence (either AgrSP or the attracting feature in CP). Apart for the specific feature configuration in Table 6.10, I claim that disruption might also be favoured by the element *Chi*. The interrogative pronoun represents a variable and is not assigned any specific feature with respect to person and number (although it agrees in third person singular with the verb) because it has no specific referent. Therefore, PADs might experience difficulties in realizing the agreement relation between *Chi* and the verb, independently of the presence of an intervening object.

In contrast, *QualeNP* provides the parser with clear features for subject-verb agreement. Moreover, a mismatch in lexical restriction crucially changes the feature configuration of the subject questions in use. This can be observed in WhichS questions:

Table 6.11: feature configuration in WhichS questions.

WH-ELEMENT IN CP	OBJECT	
Quale dottore	*mi*	*fotografa?*
+ Q	- Q	
+ NP	+ ProN	

In WhichS questions, the two arguments differ in two relevant features: the interrogative [+Q] and the lexical restriction [+NP]. *Quale dottore* bears both features because it is interrogative and the Wh-element entails a lexical restriction. In contrast, the object is pronominal in nature and has no lexical restriction; plus, it does not bear any feature of the scope-discourse type. The two elements are therefore in a configuration of disjunction in the sense defined by Friedmann et al. (2009). Disjunction hinders the possibility for *mi* of acting as an intervener, as *mi* does not share any of the relevant features responsible for attracting the extracted argument to CP. It follows that the dependency between *QualeNP* and *pro*, and QualeNP and the attractor in CP, can be successfully built.

Altogether, results from the subject conditions are consistent with results from the object conditions and confirm the crucial role of the mismatch in relevant features in the computation of non-local relations. In particular, data on the conditions in use attest the relevance of lexical restriction in feature arrays. As exemplified, the Principle of Relativized Minimality can be implemented in different steps of the derivation and in different sentences: for instance, subjects can intervene with respect to the dependency

between the attractor in CP and the object endowed with a scope-discourse feature in VP (as in the case of WhoO questions), or an object can intervene in Subject-Verb agreement (as hypothesized for WhoS questions). In both cases, the computation of the feature arrays of the arguments is crucial for determining the strength of the intervention. In particular, it has to be determined which kinds of configurations can be computed by the grammar at stake (children grammar, adult grammar, patient grammar, etc.). The feature configuration of the conditions in use in the present study are summarized in (13):

(13) a. WhoS

Chi *mi* *sta* *fotografando?*

+Q, +ProN +ProN INCLUSION

who I-ACC.1SG AUX-3PS photograph-PROG

b. WhoO

Chi pro *sto* *fotografando?*

+Q, +ProN +ProN INCLUSION

Who pro AUX-1PS photograph-PROG

c. WhichS

Quale dottore *mi* *sta* *fotografando?*

+Q, +NP +ProN DISJUNCTION

Which doctor I-ACC.1SG AUX-3PS photograph-PROG

d. WhichO

Quale dottore pro *sto* *fotografando?*

+Q, +NP +ProN DISJUNCTION

Which doctor AUX-1PS photograph-PROG

In light of the comparisons presented above, I claim that patients perform well on disjunction configurations, while they are impaired in inclusion configurations. Whenever the moved argument and the potential intervener differ only for the presence of a scope-discourse feature ([+Q] in the present study), they are in a configuration of feature inclusion, which triggers intervention effects. In contrast, when the potential intervener and the moved argument differ in their relevant features, such that their feature arrays are in a configuration of disjunction, no intervention effect emerges and sentence processing is successful.

6.4.2.3 Alternative accounts

In the previous section I considered the Principle of Relativized Minimality in order to account for the performance of PADs. Since the seminal works of Friedmann, Belletti & Rizzi (2009) and Garraffa & Grillo (2008), speakers' difficulties in the processing of specific syntactic structures entailing extraction have been accounted for in terms of locality effects triggered by the presence of relevant features both on the intervener and on the target position.

However, previous studies proposed a number of other accounts, which I reviewed in Chapter 5. I will now turn back to the previously illustrated proposals and point out the reasons why they cannot account for the performance of patients in the present study.

The good performance of PAD1 and PAD2 on WhichO questions allows discarding two hypotheses, namely the Minimal Chain Principle (De Vincenzi, 1991a and subsequent work) and the Trace Deletion Hypothesis (Grodzinsky, 1986, 1990). Both accounts predict a poor comprehension on sentences entailing object fronting, even though on the basis of different premises: the former assumes that the parser aims at building the shortest possible chain and therefore assigns a subject reading as a default strategy for the processing of moved elements. The latter ascribes difficulties to the impossibility of retrieving traces located deep in the structure. In the present study, both WhichO and WhoO questions rely on object extraction, such that the position of the trace with respect to the filler is identical for the two. According to the hypotheses, speakers should perform the same in the two conditions, as far as the syntactic structures in use are concerned. Along these lines, the two hypotheses do not predict the results I found, with a better comprehension of WhichO with respect to WhoO questions. Therefore, the two

mentioned accounts should be excluded from the interpretation of the data collected in the present study[64].

The high level of accuracy of PAD2s on questions with *quale* 'which' (both for the subject and the object condition) disconfirms the hypothesis of an intrinsic complexity brought up by *Which*-elements and their lexical restriction in reconstruction. Avrutin (2000) previously claimed that WhichNP questions are intrinsically difficult because they require the integration of detailed semantic information in the sentence computation. In the author's view, items of the Who type entail less semantic information and therefore are easier to integrate in the sentence. Given the good comprehension of questions introduced by Which-elements, data from participants in the present study are not compatible with Avrutin's hypothesis. The good results on WhichO questions show that complexity is not intrinsic to lexically-restricted items. It is rather determined at the structural level, depending on the features of the elements involved. In other words, the complexity of an element is determined by the comparison with the other elements in the sentence and the kind of configuration they enter. This holds true for the present test design, where out-of-the-blue sentences do not posit specific constraints at the semantic and pragmatic level. Different results might be obtained in discourse contexts; where elements might trigger *a priori* facilitating effect or difficulties at some levels.

In any case, the results from the present study are even more interesting if we think of the fact that PADs notoriously suffer from anomia (Chertkow & Bub, 1990, Hodges et al., 1991, Almor et al., 2009; among many others). Within their pattern of impairment, a specific difficulty in phrases entailing lexical restriction could be attended. As shown above, this is not the case. Results strengthen the hypothesis that the lexical restriction actually improves comprehension under the appropriate condition, namely in a configuration of mismatch in lexical restriction, as described above.

On the basis of this observation (founded in turn on the Principle of Relativized Minimality), I predict that PADs should show the reversed pattern of performance (i.e., a better comprehension of Who questions over Which questions) at the presence of full lexical DPs as second argument in the sentence. For instance, in the case of object questions entailing a fully-fledged lexical subject, patients should comprehend WhoO

[64] However, it was brought to my attention that offline comprehension tasks are not the most suitable technique for the evaluation of the two mentioned hypotheses. The two proposals focus on temporal aspects of processing (e.g., error reparation and trace activation), which can be better investigated in online tasks.

questions with a higher accuracy with respect to WhichO questions, because the former would allow for a clear mismatch in features with respect to a lexical subject, while the latter would not. The hypothesis is compatible with results from previous studies. For instance, De Vincenzi et al. (1999) and Guasti et al. (2012) found a better accuracy on Who object questions than on WhichNP object questions (in Italian-speaking children). Crucially, the experimental sentences in use contained a lexically restricted subject, such that mismatch with Who elements followed.

Finally, the fact that PAD1 and PAD2 participants have a good comprehension of all (or some of) the experimental conditions speaks against the Tree Pruning Hypothesis put forth by Friedmann (2002) for aphasic speakers. In the author's view, aphasic speakers cannot project any constituent higher than TP, thus failing at processing any sentence structure entailing movement to CP. In Italian, all root Wh-questions require the movement of a Wh-phrase to the upper portion of the sentence structure, independently of the position from which the element is initially extracted. The proposal is therefore incompatible with the patterns of performance of patients with mild or moderate impairment. In contrast, the proposal is compatible with the pattern of performance of PAD3 participants, who have only at-chance comprehension of all the experimental conditions in use (except for the the three patients who have a high performance in most conditions). Based on the data from the present study, it is therefore possible to hypothesize that the disease causes a reduction of the ability to move elements to the CP layer at its severe stage. Yet, further research is needed in order to verify whether the hypothesis holds true. In particular, it will be necessary to verify whether patients can compute matrix clauses that do not entail any movement to the CP. This will be possible through the use of matrix clauses in the experimental design in Chapter 7.

6.4.3 Question (iii): At which stage of the disease are PADs affected by syntactic impairment?

As previously illustrated, participants in the task subdivided into three groups according to their level of dementia, namely PAD1, PAD2 and PAD3. PAD1 consists of patients with mild impairment: they perform with high levels of accuracy in all conditions, except for a slightly lower amount of correct answers in correspondence of WhoS questions. PAD2 is the more interesting group for the discussion: it includes patients with moderate

dementia, whose comprehension accuracy sharply varies across conditions. Finally, patients in the PAD3 group are characterized by severe dementia and variability in their performance. Overall, the last two groups (PAD2 and PAD3) perform differently from their healthy peers in all conditions.

Based on the results, I can therefore conclude that major signs of disruption at the syntactic processing level emerge in patients affected by moderate impairment. At this stage of the disease, reduced computation abilities can still cope with some of the sentence structures in use in the experiment, while they are inadequate for others, depending on their specific syntactic characteristics. The deficit can be described along the lines of Relativized Minimality. As described in the previous section, the processing of crossing movements between arguments whose feature sets are in a configuration of inclusion is difficult for PADs. In contrast, moderately impaired PADs are facilitated at processing sentences that entail argument extractions whenever the feature sets of the elements involved are in a configuration of disjunction.

The severe stage of the disease (PAD3) is characterized by a generalized decrease in question comprehension. The group analysis reveals that levels of accuracy are clearly different from those of controls for all conditions. In addition, variability emerges from the individual analysis. As a result, clear asymmetries across conditions cannot be detected anymore.

Finally, the answer to the question might not be as straightforward as it looks like and a word of caution is necessary. It is well-known that the symptoms of the disease tend to appear clearly only long after the onset of neural damages (Braak & Del Tredici, 2006). It is then more appropriate to say that the impairment becomes evident when dementia becomes moderate, although the disruption most probably starts earlier.

6.5 Conclusions

In the present chapter I discussed results from a study on the comprehension of different kinds of Wh-questions by patients affected by Alzheimer's disease. The aim of the study was to sample patients' ability in the computation of extraction of wh-items. In order to do it, I used four different experimental sentences, which were obtained through the manipulation of two factors, namely: extraction site (subject vs object questions) and the

Wh-element in use (Who vs WhichNP questions). A further relevant factor for data analysis in the study is the level of dementia of participants.

The first result of the study is that patients complete the task according to different patterns of performance, which in turn depend on their level of dementia.

A mild level of dementia corresponds to a pattern of performance that closely resembles the one of healthy controls. In contrast, patients with moderate and severe impairment perform differently from controls in all conditions.

Moderate impairment brings along lower levels of comprehension accuracy; and extraction is spared only under specific syntactic conditions. These can be explained along the lines of Relativized Minimality (cf. Friedmann et al., 2009; Rizzi, 1990). Patients apply the principle in a strict version and allow the computation of non-local relations only in the case the feature arrays of the arguments are sharply different. According to the conditions in use, a mismatch in lexical restriction is a crucial factor for the successful completion of the sentence computation.

Finally, the majority of patients with severe impairment does not show signs of sensitivity to syntactic manipulation and performs around chance level in most conditions. At the group level, no significant asymmetry across conditions is to be detected, while the individual analysis reveals a certain degree of variability within the group.

6.6 Open questions

In the previous sections I interpreted the data collected in order to provide a first answer to the research questions addressed in this study. However, data also rose some further questions.

As for the general design of the study, the present task only considers extraction in root questions. In order to corroborate the significance of the results, it would be useful to replicate them with other forms of extraction, in particular with cases of extraction from embedded clauses. For this reason, the following chapter will focus on the computation of relative clauses.

The task in use brought up the relevance of the mismatch in lexical restriction between arguments in order to hinder intervention and to favour the computation of crossing movements. However, all Wh-questions in use in this task entail a pronominal

potential intervener and either a pronominal or a lexically restricted moved element. In the next section I would like to reverse the pattern by introducing relative clauses characterized by a lexically restricted intervener and the movement of a functional element with no lexical restriction. I predict that patients will benefit from a mismatch in lexical restriction, independently of which of the two involved elements is lexically restricted (either the intervener or the target position). The issue is therefore further addressed in Chapter 7.

In the next task, I will also try to replicate the results with respect to the different patterns of performance in correspondence to different levels of impairment. Moreover, I would like to verify whether the generalized decrease in comprehension in PAD2 and PAD3 participants only concerns sentences entailing argument extraction or not. The introduction of matrix declarative clauses among the experimental conditions will help answer the question.

The open questions will be further discussed in Chapter 7.

7 RELATIVE CLAUSE COMPREHENSION IN PATIENTS WITH ALZHEIMER'S DISEASE

7.0 Introduction

In Chapter 7 I present experimental data from a test on the comprehension of relative clauses in patients with Alzheimer's disease.

The reasons to test PADs on RCs are numerous. The main goal of the present study is to verify whether patients suffer from a syntactic deficit; and RCs represent the ideal field for the purpose for at least three different reasons. As discussed in Chapter 5, RCs entail a rather complex derivation from the syntactic point of view. Their computation requires the completion of steps that matrix clauses lack; among these, an instance of Wh-movement: this implies that a gap must be interpreted through reconstruction of the moved NP in its original position. After testing extraction in root clauses (see the study on Wh-question comprehension in Chapter 6), the attention shifts now to extraction in embedded clauses, such that the two phenomena (extraction and embedding) can be observed jointly. Moreover, the nature of RCs is *per se* atypical, given that they enter the derivation as part of a noun phrase, either in the complement position or in a higher functional projection of the DP (depending on whether Kayne's (1994) or Cinque's (2008, 2014) analysis is assumed).

The syntactic complexity leads directly to the second reason why RCs are the suitable field for the sample of syntactic impairment. Their structure offers the possibility to manipulate a variety of factors in order to test speakers' sensitivity to different features. The manipulation of factors thus allows for the creation of different experimental conditions entailing variable degrees of complexity.

The third reason to introduce RCs among the stimuli is a very wide literature on the topic, which allows for a valuable comparison of the data collected in this study to those from other works. By building on previous research, it will be possible to compare the performance of PADs to those of typically developing young and older children and adults, even in a cross-linguistic perspective.

Concerning this last point, I would also like to underline the fact that by collecting data on the processing of RCs in healthy and impaired elderly adults, the present study contributes to the understanding of how language evolves in aging speakers. Most of previous research on Wh-movement addressed issues concerning L1 acquisition and observed adults in the role of controls. The present study covers the further steps taken by adult speakers, either in a healthy or in an impaired status. As a result, the full lifespan is covered with respect to the ability to process RCs. Although the issue will not be directly addressed in the present work, which rather focuses on Alzheimer's disease, data from the performance of healthy elderly controls will also be presented.

The present chapter is organized as follow: research questions are addressed in 7.1. Section 7.2 describes the task in use, the material, the participants, the procedure and the guidelines for coding. In 7.3 I present the collected data, which are discussed in 7.4. Section 7.5 summarizes temptative answers to each experimental question.

7.1 Research questions

The task in use will help answering the research questions addressed in the Introduction. In addition, a further question is pointed out (*ii*), this concerns patients' sensitivity to syntactic manipulations. Overall, the following research questions will be considered:

i. Do PADs suffer from syntactic impairment?
ii. What kind of syntactic manipulations are PADs sensitive to?
iii. How does the syntactic deficit manifest itself?
iv. At which stage of the disease are PADs concerned by syntactic impairment?

The reason for running a second study on the same questions is due to the intention to verify whether results can be replicated in structures that have similarities and differences with respect to Wh-questions. The conclusions reached in Chapter 6 would gain more strength thanks to similar results on Relative Clauses.

The task will also allow us to observe patients' sensitivity to different kinds of syntactic features. The topic has been addressed throughout in Chapter 5: young speakers (and adults) benefit from specific feature configurations in the computation of extraction movements. The present study takes into consideration some of those factors in order to

allow for a comparison between PADs, on the one hand, and children and adults on the other.

Moreover, the task in use will offer the opportunity to answer some of the questions left open in Chapter 6, in which it was not possible to verify whether PADs only have difficulties at extracting phrases to be moved to the CP-layer, or their impairment also disrupts the comprehension of structures that do not involve A'-movement. The use of declarative sentences among the experimental conditions in the present study will provide an answer to that question. In particular, I intend to observe whether the comprehension deficit of PAD3 patients also affects simple declarative sentences.

7.2 The task

The task in use is a sentence-to-picture matching task that aims at sampling the comprehension of different kinds of relative clauses. The visual material and the general design are taken from BAMBI, a well-known task developed by Friedmann & Novogrodsky (2002), and subsequently employed in many languages, among which Italian (Contemori & Belletti, 2014).

7.2.1 Design

Each experimental stimulus in the task is composed of two elements: a visual and an oral one.

The visual part consists of two coloured drawings printed on the same A4-sheet. In both drawings, two characters are involved in the same action (e.g., combing, taking a picture, kissing, etc.). The two images differ in the character that performs the action; for instance, in the first picture character A kisses character B, while in the second picture character B kisses character A.

All verbs used in the stimuli are transitive and refer to a concrete action that can easily be acted out and represented in a drawing.

The oral content of the experimental stimuli consists of a sentence read aloud by the experimenter. Participants' task is to listen to the sentence and point to the character that correctly represents the sentence meaning. Only one of the two images corresponds to the target, while the second one corresponds to the reversed interpretation of the sentence.

The characters depicted in the drawings are all animate and human, except for the presence of a dog in one pair of images. The original task design actually includes a variety of animals (e.g. penguins, rabbits, giraffes, etc.), but these were all eliminated from the present version of the task because they might be not familiar animals for PADs. Such animals could have represented a confounding factor for patients: the risk is to collect data mirroring an impairment on the lexicon of exotic animals, rather than on syntactic competence.

The complete task includes 48 sentences, subdivided into six different experimental conditions, including eight trials each. Trials are distributed across twelve pairs of images. Each pair of images is used four times, in association to four different experimental conditions.

Trials are randomized according to the sentence type, so that more than two trials from the same condition are never performed in a row. Moreover, target answers are equally distributed across the two characters and the two images (the upper and the lower one, according to the distribution on the paper sheet).

Short attention span in PADs and the desire not to burden participants with excessive demands are the main reasons for not including fillers in the task.

7.2.2 Material

As anticipated, the task includes six different conditions: a declarative sentence, a subject relative clause and four different kinds of object relatives.

Examples of the conditions in use are in (1). For the sake of clarity, all examples presented here assign the agent *theta*-role to the same character (the granny). However, experimental trials were designed in order to equally assign the agent *theta*-role to the two characters across conditions and to distribute the answers across the two pairs of images.

(1)

a. SVO *La nonna bacia la bambina*
The granny.SG kiss-PRS.3SG the child.SG
'The granny kisses the child'

b. SR *Mostrami la nonna che bacia la bambina*
Show-IMP.2SG I-OBL.1SG the granny.SG Rel kiss-PRS.3SG the child.SG
'Show me the granny that kisses the child'

c. OR *Mostrami la bambina che la nonna bacia*
Show-IMP.2SG I-OBL.1SG the child.SG Rel the granny.SG kiss-PRS.3SG
'Show me the child that the granny kisses'

d. POR *Mostrami la bambina che è baciata dalla nonna*
Show-IMP.2SG I-OBL.1SG the child.SG Rel AUX kiss-PTCP by.the granny.SG
'Show me the child that is kissed by the granny'

e. ClOR *Mostrami la bambina che la nonna la bacia*
Show-IMP.2SG I-OBL.1SG the child.SG Rel the granny.SG she-ACC kiss-PRS.3SG
'Show me the child that the granny kisses *her*'

f. ORdem *Mostrami quella che la nonna bacia*
Show-IMP.2SG I-OBL.1SG that-F Rel the granny.SG kiss-PRS.3SG
'Show me the one that the granny kisses'

The first condition to be introduced (1a) is a declarative sentence with the unmarked Subject-Verb-Object constituent order. Its function is to play the role of the baseline condition with respect to all the other conditions, which include a matrix and a relative clause. It represents the benchmark for the evaluation of all other sentence types. Moreover, as previously pointed out, the performance on matrix clauses can be particularly informative about the status of syntactic computation in severely impaired speakers. Patients with severe impairment have around-chance comprehension of Wh-questions. Data on matrix clauses will tell us whether patients are impaired only at A'-movement to the CP-level, or they are generally impaired at carrying out the task. The first scenario would correspond to high levels of accuracy in the comprehension of simple declarative sentences, while in the second all conditions will be equally impaired.

Relative clauses in conditions (1b) to (1f) are introduced by a matrix clause; this is composed of a predicate at the imperative form (*Mostrami* 'Show me') plus the head of the relative. The reasons for choosing this kind of introductory sentence is twofold: on one hand, it reminds participants of how they are expected to perform in order to fulfill the task's requirements (namely by pointing to target character). This is particularly useful given that PADs have short attention spans and often suffer from apraxia, which inhibit them from completing the procedure they are executing. On the other hand, the matrix clause in use does not contribute to the semantic interpretation of the sentence. This means that participants do not need to activate further lexical items and to process extra semantic information in order to complete the task, and therefore they can allocate all their efforts to the comprehension of RCs.

Conditions (1b) and (1c) represent a minimum pair: a subject relative clause (*Mostrami la nonna che bacia la bambina*, 'Show me the granny that kisses the child'), and an object relative clause (*Mostrami la bambina che la nonna bacia*, 'Show me the child that the granny kisses'). The two sentences include two lexically restricted DPs as subject and object. The two conditions only differ in the site in which the relative head noun is initially merged in the RC: as for the transitive verbs in use, either the external argument in SRs, or the internal one in ORs[65]. In the light of the previously reviewed literature, I expect this pair of conditions to be highly informative about participants' syntactic processing abilities. Indeed, in name of the well-known asymmetry between subject and object extraction (see Chapter 5 for an overview), if patients suffer from any kind of impairment, signs of it are expected to emerge at least in this minimal pair, i.e. in the form of an enhanced asymmetry between the two conditions with respect to the levels of accurate comprehension.

The remaining three conditions (1d to 1f) are all instances of ORs, which result from different kinds of syntactic and feature manipulations. The condition in (1d) uses a passive voice to convey the meaning of an OR through a SR (*Mostrami la bambina che è baciata dalla nonna*, 'Show me the child that is kissed by the granny'). The structure, known in the literature as Passive Object Relative (POR) since Belletti (2014), has given

[65] Although the issue is not addressed in the discussion, it should be noted that subject and object extraction follow slightly different procedures. As discussed by Rizzi & Shlonsky (2007), subject extraction posits the necessity to fill the subject criterial position with expletives, in order for the subject DP to skip the criterial position and move to the CP layer. Being Italian a null subject language, *pro* is the expletive in the criterial position.

proof of ameliorating the processing of ORs, thanks to the *smuggling* operation it entails[66], which hinders the intervention of the subject. Looking at data from Italian (Belletti & Contemori, 2009; Contemori & Belletti, 2014), PORs emerge systematically in adult production, they appear in child production starting from age 5; and they also facilitate child comprehension of ORs. We can assume that children (starting from a certain age on) and adults behave alike with respect to this condition, which, in the present study, assumes a crucial role in evaluating whether PADs perform qualitatively differently from healthy speakers.

Condition (1e) corresponds to a kind of OR that is attested only in substandard varieties of Italian (Guasti & Cardinaletti, 2003; Contemori & Belletti, 2014): its peculiarity consists in entailing a resumptive clitic of the moved object (*Mostrami la bambina che la nonna la bacia*, 'Show me the child that the granny kisses her'). From the syntactic point of view, two alternative analyses of the structure have been proposed. According to the first one, ClORs are derived exactly as ORs are, with the only difference that the moved Object is a big DP[67] which strands a clitic in the relative clause (Belletti 2005, 2006; Boeckx, 2003; see also Kayne (1975) and Uriagereka (1995) on clitics). Alternatively, no Wh-movement is at work and the head noun (externally generated) is linked to the clitic through a semantic bind (McCloeskey 2001, 2002). Leaving aside the syntactic dispute, studies on Italian acquisition already addressed the issue because it was noticed that children often use resumption when ORs are elicited (Guasti & Cardinaletti, 2003; Contemori & Belletti, 2014). In Contemori & Belletti (2014) resumption can take two forms: either a full DP is repeated within the RC or an object clitic is realized in preverbal position, and children produce them at the same rate. As for comprehension, authors find a facilitating effect given by resumptive clitics in ORs in comparison to ORs with gap. The effect is significant at least in older children (aged 6;5 – 8;10), although comprehension of ORs with resumption does not improve as much as in the case of PORs. For instance, in older children comprehension of OR is at 64%, comprehension of OR with resumption is at 77%, and comprehension of POR is between 92% and 95%, depending on the kind of passive in use (Contemori & Belletti, 2014: 1042). In the present study ClORs can be useful to verify whether the resumptive clitic helps participants to

[66] See Section 5.5.3.6 for details on how passive derivation through *smuggling* can improve the comprehension of ORs.
[67] See also Poletto (2006), where clitic doubling is accounted for without resorting to a big DP.

reconstruct the interpretation of the relative head within the RC and to draw a comparison between the magnitude of the effect in Italian-speaking children and in PADs.

ORdem, the last condition in the experiment, builds on the Relativized Minimality account for RCs and aims at verifying whether patients benefit from enhanced asymmetries between the feature arrays that characterize the subject and the object. In particular, in the proposed condition (*Mostrami quella che la nonna bacia*, 'Show me the one that the granny kisses'), the moved object consists of a light noun and the head is spelled out by a demonstrative pronoun. Building on Benincà (2010, 2012), Sanfelici & Poletto (in press) analyze ORdem[68] as entailing a light noun of the [PERSON] or [THING] type, which raises high in the spine of the relative clause, but it is crucially not extracted from it. The relevant steps for the derivation of ORdem are represented in (2):

(2) $[_{DP}$ [quella] D° $[_{CP}$[PERSON]$[_{TP}$ DP_i V_t [vP i t <[PERSON]>]]]

The subject and the object of the RC differ in at least two important features: the subject is lexically restricted but lacks the [+Rel] feature; the object is not lexically restricted and is endowed with [+Rel]. Altogether, the subject and the object in the ORdem condition have feature arrays which are in a disjunction relation, in the sense that they differ both for the lexical restriction and for the [+Rel] feature. Data will tell us whether this kind of feature configuration can help PADs to improve their accuracy in OR comprehension or not.

7.2.3 More on the experimental conditions

Preverbal subjects, independently of the syntactic manipulation in use, characterize all ORs inserted in the list of experimental stimuli. Some Italian native speakers judge ORs with a preverbal subject of the type in (3) less natural than ORs with a post-verbal subject of the type in (4). It could be claimed that poor performance on ORs registered in previous studies (for an overview see Contemori & Belletti, 2014) is due to the fact that experimenters did not use the most natural form of ORs and that the participants were tested on confounding material.

[68] ORdem is a label used in the present study in order to refer to the structure in (1f). Sanfelici & Poletto (in press) simply refer to it as restrictive relative.

(3) *Mostrami la mamma che la bambina bacia*
Show-me the mother that the child kisses
'Show me the mother that the child kisses'

(4) *Mostrami la mamma che bacia la bambina*
Show-me the mother that kisses the child
'Show me the mother that the child kisses'

However, different reasons support the choice of introducing only ORs of the type in (3). First, ORs with post-verbal subjects are ambiguous between a SR and an OR reading. The ambiguity could be solved by using a mismatch in number between the two DPs; in that case, verb agreement would provide for efficient disambiguating cues towards the DP that should be interpreted as the sentence subject. This solution is not desirable though in the present study (see below about the reason for not including number mismatch as a factor in the study).

Moreover, Belletti et al. (2012) point out that on the same vein, SRs with post-verbal object, could result ambiguous and be interpreted as ORs, while this is never the case. The interpretation of (4) as a SR consistently emerges across studies and no control group ever showed signs of uncertainty between the two interpretations (SR/OR). A third observation concerns post-verbal subjects: this is not their usual position in standard Italian, at least not for transitive verbs like the ones in use in the present study. It has been widely proved that post-verbal subjects of transitive verbs play a specific role at the pragmatic level: they express Focus of New Information (Belletti, 2004; Belletti & Leonini, 2004; Belletti et al. 2012). In the task in use, such an interpretation of OR subjects is not available because both characters appear simultaneously in the pictures.

Further support comes from Guasti, Stavrakaki & Arosio (2008), a study that provides results from a direct comparison between ORs with pre-verbal subjects and ORs with post-verbal subjects. The comparison is allowed by the fact that their stimuli are disambiguated through verb agreement in presence of two arguments characterized by number mismatch (e.g., singular subject and plural object or *viceversa*). The results from the study show that both in Italian and in Greek, the two languages at stake in the study,

post-verbal subjects in ORs bring along lower levels of accuracy in a sentence-to-picture matching task in comparison to ORs with pre-verbal subjects.

For all these reasons taken together, I claim that the use of preverbal subjects in ORs is more appropriate in the present task with respect to post-verbal subjects. The fact that some native speakers judge the structure not natural should be interpreted as a further sign that RM makes the derivation difficult even for adult speakers, although it is grammatical.

So far, I have illustrated the conditions included in the study and their major characteristics. I would now like to spend a few words on why I decided to exclude another relevant factor from the present study. Among the features that have been proved to be responsible for reducing processing complexity in ORs, I deliberately avoided employing number mismatch as a factor for manipulation in the present task.

The reason for not including number mismatch is the following: in comprehension studies based on a sentence-to-picture matching task, number mismatch posits constraints on the number of images that have to be shown in accordance with the oral stimulus. In particular, in order to represent the complete range of possible interpretations at least four different images must be reproduced (Adani et al., 2010), namely: one in which character A is the agent and is singular (*a grandmother kisses two girls*) and one in which it is plural (*two grandmothers kiss a girl*), and additionally two images in which the agent *theta*-role is assumed by character B, first as a singular (*a girl kisses two grandmothers*) and then as a plural (*two girls kiss a grandmother*). In other words, three different *foils* should accompany every target image. This kind of design is certainly possible and, indeed, it was already employed with children with good results (Adani et al., 2010; Adani, 2012). However, a design that entails the analysis of four different pictures at the same time is not desirable with PADs, because it could represent a confounding factor. This is due to the fact that the status of visual perception in PADs has not been completely defined yet (Hodges et al., 1991). Therefore, I decided to leave aside number mismatch and exclude it from the pool of tested factors, in order to avoid the risk that a higher amount of visual information might confound data, which would result in being impoverished by the overload of stimuli at the visual-perceptive level rather than by syntactic complexity.

Two further factors from the list of those reviewed in Section 5.5.3 are missing in the present study; these are case and gender. Concerning the former, its omission is rather

straightforward, given that Italian lexical DPs do not have overt morphological markers for case. As for grammatical gender, although it is marked on DPs, its manipulation in Italian stimuli did not allowed for improved comprehension, at least not in children (Belletti et al., 2012). Belletti et al. (2012) claim that the feature is not relevant in the computation of RM because it does not play a role in triggering movement within the syntactic derivation of Italian structures, while it certainly does in other languages (Hebrew for instance). This being a parametric difference among languages, I have no reasons to expect a different result from PADs in comparison to children, and therefore I decided not to test gender mismatch in the present study.

7.2.4 Participants

The same participants as in Experiment 1 on Grammatical Gender retrieval were enrolled also for the present task on RC comprehension. This holds true both for patients and for their healthy peers. Criteria of enrolment and data with respect to age, level of education and dementia are reported in Section 4.2.2.

Results are presented for the three groups in which PADs have been subdivided according to their level of dementia. I am going to refer to them again with the labels 'PAD1', 'PAD2' and 'PAD3'. The group PAD1 (mild impairment) includes four subjects with MMSE scores between 28 and 25. The most copious group is the one of moderately impaired patients, PAD2, with MMSE scores between 24 and 19. Finally, the remaining 12 subjects are severely impaired: they score between 18 and 12, and form the PAD3 group. Unfortunately, it was not possible to balance the number of participants across groups, because of the reason discussed in Chapter 6.

7.2.5 Procedure

The task was administered individually, either at the hospital in the case of patients, or at local associations, in the case of controls. Sessions took place in a private and quite room.

A warming-up phase including four trials preceded each experimental session, in order to give participants the possibility to familiarize with the task. Right afterwards, the researcher started showing the experimental stimuli.

The experimental stimuli were presented in row and no break was planned, but patients were allowed to have a short break whenever they asked for it, showed fatigue

or gave signs of behavioural disorders. After the break, patients were brought back to the task by repeating the four sentences from the warming-up phase, before continuing with the experimental material.

No feedback to answers was provided during the task. In case patients did not answer or could not hear the sentence stimulus, this was repeated at most twice, although correct answers were registered only if provided within the first repetition. The function of the second repetition was only to avoid frustration in patients that could not provide an answer; afterwards, participants were kindly invited to move to the following trial.

Task performance required six to eight minutes per participant.

7.2.6 Coding

Answers are coded in a binary way. One point is assigned whenever the participant can point to the target character, namely the one correctly representing the sentence content. Non-target answers or answers provided after the first sentence repetition correspond to zero.

7.3 Results

In what follows, I present the data provided by thirty-eight PADs out of the forty-one initially enrolled in the study. The reason for this is that three patients did not complete the task: in one case the decision to interrupt the task administration was taken by the experimenter because of behavioural disorders in PADs; one participants did not complete the task for independent reasons.

In what follows, I am going to present a group analysis (7.3.1) first, and then some in-depth data from the individual analysis (7.3.2).

7.3.1 Group analysis

Table 7.1 shows results organized per group and condition. In order to allow for a comparison, raw numbers are transformed into percentages, which correspond to the amount of target answers provided by participants in each group, and therefore represent their levels of accuracy.

Table 7.1. Percentages of accuracy (sd) per group in the six experimental conditions.

	CO	**PAD1**	**PAD2**	**PAD3**
n°	21	4	22	12
MMSE	(29-30)	(28-25)	(19-24)	(12-18)
SVO	97.02% (6.7)	75.00% (30.6)	84.09% (14.6)	70.83% (19.4)
SR	95.83% (9.9)	90.63% (11.9)	81.25% (16.7)	62.50% (24.4)
OR	92.86% (14.0)	90.63% (11.9)	63.64% (19.5)	60.42% (20.5)
POR	97.02% (8.7)	81.25% (23.9)	66.48% (22.4)	59.38% (23.3)
ClOR	90.48% (10.3)	90.63% (11.9)	67.61% (19.0)	61.46% (14.5)
ORdem	95.83% (7.2)	84.38% (18.7)	76.70% (15.2)	61.46% (19.5)

Data are subsequently analysed according to two factors: Group membership and Sentence Type. Significant results are detected with respect to both factors.

Starting from the Group comparison, the Mann-Whitney test shows that PAD1s do not perform differently from COs on any condition. However, PAD2s perform differently from PAD1s on ORs (Mann-Whitney, OR: U=12.500, Z=-2.290, p=.021) and on ClORs (Mann-Whitney, ClOR: U=15.500, Z=-2.073, p=.039). Significant differences emerge between PAD1 and PAD3 on SR (Mann-Whitney, SR: U=7.000, Z=-2.088, p=.042), on OR (Mann-Whitney, OR: U=4.500, Z=-2.409, p=.013), and on ClOR (Mann-Whitney, ClOR: U=3.000, Z=-2.597, p=.008). The Mann-Whitney U test also reveals that PAD2 perform differently from CO on SVO (Mann-Whitney, SVO: U=355.000, Z=3.343, p=.001), on SR (Mann-Whitney, SR: U=366.500, Z=3.584, p=.000), on OR (Mann-Whitney, OR: U=404.500, Z=4.384, p=.000), on POR (Mann-Whitney, POR: U=394.500, Z=4.351, p=.000), on ClOR (Mann-Whitney, ClOR: U=378.000, Z=3.663, p=.000), and on ORdem (Mann-Whitney, ORDem: U=395.000, Z=4.198, p=.000). PAD3 differ from PAD2 on SR (Mann-Whitney, SR: U=71.500, Z=-2.245, p=.028) and on ORdem (Mann-Whitney, ORDem: U=74.000, Z=-2.137, p=.037). PAD3s differ from CO on all conditions (p=0.000 in all conditions).

When the analysis focuses on Sentence type, comparison within groups reveals

significant differences within COs and PAD2s. COs are less accurate on ClOR in comparison to SVO (Wilcoxon related samples, CO: Z=-2.673, p=.008), SR (Wilcoxon related samples, CO: Z=-2.496, p=.013) and POR (Wilcoxon related samples, CO: Z= -2.484, p=.013).

Comparisons within PAD2 reveal that participants in this group perform significantly lower on OR in contrast to SVO (Wilcoxon related samples, PAD2: Z= -3.290, p=.001) and to SR (Wilcoxon related samples, PAD2: Z=-2.694, p=.007); on POR in contrast to SVO (Wilcoxon related samples, PAD2: Z=-2.633, p=.008) and to SR (Wilcoxon related samples, PAD2: Z=-2.326, p=.020). Crucially, for PAD2, no difference between OR and POR is found. Moreover, PAD2 perform with less accuracy on ClOR with respect to SVO (Wilcoxon related samples, PAD2: Z=-3.008, p=.003) and SR (Wilcoxon related samples, PAD2: Z=-2.346, p=.019). Finally, their performance on ORdem is significantly better than their performance on OR (Wilcoxon related samples, PAD2: Z=2.485, p=.013), on POR (Wilcoxon related samples, PAD2: Z=2.088, p=.037), and on ClOR (Wilcoxon related samples, PAD2: Z=2.161, p=.031).

7.3.2 Individual analysis

The group analysis reveals some tendencies that are worth further in-depth observations through the individual analysis. I will start by counting how many subjects in each group perform above chance on the different conditions (Table 7.2). For the sake of the present analysis, the 'above-chance' status is assigned to participants that provide at least six correct answers out of the eight trials included in each condition.

Table 7.2. Number of participants that perform above chance in the different conditions.

	CO (=21)	PAD1 (=4)	PAD2 (=22)	PAD3 (=12)
SVO	21	2	18	5
SR	20	4	16	5
OR	19	4	9	4
POR	20	3	7	4
ClOR	21	4	10	4
ORdem	21	3	14	5

Data in Table 7.2 confirm that within CO, PAD1 and PAD3, the number of participants that perform above chance ir rather balanced across conditions, either because of good mastery of all sentence types (as in the case of CO and PAD1), or because of generalized impaired comprehension of all conditions (as in PAD3).

As expected, data confirm asymmetries across conditions within PAD2. In particular, more than half of the participants in PAD2 perform above chance on SVO (18 out of 22 participants), SR (16/22) and ORdem (14/22), but not on OR (9/22), POR (7/22) and ClOR (10/22).

The performance of participants in this group deserves further in-depth analysis through contingency tables. The analysis will focus on the following pairwise comparisons: SR versus OR (Table 7.3), OR versus POR (Table 7.4), OR versus ClOR (Table 7.5) and OR versus ORdem (Table 7.6).

Table 7.3 shows that 10 PAD2 participants perform better on SR than on OR; six perform above chance on both conditions; 3 perform better on OR than on SR and three perform at chance on both conditions.

Table 7.3. Target comprehension of SR and OR in PAD2.

	SR +	SR -
OR +	6	3
OR -	10	3

Table 7.4 shows that the most common pattern is the one in which participants perform at chance both on OR and on POR; all other possible patterns being equally adopted: four participants perform above chance in all conditions; five comprehend OR better that POR and three present the reversed pattern, namely higher accuracy on POR than on OR.

Table 7.4. Target comprehension of OR and POR in PAD2.

	OR +	OR -
POR +	4	3
POR -	5	10

Table 7.5 shows that very few participants (3/22) perform above chance both on OR and on ClOR. The others equally distribute across the alternative patterns: six perform above

chance on OR but not on ClOR; seven show the reversed pattern with higher accuracy on ClOR than on OR; finally, six participants perform at chance both on OR and on ClOR.

Table 7.5. Target comprehension of OR and ClOR in PAD2.

	OR +	OR -
ClOR +	3	7
ClOR -	6	6

Table 7.6 shows that the most frequent pattern is the one in which participants comprehend ORdem above-chance and OR only at-chance. In contrast, the reversed pattern is the least frequent in PAD2 (i.e., above chance comprehension of OR and at-chance comprehension of ORdem).

Table 7.6. Target comprehension of OR and ORdem in PAD2.

	OR +	OR -
ORdem +	5	8
ORdem -	3	6

Overall, contingency tables contributes to a better understanding of performance patterns among PAD2 participants. Results from the group analysis and from the individual analysis are discussed in the following section.

7.4 Discussion

The present task aims at investigating four aspects of the syntactic competence of patients with Alzheimer's disease. Questions deal with (*i*) the presence of a syntactic deficit in PADs, (*ii*) the sensibility shown by participants towards different kinds of syntactic factor manipulation, (*iii*) the form that the deficit takes, and (*iv*) the level of dementia at which major changes appear at the syntactic level.

In order to answer those questions, data collected from 38 PADs are analysed.

7.4.1 Question (i): Do PADs suffer from a syntactic deficit?

In order to answer the first question concerning the presence of a syntactic impairment in PADs, I start from a comparison between the patterns of performance in the four different groups.

The statistical analysis does not provide signs of discontinuity between COs and PAD1s; however, differences between controls on one hand, and PAD2 and PAD3 on the other clearly emerge. This first observation is sufficient to reject the hypothesis that COs and PADs perform the same. Therefore, some kind of difference between the two groups is present.

The second important thing to consider is whether the difference between COs and PADs is of a quantitative or of a qualitative type. If I only compared PAD3 to CO, I might get to the conclusion that the two groups perform alike (i.e., without major asymmetries between conditions), and therefore are only quantitatively different. The conclusion would be based on the fact that, according to the within group analyses, no significant difference across conditions appear neither in PAD3, nor in COs (except for a dis-preference for the substandard ClOR form in COs). However, this is not the correct conclusion. The observation of the performance pattern in PAD2 calls for a qualitative difference between CO and PAD2. Crucially, asymmetries across conditions are attested in the latter (PAD2) but not in the former group (CO), which supports the claim for a qualitative difference in performance patterns between healthy and cognitively impaired elderlies. Ultimately, the first conclusion of the present study is that PADs suffer from a syntactic deficit, because they have a pattern of performance which differs from the one of COs in many different ways.

Overall, results from RC comprehension resemble results from the task on Wh-question comprehension. As discussed in section 6.4, PAD1s are accurate in all Wh-questions in use; PAD2s show asymmetries across conditions (section 6.4.2.2), mainly in the form of a better accuracy on WhichNP questions than on Who questions. Finally, PAD3s show a generalized impoverished comprehension of all conditions, without major asymmetries. That is precisely the general pattern found also in the present task on RC comprehension. Results from the two studies taken together suggest that a syntactic impairment is at play. The issue is further discussed in the following sections.

7.4.2 Questions (ii) and (iii): What kind of manipulations are PADs sensitive to and how does their syntactic deficit manifest itself?

In order to answer the questions concerning the sensitivity to the manipulation of different factors and the characteristics of the syntactic deficit, I will go through within group analyses across conditions.

7.4.2.1 Controls

As already pointed out, participants in the control group perform alike in all conditions except one: ClOR, namely object relatives with a resumptive clitic. Overall, their comprehension of the experimental stimuli reaches clearly above 90%. Based on the results, I assume that elderly speakers (with low levels of education) perform the task successfully, without major difficulties. In addition, the lack of ceiling effects suggest that the task is not trivial and therefore it is appropriate for testing comprehension in aging speakers[69].

The lower performance on ClOR and its asymmetry with respect to SVO, SR and POR conditions, is certainly due to the fact that ORs with clitic resumption are not in use in Standard Italian. The structure has been classified as substandard (Guasti & Cardinaletti, 2003), but, except for data on children' elicited production and comprehension (Contemori & Belletti, 2014; Guasti & Cardinaletti, 2003 for production), there are not available data on its frequency in adult spontaneous speech. Moreover, what might appear as a very specific deficit in the performance of controls, it is actually the result of a coding procedure. The level of accuracy strictly reflects the number of provided answers that correspond to the target image, while it does not take into account the (few) occurrences in which COs do not provide an answer. Indeed, the sentences classified as "non-target" also include trials for which COs expressed grammaticality judgements, instead of answering. In other words, participants in the CO group sometimes autonomously switch to a Grammatical Judgement Task, although this kind of performance was neither requested nor awaited. From the statistical point of view, this translates into a lower level of accuracy on the condition at stake.

[69] The present study does not include a control group of young adult speakers. However, data from previous study (see Contemori & Belletti 2014 for an overview) confirm that healthy native speakers of Italian complete comparable tasks with levels of accuracy that are always at ceiling.

From the linguistic point of view, controls' attitude is informative of the fact that resumption in OR is not a much preferred option among adult speakers, at least not for the diatopic variety spoken by the enrolled participants. This reinforces the idea that the presence of resumptive clitics in child ORs does not strictly mirror adult speech; rather, it is a specific characteristic of the acquisition path, with the clitic playing the function of favouring the extraction of the object in sentence production (Contemori & Belletti, 2014).

In what follows, we will see whether PADs benefit from a facilitating effect similar to the one found in children or they pair their healthy peers with respect to object clitic resumption.

7.4.2.2 PAD1

The within group and across conditions analysis is rather straightforward when it comes to the PAD1 group with patients characterized by a mild level of impairment. In this group, no significant difference across conditions is to be detected, although a lower level of accuracy (75%) is registered in correspondence to SVO, the baseline condition.

The individual analysis reveals that this lower percentage is mainly due to the poor performance of one single individual (VI_16), who provides target answers only to 3 out of 8 experimental trials in the SVO condition. However, the same individual performs much better in all other conditions and provides a number of target answers that overall resembles the ones of the other participants in the same group. In light of this data, it cannot be concluded that patient VI_16 suffers from a specific impairment, whose seriousness highly exceeds the profile expected in correspondence to her reported level of dementia (25/30 in the MMSE). Moreover, the subject speaks the same Italian variety as the other participants involved in the study, such that it is also necessary to discard the hypothesis that she might have a different syntax of subject positions. The question concerning why the patient performs poorly on SVO sentences is left unanswered.

Leaving aside this fact, participants in PAD1 perform much alike to COs, except for the lack of a peculiar attitude towards the ClOR condition, otherwise registered in COs. Based on the data, it can then be assumed that there is some kind of continuity between the comprehension abilities of controls and patients in the first stage of the disease. However, a word of caution is necessary here: given the small number of

participants included in the PAD1 group, it cannot be excluded that different results might be reached through the enrolment of more participants. However, in light of the available data, I can conclude that no major sign of syntactic impairment appears in PAD1.

7.4.2.3 PAD2

The observation of participants in the PAD2 group is highly informative. In this group, performance is overall significantly different from that of COs and various asymmetries across conditions emerge too.

The most striking (although not surprising) asymmetry is the one between SRs and ORs. Participants in the PAD2 group have a specific impairment on the comprehension of ORs. The claim is based both on the percentage of accuracy from the group analysis and on the individual analysis. Overall, comprehension is at 79.5% on SRs, while it drops to 65.1% on ORs. Moreover, the number of participants who perform above chance is sixteen in the former, versus nine in the latter (out of 22). Overall, only six participants can perform above chance on both conditions; ten perform above chance on SR, but crucially not on OR. The reversed pattern, with above chance comprehension of ORs but not of SRs, is very rare: only three patients show it.

In line with Friedmann, Belletti & Rizzi (2009, and much subsequent work), I adopt a RM account for the processing of ORs. I take the subject/object asymmetry as the sign of a major difficulty at extracting the internal DP from its original position and moving it across the subject DP, which intervenes in the derivation. The intervention is to be ascribed to the particular configuration that emerges from a comparison between the two relevant phrases: the one corresponding to the internal argument and the one for the external argument. The configuration, reproduced in Table 7.7, takes into account the [+Rel] scope-discourse feature for RCs, and lexical restriction ([+NP]). A mismatch in number and animacy is not taken into consideration in the present task because both DPs are singular and animate, while gender is not relevant for the computation of locality effects in Italian (Belletti et al., 2012). The two elements differ only for the presence of a [+Rel] feature on the head of the RC, such that their feature arrays are in a relation of inclusion, in which the potential intervener (the subject) counts on a subset of the features that characterize the moved DP.

Table 7.7. Feature arrays of the two relevant positions in the derivation of OR clauses.

HEAD OF THE RC		SUBJECT DP		(EXTRACTED) OBJECT DP
la nonna	*che*	*la bambina*	*bacia*	*<la nonna>*
+ Rel		-Rel		+Rel
+ NP		+ NP		+ NP

The difficulty that raises from a configuration of inclusion is widely attested in children (see Chapter 5 for an overview), while adult speakers seem to cope well with it (although asymmetries might nonetheless emerge in on-line tasks). Here, the re-emergence of the deficit in the considered population supports the idea that the derivation of ORs entails a higher degree of complexity and therefore it is more sensitive to disruption in case of processing impairment.

The assumption is confirmed by the low levels of accuracy on other conditions that manipulate factors within the class of object relatives. I refer in particular to two conditions, namely POR and ClOR sentences. While this is not the case for ORdem sentences, which have much higher levels of accuracy in comparison to the other conditions in the OR category. In what follows, I will analyse each condition in turn, in order to evaluate their informative value for the study.

I introduced POR in the present study in order to verify whether the use of a passive voice can facilitate sentence processing in PADs. This sentence manipulation is particularly interesting because it is effective both in adults and children. As for acquisition, previous research (Contemori & Belletti, 2014) proved that the presence of a passive voice improves the comprehension in young Italian-speaking children; moreover, the strategy also emerges in elicited production as a preferred option over canonical ORs. Cross-linguistic studies covering European Portuguese (Costa et al., 2011) and German (Adani et al., 2013; Sanfelici et al., 2014) further confirm the claim. What is even more interesting for the present study is the fact that the preference for POR over OR is observed in adult speakers too. Contemori & Belletti (2014) show that PORs are spontaneously generated in correspondence to stimuli eliciting ORs to a rate that corresponds to ca. the 90% of the total amount of expected ORs. This assures that the strategy cannot be confined to acquisition because it is a well-established option in the grammar of adults, who make wide use of it. A previous study with a small group of ten speakers with dmeentia (Caloi, 2013) reported good comprehension of PORs (ca. 77%),

despite poor accuracy on ORs (ca. 57%). With these premises, a facilitating effect for PORs is expected in in the present study too. Data from the present study do not confirm the hypothesis, though. Accuracy on PORs is at 68.6%, only 3.5 points above the level of accuracy on ORs, a difference that does not reach statistical significance. Moreover, seven PAD2 participants perform above chance in this condition, versus the 9 participants who reach the same result on ORs. Indeed, the contingency table (Table 7.4) shows that only four participants perform above chance on both conditions, while almost half of them (10 participants) do not comprehend either of the two conditions above chance. In a few cases, participants perform better on one or the other: three PAD2 perform better on PORs than on ORs, while five show the opposite pattern. These observations lead to the conclusion that PAD2 participants are as impaired at POR comprehension as they are at OR comprehension. In other words, the use of a passive voice does not improve OR comprehension in PAD2. Why are PORs not an effective strategy for PAD2?

I can think of a variety of reasons that might hide behind the poor performance of PAD2 on PORs. First of all, it might be the case that passive is impaired *per se* and therefore its use is not available within a RC. In addition, the inefficiency of the *smuggling* strategy can in principle be ascribed to different reasons too. For instance, a specific impairment might occur at the VoiceP level, the functional projection that allows the realization of a passive. Alternatively, if we assume that it is a discourse feature that triggers the derivation of passive and attracts the whole VP across the subject and moves it higher in the structure (Collins, 2005; Gehrke & Grillo, 2009), it is also quite likely that the sentence processing crashes during the analysis of the feature arrays that regulates the derivation. If this speculation were on the right track, it would be probably possible to trace a parallelism between the deficit in passive and the one in OR, in consideration of the demanding analysis of their feature arrays. However, the hypothesis should be discarded based on previous studies on the processing of passive sentences in German- and English-speaking PADs: Bates et al. (1995), Bickel et al. (2000), Grossman & White-Devine (1998), Small et al. (2000) and Waters et al. (1998) all fail at finding significant asymmetries between the comprehension of active and passive sentences. Although those studies do not cover Italian, I will assume that PADs are spared at processing passive structures.

One more reason that could lay behind the poor performance of PAD2s on PORs is the kind of passive in use in the experimental sentences. Italian counts different forms of passive: the copular form in use in the present task (e.g. *La nonna è baciata dalla bambina* 'The granny is kissed by the child'), a copular form with the auxiliary *venire* (e.g. *La nonna viene baciata dalla bambina*, 'The granny comes kissed by the child'), and the *si*-causative form (e.g. *La nonna si fa baciare dalla bambina*, 'The granny makes herself kiss by the child'). The three forms are not equivalent from the point of view of acquisition, given that they neither emerge simultaneously in the spontaneous speech of Italian children nor their comprehension is mastered at the same age. In particular, Contemori & Belletti (2014) find that *si*-causative passive forms are attested earlier than copula passives. The result finds a clear counterpart also in a comprehension task, in which the *si*-causative construction collects better results than the copular passive does (both in the full and in the reduced form). Further support in favour of the claim that *si*-causative structures precede other forms of passive comes from another elicitation task. Manetti & Belletti (2015) try to prime different forms of passive in children. In this, *si*-causatives prime more target outputs than the passive form with *venire*. In turn, *venire* primes more target outputs than the copular passive with *essere* ('be') does. Together, results from the two studies (Manetti & Belletti, 2015; Contemori & Belletti, 2014) show that children begin to produce passive structures at around age five (although first productions can appear earlier), starting with the *si*-causative form, and following with the copular passive with *venire*; finally, the form with *essere* is acquired too.

With respect to PADs, Bates et al. (1995), previously found that English speakers rarely produce passive structures, but when they do, these are more often of the *get*-type. In light of these data, it cannot be excluded that the use of a different form of passive, for instance the *si*-causative one, might give better results with PADs too. Unfortunately, the gathered data are not informative about the mastery of passive *per se* and the status of the different forms of passive in the competence of PAD2. Future research will certainly need to take into account the issue by testing passive independently of its use in RC. Results from such studies will therefore shed new light on the results discussed here and finally reveal whether the poor comprehension of PORs is caused by a specific impairment of passive derivation or not. In the former case, a common explanation behind the

impairment on passives and PORs will be pursued. In the latter case, it will be necessary to discuss the reasons why passive derivation works fine in matrix clauses but not in RCs.

For the moment, I can only speculate on the latter scenario and say that the poor comprehension might also be due to patients' difficulties in dealing with a structure that requires the satisfaction of, at least, two discourse features: one for the passive form and one for the relative clause. If we assume that a discourse feature attracts the VP to a higher functional position (Gehrke & Grillo, 2009) and that the head of the RC is endowed with a [+Rel] feature (Rizzi, 2006), it is conceivable to think that the parser might suffer from a cumulative effect. The presence of more discourse features requires the parser to take crucial steps for the achievement of two different goals. The impossibility to correctly complete and satisfy both procedures (the one for the RC and the one for the passive) might follow from the apraxia that affects patients. This is the inability to take the necessary steps in the required order with the purpose of completing a task (see Chapter 2). In the case of PORs, the presence of two tasks to be completed (in the sense of satisfying two discourse features), might overwhelm PAD2 and exceed their abilities, which are limited by apraxia in the form of an impossibility to follow the required procedure for the derivation. The hypothesis could also be formalized from a theoretic point of view as an instance of the Universal Freezing Hypothesis proposed by Hyams & Snyder (2005) for language acquisition: according to the authors, under certain circumstances (i.e. acquisition), sub-extraction is not allowed from a phrase that has already been moved. In the authors' view, this is the main reason why young children cannot complete the *smuggling* operation[70]. Correct processing of passive but failure on POR for PAD2 patients would indeed perfectly fit the idea that PADs correctly smuggle the VP across the subject (for the passive part), but then fail at sub-extracting the Object-DP from the moved phrase in order to move it to the head of the relative clause.

It is probably not worthy at this point to further develop the discussion on the reason why passive does not ameliorate the comprehension of ORs, given the lack of relevant

[70] More recently, Snyder & Hyams (2015) discussed the idea that children might process passive structures correctly under a strict version of RM and despite the Universal Freezing Hypothesis. Through the review of previous studies, Snyder & Hyams (2015) pointed out that even 3-year-old children can succeed in passive production when the experimental conditions endow the derived subject of a quantificational feature of the [+Wh] or [+Topic] kinds. The feature would make the logic object different enough from the logic subject in order to allow the movement of the former across the latter. Authors discussed cases similar to the one in use in the present study: the derived subject is endowed with a [+Q] feature, such that the passivization should be allowed under RM in their view. However, this was not the case for PAD2 patients.

data. I therefore postpone the discussion to the moment when data on the processing of passive will be available. For the time being, it can only conclude that copular *be* passive does not improve the comprehension of ORs in patients affected by a moderate level of dementia.

Another factor I manipulated within OR is the presence of a clitic resumptive pronoun in the RC. Again, the reason for introducing the condition in the experimental materials is that resumption often appears in children production, while it never does in adults (Contemori & Belletti, 2014). As for comprehension, Contemori & Belletti (2014) found that resumption (slightly but significantly) improves comprehension in Italian-speaking children between the age of 6;5 and 8;10. Thus the question was whether clitic resumption can be of any help for the comprehension of ORs in PADs. Results say that this is not the case, though. Accuracy on ClOR is only (nearly) 2 points higher than on OR, and the difference does not reach statistical significance. The number of participants that provide at least six target answers out of eight trials is nine for ORs and ten for PORs. Among these, only three participants perform above chance on both conditions. The other participants distribute across the free alternative patterns: six perform above chance only on OR; seven PAD2s perform above chance only on ClOR, while six subjects do not reach above chance accuracy in either of the two conditions. According to the data, there is no facilitating effect provided by the presence of a resumptive clitic in the embedded clause[71].

While discussing data on children production and comprehension, Contemori & Belletti (2014) speculated on the idea that the stranded clitic might enrich the feature arrays of the extracted big DP, thus favouring its movement across the subject. Under this view, the presence of the clitic prevents the derivation from crashing, in accordance with the Principle of RM. I assume that, with respect to this phenomenon, PAD2s do not benefit from the presence of a resumptive clitic. Their parser blocks the movement of the big object DP across the subject DP in name of the fact that the two elements share relevant features (see discussion above on ORs). The optional presence of a stranded clitic does not play any role with respect to this. In addition, future research should investigate

[71] Lack of facilitated comprehension of OR with resumption is reported also for Hebrew-speaking children in Friedmann, Belletti & Rizzi (2009) and in Costa et al. (2014).

in depth the status of clitic processing in PADs, given that the present study and previous literature fail at providing information on the status of clitic processing in PADs.

The last experimental condition to be analysed here is the one that entails the presence of a distal demonstrative pronoun (*quello*/*a*) in the determiner of the modified DP. The reason for introducing this condition in the study is to verify the relevance of mismatch between arguments' feature sets for the derivation of ORs. Sanfelici & Poletto (in press) claim that sentences of the kind *Mostrami quella che la bambina bacia* ('Show me the one that the child kisses') entail a light functional noun of the [PERSON] or [THING] type, which merges in the internal argument position (in the case of ORs) and moves to the CP-layer of the RC. The noun is functional in nature and does not entail a lexical restriction as traditionally meant (e.g., *child*, *doll*, *table*, etc.). Table 7.8 reproduces the relevant feature arrays of the elements involved in the derivation:

Table 7.8. Feature arrays of the arguments involved in the derivation of ORdem sentences.

DETERMINER	SPECCP		SUBJECT DP		(MOVED) OBJECT DP
quella	[PERSON]	*che*	*la bambina*	*bacia*	<[PERSON]>
	+Rel - NP		- Rel +NP		+ Rel - NP

The two relevant feature arrays, the one corresponding to the Subject and the one corresponding to the Object, differ for two features. The moved Object is a light noun, which entails the [+Rel] feature that triggers the movement. In contrast, the potential intervener in the derivation, namely the Subject DP, is endowed with a lexical restriction, while it obviously lacks [+Rel] for relative clauses. Therefore, with respect to the two relevant features, the arguments are in a configuration of disjunction.

PAD2 participants in the present study show to be sensitive to this configuration and indeed provide high levels of accuracy in the ORdem condition. In particular, accuracy on ORdem (75%) sharply differs from that on OR (65.1%). The number of participants in the PAD2 group that perform above chance level on ORdem is fourteen, against the nine participants performing above chance on ORs (out of twenty-two). Out of these, only five PAD2 subjects perform above chance on both conditions; six perform at chance on neither of the two conditions; eight patients perform above chance on ORdem but not on OR, while the reversed pattern (above chance performance on OR but

not on ORdem) emerges only in three participants. Moreover, the level of accuracy on ORdem does not reach a statistically significant difference when compared to results on SVO and SR. This observation supports the claim that ORdem can sharply improve the comprehension of ORs, bringing comprehension to levels that resemble the ones found in simple declarative sentences (SVO) and in SR. I assume that the successful comprehension of the ORdem condition is due to the specific feature configuration that the presence of a demonstrative and a light noun create. The observation pairs the results from the task on Wh-question comprehension in which a mismatch in lexical restriction improves the comprehension of interrogatives of the Which-type with respect to Who questions, both within the subject and the object conditions.

In addition, I should also remark that the use of a distal demonstrative is particularly natural in the experimental design in use because participants are forced to choose between two characters. Pragmatic felicity might contribute to an improvement of sentence comprehension *per se*. In order to evaluate the real size of the phenomenon of feature mismatch with respect to pragmatic felicity, it would be useful in future to compare ORdem sentences to SR with a demonstrative pronoun (5):

(5) *Mostrami quella che bacia la nonna*
'Show me the one that kisses the granny'

A higher accuracy on SRs entailing distal demonstratives with respect to SRs with lexically restricted heads would confirm the role played by pragmatic felicity in sentence comprehension. However, given that ORs are harder to process for adults (see Chapter 5 for considerations on adults) but not infelicitous from the pragmatic point of view, the asymmetry between OR and ORdem cannot stem only from pragmatic felicity.

With respect to the results gathered in correspondence of the ORdem condition, I can therefore claim that PADs are sensitive to enhanced mismatch in features between the two relevant positions in the derivation and that their capacity to process feature arrays in a configuration of disjunction is spared, while they are impaired at processing structures with feature arrays in a configuration of inclusion (see the discussion for ORs).

Summing up, PAD2 participants have a specific impairment at OR comprehension with respect to SVO sentences and SRs. Neither passive voice (POR) nor resumptive clitics (ClOR) improve comprehension. The only strategy that improves comprehension of ORs in PAD2 is the presence in ORdem of a lexical mismatch between the two relevant positions.

7.4.2.4 PAD3

After the review of the results produced by participants in the PAD2 group, I now address my attention to the third sub-group of PADs, namely PAD3. This last group counts participants characterized by a severe level of dementia, which corresponds to MMSE scores ranging between 12 and 18. PAD3 significantly differ in their performance from COs in all conditions; however, the within group analysis across conditions does not reveal any significant difference across the structures considered in the study. As a group, PAD3 reaches similar levels of accuracy in all conditions. Moreover, the individual analysis reveals that in all conditions, more than half of the participants performs at chance level. As a group, performance is around chance level on most experimental conditions, except for SVO sentences.

Indeed, the higher level of accuracy (70.8%) corresponds to SVO sentences. However, this is only descriptively higher than the accuracy on other conditions. I assume that the lack of A' movement keeps the level of complexity rather low, thus favouring a better comprehension of simple declarative sentences, with respect to the relative clauses in use. However, there is no statistical difference between SVO and all other conditions. I take this to be a sign of the fact that PAD3 participants are generally impaired at coping with the task.

At first, these results might resemble the ones found in PAD1, in the sense that no significant difference across conditions is detected. However, results in the two groups are not similar at all. Rather, they are the output of two diametrical different situations. PAD1 do not show any statistical difference across conditions because their unimpaired processing abilities allow for a good comprehension of all kinds of sentences. In contrast, in PAD3 the lack of significant differences across conditions is due to a generalized impairment that disallows for successful task comprehension and completion.

7.4.3 Question (iv): At which stage of the disease are PADs concerned by syntactic impairment?

I now turn to the last issue I addressed in the introduction, namely the one concerning the level of disease severity at which major changes take place in syntactic processing abilities. An answer to this question already emerged from the data discussed so far; still, some explicit observations follow.

I subdivided the participants enrolled in the present study into three groups according to their level of dementia: mildly (PAD1), moderately (PAD2) and severely (PAD3) impaired subjects. The three groups overall provide patterns of performance that significantly differ one from the others. In particular, the first group (PAD1, MMSE scores 25-28) performs quantitatively and qualitatively similarly to controls. I interpret this as a sign of spared processing abilities: no impairment in the comprehension of RCs is visible in the first stage of the disease.

Results are quite different in the case of subjects with a moderate level of dementia (PAD2, MMSE scores 19-24): major changes to the capacity of sentence processing take place at this level of the disease. These manifest themselves in the form of a specific impairment on the computation of ORs, as described above.

Finally, in the third group (PAD3, MMSE scores 12-18), the capacity to correctly comprehend RCs is crucially damaged. The impairment is so severe that it causes a generalized drop of the performance on all conditions, and it ultimately nullifies the subject/object asymmetry. Even the comprehension of SVO sentences is reduced: it only descriptively exceeds that of RCs.

Capitalizing on these data, I claim that major changes to the ability to correctly comprehend RCs take place in a stage of the disease in which subjects are characterized by a moderate level of dementia. The transition to very low levels of performance is very fast afterwards, so that patients soon lose the capacity to meet the requirements for achieving accuracy on all RCs.

In order to complete the overview, I should point out that the study was initially meant to enrol also patients with MMSE scores lower than 12; however, all attempts made in the pilot phase of the study were unsuccessful, due to the subjects' inability to comprehend the task and participate in the experimental section with a collaborative attitude.

I would also like to remark that the results from the three groups resemble those presented in Chapter 6 with respect to the comprehension of Wh-questions. The detected pattern of disruption through levels of dementia holds true across different tasks. This corroborates the claims above.

7.5 Conclusions

In order to answer the four research questions addressed in the introduction to the present chapter, I sampled the comprehension of RCs in PADs. In light of the collected data, I provided an answer for each question, although few unclear aspects certainly persist and require further research.

The first goal was to determine whether patients suffer from a syntactic deficit. This first question receives a positive answer in light of the fact that PADs perform quantitatively (PAD2, PAD3) and qualitatively (PAD2) differently from participants in the control group.

Concerning question number two, I claim that the performance of PAD2 participants is characterized by sensitivity to the extraction site: this manifests in a sharp asymmetry between the comprehension of SRs and of ORs, with the latter comprehended only with a low level of accuracy. Moreover, among the features manipulated in the different conditions, PAD2s do not show sensitivity to the use of a passive voice and of a resumptive clitic in ORs. By insensitivity, I mean that the two strategies do not contribute at ameliorating the comprehension and are therefore ineffective. In contrast, PAD2 participants process ORs with better accuracy in presence of a lexical mismatch between the moved object and the intervening subject. That is the case of ORdem. This brings comprehension to levels similar to the ones found in SVOs and SRs.

In light of the pattern of sensitivity summarized above, I claim that the performance of PAD2 can be accounted for by adopting an approach based on the Principle of Relativized Minimality (Friedmann et al., 2009; Rizzi, 1990; and much subsequent work). Participants fail at comprehending conditions that entail the movement of the Object DP across the Subject DP. I refer in particular to the OR, POR and ClOR conditions. I assume that the derivation of ORs and ClORs is blocked due to the fact that in both structures the relevant feature arrays are in a relation of inclusion. As for PORs, the reasons why *smuggling* does not allow for a successful extraction of the Object DP in PAD2 group are

not clear yet, but I suggest that the results might be due to the kind of passive in use, namely the copular one. However, this is precisely the aspect of question (*iii*) that is left partially unanswered and that deserves further investigation in future research. What is clear is that, despite very poor performance on feature configurations of inclusion, the parsing of features arranged in a configuration of disjunction is spared in PAD2. That is precisely the case of ORdem sentences, which are accurately comprehended by the participants. The nature of the impairment in PAD2 must therefore be ascribed to a stricter version of RM that disallows object extraction in a configuration of inclusion, but allows disjunction. Indeed, the mismatch in lexical restriction appears as an effective strategy for ameliorating the comprehension of ORs in PADs. Moreover, the observation confirms data from the previous task on Wh-question comprehension about the high sensitivity of patients to the mismatch in lexical restriction.

Alternative accounts to poor syntactic comprehension in aphasic speakers were not considered because these had already been discarded in the discussion of results from the task on Wh-question comprehension (see Section 6.4.2.3).

Finally, the present study addresses the issue concerning the stage of the disease at which syntactic computation abilities undergo a process of disruption. Data analysis based on the level of dementia of each participant allowed for the individuation of the phase in which patients start showing difficulties in the comprehension of RCs. According to the data, comprehension starts dropping in patients in the PAD2 group, namely in subjects affected by a moderate level of dementia, corresponding to MMSE scores ranging from 19 to 24. In contrast, mildly impaired participants (PAD1) have unimpaired comprehension of all conditions, according to a pattern that closely resembles the one of COs. Finally, comprehension drops in all conditions for participants affected by severe impairment (PAD3); this shows that the syntactic principles that regulate processing are ultimately disrupted along with the worsening of the disease.

8 CONCLUSIONS

In this work I presented results from three experiments that dealt with the linguistic deficit of Italian-speaking patients with Alzheimer's disease. The study was an attempt to answer three research questions that concern patients' syntactic competence.

The first question dealt with the emergence in PADs of a syntactic deficit. The issue was addressed under two perspectives: on one hand, the ability to retrieve syntactic information attached to entries in the mental lexicon, on the other hand, the ability to process sentence structures characterized by argument extraction. As for the first aspect, in Chapter 4 I showed that Grammatical Gender retrieval is spared in Italian-speaking patients. The claim is based on their mean accuracy on the experimental trials, which was at 93.9%. However, an analysis of the 134 mistakes they made offered the opportunity to get an insight into the (few) difficulties they experience. For instance, mistakes on regular nouns were virtually absent, while mistakes on irregular and opaque nouns were more frequent (although as low as 3.5% with respect to the whole of experimental trials). This suggests that the form-driven strategy for GG retrieval is completely spared in PADs, while direct access to the information in the mental lexicon occasionally fails. The hypothesis is supported by a further observation, which point in the direction of a tendency to adopt over-regularization strategies. Patients adopt the form-driven strategy and base their answer on the final vowel of the irregular nouns presented in the task. Because of the asymmetry between regular and opaque/irregular nouns, results resemble those collected in studies on regular and irregular verbal morphology (Colombo et al., 2009; Walensky et al., 2009). Therefore, the study further supports the assumption of an asymmetry between mental lexicon and mental grammar, as proposed by Ulmann (2001, and subsequent work). PADs suffer major disruptions at the declarative knowledge, such that direct access to syntactic information is blocked. In contrast, the procedural mechanism for form-based retrieval is spared.

However, the study also pointed out an increased percentage of errors on derived nouns with respect to simple nouns. Although each derivational suffix in use can assign only either masculine or feminine GG, PADs do not systematically benefit from their presence. If their direct access to the GG entailed in derivational suffixes were spared,

they would be able to use it in order to assign the correct GG of the derived noun. In contrast, the higher number of mistakes in correspondence to derived nouns suggests that PADs are sensitive to morphologically complex words, but in a negative way because they cannot successfully analyse the word components. In this, PADs differ from an aphasic speaker previously tested by Luzzatti & De Bleser (1996).

Overall, I conclude that patients are not impaired with respect to their ability to retrieve syntactic information about nouns: the observations above are based on a small number of mistakes and therefore represent only slight tendencies. Finally, with respect to this first task, it is necessary to keep in mind that grammatical gender is a highly language-specific system, such that conclusions from the present study cannot be extended to speakers of other languages, for whom further research is desirable.

The second aspect of syntactic competence I investigated is the ability to process sentences characterized by argument extractions. For this purpose, PADs completed two comprehension tasks on Wh-questions and on relative clauses. The tasks unveiled specific impairments. The claim is based on patients' performance, which was different from that of controls under various aspects. First, patients performed differently according to their level of dementia: patients with moderate and severe impairment were overall less accurate than controls. Second, moderately impaired patients showed significant asymmetries among conditions, while comprehension was generally reduced in all sentence types for severe patients. The latter group was also characterized by higher variability, especially in the case of Wh-question comprehension. I conclude that patients performed both quantitatively and qualitatively different from controls and for this reason they must be considered impaired at sentence processing.

Once signs of impairment were unveiled in a specific area, I could address a second research question, namely the one concerning the disruption pattern behind the impairment. The in-depth observation of patients with moderate dementia was particularly helpful in order to individuate which account can better describe their performance.

In Chapter 6 I reported that PAD2s perform better on WhichNP questions than on Who questions. For both WhoO and WhoS questions, I ascribe the poor performance to enhanced sensitivity to locality effects (Rizzi, 1990) in the grammar of PADs. In the WhoO questions in use, the movement of the object DP is blocked by the subject DP. As

for WhoS questions, I hipothesized that the object clitic pronoun and the bare Wh-element enter in a configuration that disrupts the possibility for the subject *Chi* to build the relevant dependencies for Wh-movement and verb agreement (along the lines of Guasti et al., 2012). In both cases, this is due to the fact that the extracted DP and the potential intervener share relevant features. The claim is further supported by the fact that processing sharply improves whenever the feature arrays at stake mismatch for a relevant feature. As for the conditions in use, the crucial mismatch is represented by lexical restriction. This is precisely the reason why Which questions are overall better comprehended both in the subject and in the object condition: the extracted element is characterize by a lexical restriction, which makes its feature array crucially different from that of the potential intervener. Therefore, locality effects are hindered and the dependency between the extracted argument and its target position in CP can be built. The claim on WhichNP elements concerns the specific configurations in use in the present study. It is excluded that WhichNP elements can have *a priori* a facilitating or a disrupting effect.

Results in Chapter 7 are consistent with the data above. Patients with moderate dementia are specifically impaired at comprehending object relatives in comparison to subject relatives. In order to verify patients' sensitivity to syntactic factors, I introduced three more OR conditions in the task: object relatives with a resumptive clitic, passive object relatives, and object relatives with a mismatch in lexical restriction between a light functional noun and the subject DP. Only the third kind of manipulation improved patients' performance on object relatives. In contrast, PADs were insensitive to clitic resumption. As for passive object relatives, I expected PADs to perform well on this condition. The prediction was based on previous results from language acquisition and healthy adults (Contemori & Belletti, 2014), and on patients' good processing of passive structures in other languages (Bickel et al., 2000 for German; Small et al., 1998 among others for English). However, that was not the case for the present study, most probably because of the kind of passive in use, namely the copular form with *essere* ('to be'). Italian counts on more forms of passive and I cannot exclude that different results would be obtained with structures that appear earlier in L1 acquistion, as for example *si*-causative passives (Contemori & Belletti, 2014; Manetti & Belletti, 2015). Moreover, a cumulative

effect might arise in passive, due to the necessity of satisfying two discourse features (one for the relative and one for the passive) within the same sentence structure.

Overall, data on relative clause comprehension confirm the results from the previous task on Wh-question comprehension: PADs are more sensitive than controls to locality effects. Patients with moderate dementia are impaired at computing dependencies that entail a crossing movement between two arguments, when their features are in a relation of inclusion. In contrast, crossing movements are allowed when the involved feature arrays are in a relation of disjunction (in the sense exemplified in Belletti et al., 2012). In both tasks in the present study, disjunction is represented by a mismatch in lexical restriction. The observation is particularly interesting with respect to the well-known anomia that characterizes PADs. Despite anomia, the integration of fully-fledged lexical DPs represents a facilitating factor for PADs under the appropriate configuration, namely when the involved arguments mismatch in lexical restriction (with respect to conditions in which they are both either lexically-restricted or not).

The collected data can also provide an answer to the third research question, which concerns the level of impairment at which syntactic impairment emerges. In order to answer this question, I assessed patients' level of dementia through the Mini Mental-State Examination Test (Folstein et al., 1975). According to the obtained scores, I subdivided participants into mildly, moderately, and severely impaired patients. Therefore, I could observe that PADs with mild dementia (MMSE scores: 25-28) perform as accurately as controls in all conditions. Their ability to process Wh-questions and RCs, and to perform a sentence-to-picture matching task is unaltered. Major changes take place when dementia reaches a moderate stage (MMSE scores: 19-24): patients' ability to compute complex syntactic derivations is impaired in the forms described above. The worsening of the disease gradually reduces patients' ability to perform the task. The third group is characterized by variability: most PADs with MMSE scores lower that 19 can perform the task only at chance level. In these patients, comprehension is reduced even on active declarative sentences. Such poor results do not allow to determine whether patients are experiencing only syntactic impairment or if other forms of cognitive impairment disrupt the task performance.

In Chapter 6 and 7 I claim that the syntactic impairment manifests itself in correspondence to moderate impairment. However, two issues should be considered.

First, I would like to remark once again that the collected data only concern comprehension abilities. I cannot exclude that tests on production might reveal slightly different patterns, with patients showing production deficits even at an earlier stage of their disease. This brings to the second issue, namely the fact that in AD neuropsychological symptoms clearly manifest themselves only long after the onset of neurological damages (Braak & Braak, 1997). Their appearance is gradual, because patients can compensate them for a long time, before showing overt impairment. Therefore, it is plausible to assume that a deficit in sentence comprehension is present but covert in earlier stages of the disease (and for this reason not clearly detectable in spontaneous speech, as discussed in Chapter 2).

In conclusion, I claim that patients are spared at using procedural mechanisms for the retrieval of syntactic information, while they are impaired at processing sentence that entail argument extraction. The impairment manifests itself in the form of enhanced sensitivity to locality effects, such that only feature configurations of disjunction can be computed in case of dependencies to be built across a potential intervener.

Finally, the present study triggered new questions for future research. As for the supposed dissociation between semantic and syntactic information in lexical entries, the hypothesis could gain further support from a study that includes GG retrieval tasks as well as semantic tasks on the same list of words. An item-by-item analysis of results would be very informative on the issue. Moreover, being grammatical gender a highly language-specific system, a cross-linguistic comparison with studies on other languages is highly desirable. The investigation into the impaired computation of specific syntactic configuration should be enlarged in many different directions. For instance, PADs should be tested on the computation of structures characterized by passive voice. Tasks should comprehend more conditions for both Wh-questions and RCs in order to test patients' sensitivity to other syntactic manipulations: mismatch in number, person, animacy, etc. I will also need data from production in order to complete the understanding of the syntactic impairment.

REFERENCES

Adani, F. (2012). Some notes on the Acquisition of Relative Clauses: New Data and Open Questions. *ENJOY LINGUISTICS!Papers offered to Luigi Rizzi on the occasion of his 60th birthday*.Siena: CISCL Press

Adani, F., Shem, M., & Zukowski, A. (2013). How do German children and adults deal with their relatives. *Advances in language acquisition*, 14-22.

Adani, F., Van der Lely, H. K., Forgiarini, M., & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua*, *120*(9), 2148-2166.

Alexiadou, A. (2004). Inflection class, gender and DP internal structure. *Explorations in nominal inflection*, 21-50.

Alexiadou, A., Haegeman, L., & Stavrou, M. (2007). *Noun phrase in the generative perspective* (Vol. 71). Walter de Gruyter.

Alexiadou, A., Law, P., Meinunger, A., & Wilder, C. (Eds.). (2000). *The syntax of relative clauses* (Vol. 32). John Benjamins Publishing.

Almor, A., Aronoff, J. M., MacDonald, M. C., Gonnerman, L. M., Kempler, D., Hintiryan, H., Hayes, H., Arunachalam, S. & Andersen, E. S. (2009). A common mechanism in verb and noun naming deficits in Alzheimer's patients. *Brain and language*, *111*(1), 8-19.

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.

Altmann, L. J., Kempler, D., & Andersen, E. S. (2001). Speech Errors in Alzheimer's DiseaseReevaluating Morphosyntactic Preservation. *Journal of speech, language, and hearing research*, *44*(5), 1069-1082.

Anderson, R. T. (1999). Loss of gender agreement in L1 attrition: Preliminary results. *Bilingual Research Journal*, *23*(4), 389-408.

Andonova, E., D'Amico, S., Devescovi, A., & Bates, E. (2004). Gender and lexical access in Bulgarian. *Perception & psychophysics*, *66*(3), 496-507.

Antinucci, F., & Cinque, G. (1977). Sull'ordine delle parole in italiano: l'emarginazione. *Studi di grammatica italiana VI, pp. 121-146.*

Arnon, I. (2010). Rethinking child difficulty: The effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language*, *37*(01), 27-57.

Astell, A. J., & Harley, T. A. (1996). Tip-of-the-tongue states and lexical access in dementia. *Brain and language*, *54*(2), 196-215.

Avrutin, S. (2000). Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. *Language and the brain: Representation and processing*, 295-313.

Baddeley, A. (1986). Working Memory. Oxford, England: Oxford University Press.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease. *Brain*, *114*(6), 2521-2542.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review*, *105*(1), 158.

Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, *57*(2), 193-216.

Barbosa, P. (2001). On inversion in wh-questions in Romance. *Subject inversion in Romance and the theory of universal grammar*, 20-59.

Bastiaanse, R., Jonkers, R., Ruigendijk, E., & Van Zonneveld, R. (2003). Gender and case in agrammatic production. *Cortex*, *39*(3), 405-417.

Bates, E., Harris, C., Marchman, V., Wulfeck, B., & Kritchevsky, M. (1995). Production of complex syntax in normal ageing and Alzheimer's disease. *Language and Cognitive Processes*, *10*(5), 487-539.

Bayles, K. A. (1982). Language function in senile dementia. *Brain and language*, *16*(2), 265-280.

Bayles, K. A., & Tomoeda, C. K. (1983). Confrontation naming impairment in dementia. *Brain and language*, *19*(1), 98-114.

Belletti, A. (2004). Aspects of the low IP area. *The structure of CP and IP. The cartography of syntactic structures*, *2*, 16-51.

Belletti, A. (2005). Extended doubling and the VP periphery. *Probus*, *17*(1), 1-35.

Belletti, A. (2006). Extending Doubling to Non-Local Domains: Complete vs. Partial Copying+ Deletion. In *Form, Structure, and Grammar: A Festschrift Presented to*

Günther Grewendorf on Occasion of His 60th Birthday (p. 129). Walter de Gruyter GmbH & Co KG.

Belletti, A. (2014). Notes on passive object relatives. *Functional Structure from Top to Toe: The Cartography of Syntactic Structures*, *9*, 97-114.

Belletti, A., & Chesi, C. (2011). Relative clauses from the input: syntactic considerations on a corpusbased analysis of Italian. *Studies in Linguistics (CISCL Working Papers)*, *4*.

Belletti, A., & Contemori, C. (2010). Intervention and attraction. On the production of subject and object relatives by Italian (young) children and adults. In *Language Acquisition and Development. Proceedings of GALA* (pp. 39-52).

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, *122*(10), 1053-1069.

Belletti, A., & Guasti, M.T. (2015). *The acquisition of Italian*. John Benjamins Publishing Company.

Belletti, A., & Leonini, C. (2004). Subject inversion in L2 Italian. *EUROSLA yearbook*, *4*, 95-118.

Belletti, A., & Rizzi, L. (2013). Ways of avoiding intervention: Some thoughts on the development of object relatives, passive and control. *Rich languages from poor inputs*, 115-126.

Belleville, S., Peretz, I., & Malenfant, D. (1996). Examination of the working memory components in normal aging and in dementia of the Alzheimer type. *Neuropsychologia*, *34*(3), 195-207.

Bencini, G. M., Pozzan, L., Biundo, R., McGeown, W. J., Valian, V. V., Venneri, A., & Semenza, C. (2011). Language-specific effects in Alzheimer's disease: Subject omission in Italian and English. *Journal of Neurolinguistics*, *24*(1), 25-40.

Benincà, P. (2010). Headless relatives in some Old Italian varieties. *Synctactic variation. The dialects of Italy. Cambridge, UK: Cambridge University Press.*

Benincà, P. (2012). Lexical complementisers and headless relatives. *Functional Heads. The Cartography of Syntactic Structures*, *7*, 29-41.

Benincà, P., & Poletto, C. (2005). The third dimension of person features. *Syntax and variation: Reconciling the biological and the social*, 265-299.

Bentea, A., Durrleman, S., & Rizzi, L. (2016). Refining intervention: The acquisition of featural relations in object A-bar dependencies. *Lingua*, *169*, 21-41.

Bernstein, J. B. (1993). *Topics in the syntax of nominal structure across Romance*. City University of New York
(06.04.2016 www.ai.mit.edu/projects/dm/theses/more/bernstein93.pdf)

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton, A. M. (2005). Corpus e lessico di frequenza dell'italiano scritto (CoLFIS). *Scuola Normale Superiore di Pisa*.

Bever, T. G. (1974). The ascent of the specious, or there'sa lot we don't know about mirrors. *Explaining linguistic phenomena*, 173-200.

Bianchi, G. (2013). Gender in Italian–German bilinguals: A comparison with German L2 learners of Italian. *Bilingualism: Language and Cognition*, *16*(03), 538-557.

Bianchi, V. (1999). *Consequences of antisymmetry: Headed relative clauses* (Vol. 46). Walter de Gruyter.

Bianchi, V. (2000). The raising analysis of relative clauses: a reply to Borsley. *Linguistic Inquiry*, *31*(1), 123-140.

Bianchi, V. (2002). Headed relative clauses in generative syntax. Part I. *Glot International*, *6*(7), 197-204.

Bianchi, V. (2004). Resumptive relatives and LF chains. *The Cartography of syntactic Structures*, *2*, 76-114.

Bickel, C., Pantel, J., Eysenbach, K., & Schröder, J. (2000). Syntactic comprehension deficits in Alzheimer's disease. *Brain and Language*, *71*(3), 432-448.

Biran, M., & Ruigendijk, E. (2015). Do case and gender information assist sentence comprehension and repetition for German-and Hebrew-speaking children?. *Lingua*, *164*, 215-238.

Boeckx, C. (2003). *Islands and chains: Resumption as stranding* (Vol. 63). John Benjamins Publishing.

Borsley, R. D. (1997). Relative clauses and the theory of phrase structure. *Linguistic Inquiry*, 629-647.

Braak, H., & Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiology of aging*, *18*(4), 351-357.

Braak, H., & Del Tredici, K. (2006). Staging of cortical neurofibrillary inclusions of the Alzheimer's type. In *Alzheimer: 100 Years and Beyond* (pp. 97-106). Springer Berlin Heidelberg.

Bushell, C. M., & Martin, A. (1997). Automatic semantic priming of nouns and verbs in patients with Alzheimer's disease. *Neuropsychologia*, *35*(8), 1059-1067.

Cacciari, C., & Cubelli, R. (2003). The neuropsychology of grammatical gender: An introduction. *Cortex*, *39*(3), 377-382.

Caloi, I. (2013). The comprehension of relative clauses in patients with Alzheimer's disease. *Studies in Linguistics*, 5.

Cameli, L., Phillips, N. A., Kousaie, S., & Panisset, M. (2005). Memory and language in bilingual Alzheimer and Parkinson patients: Insights from verb inflection. In *ISB4: Proceedings of the 4th International Symposium on Bilingualism* (pp. 452-476).

Cantone, K. F., & Müller, N. (2008). Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua*, *118*(6), 810-826.

Caplan, D. (2001). The measurement of chance performance in aphasia, with specific reference to the comprehension of semantically reversible passive sentences: A note on issues raised by Caramazza, Capitani, Rey, and Berndt (2001) and Drai, Grodzinsky, and Zurif (2001). *Brain and Language*, *76*(2), 193-201.

Caramazza, A. (1997). How many levels of processing are there in lexical access?. *Cognitive neuropsychology*, *14*(1), 177-208.

Caramazza, A., Capitani, E., Rey, A., & Berndt, R. S. (2001). Agrammatic Broca's aphasia is not associated with a single pattern of comprehension performance. *Brain and Language*, *76*(2), 158-184.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from thetip-of-the-tongue'phenomenon. *Cognition*, *64*(3), 309-343.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and language*, *3*(4), 572-582.

Cardinaletti, A. (2004). Towards a cartography of subject positions. *The structure of CP and IP: The cartography of syntactic structures*, *2*, 115-165.

Carminati, M. N. (2005). Processing reflexes of the Feature Hierarchy (Person> Number> Gender) and implications for linguistic theory. *Lingua*, *115*(3), 259-285.

Carreiras, M., Duñabeitia, J. A., Vergara, M., De La Cruz-Pavía, I., & Laka, I. (2010). Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition*, *115*(1), 79-92.

Carrillo, M. C., Dean, R. A., Nicolas, F., Miller, D. S., Berman, R., Khachaturian, Z., ... & Knopman, D. (2013). Revisiting the framework of the National Institute on Aging-Alzheimer's Association diagnostic criteria. *Alzheimer's & Dementia*, *9*(5), 594-601.

Carroll, S. (1989). Second-Language Acquisition and the Computational Paradigm. *Language Learning*, *39*(4), 535-594.

Carstens, V. (2000). Concord in minimalist theory. *Linguistic Inquiry*, *31*(2), 319-355.

Castelli, I., Pini, A., Alberoni, M., Liverta-Sempio, O., Baglio, F., Massaro, D., ... & Nemni, R. (2011). Mapping levels of theory of mind in Alzheimer's disease: A preliminary study. *Aging & mental health*, *15*(2), 157-168.

Cheng, L. L. S. (1997). *On the typology of wh-questions*. Taylor & Francis.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. *Brain*, *113*(2), 397-417.

Chertkow, H., Bub, D., & Seidenberg, M. (1989). Priming and semantic memory loss in Alzheimer's disease. *Brain and language*, *36*(3), 420-446.

Chiarelli, V., Menichelli, A., & Semenza, C. (2007). Naming compounds in Alzheimer's disease. *The Mental Lexicon*, *2*(2), 259-269.

Chini, M. (1995). Genere grammaticale e acquisizione. *Aspetti della morfologia nominale in italiano L2*.

Chini, M. (1998). Genuserwerb des Italienischen durch deutsche Lerner. *Eine zweite Sprache lernen: empirische Untersuchungen zum Zweitspracherwerb*, *24*, 39.

Chomsky, N. (1977). On WH-movement. *Formal syntax*, ed. by Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 71-132.

Chomsky, N. (1980). On binding. *Linguistic inquiry*, 1-46.

Chomsky, N. (1981). Lectures on binding and government. *Foris, Dordrecht*.

Cinque, G. (1990). *Types of Ā-dependencies*. MIT press.

Cinque, G. (2008). More on the indefinite character of the head of restrictive relatives. *Rivista di grammatica generativa*, *33*, 3-24.

Cinque, G. (2014). *Typological studies: Word order and relative clauses*. Routledge.

Clifton Jr, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In *Linguistic structure in language processing* (pp. 273-317). Springer Netherlands.

Cole, P., & Hermon, G. (1998). The Typology of Wh-Movement, Wh-Questions in Malay. *Syntax*, *1*(3), 221-258.

Collins, C. (2005). A smuggling approach to the passive in English. *Syntax*, *8*(2), 81-120.

Colombo, L., Fonti, C., & Stracciari, A. (2009). Italian verb inflection in Alzheimer dementia. *Neuropsychologia*, *47*(4), 1069-1078.

Comrie, B. (1999). Grammatical gender systems: a linguist's assessment. *Journal of Psycholinguistic Research*, *28*(5), 457-466.

Contemori, C., & Belletti, A. (2014). Relatives and passive object relatives in Italian-speaking children and adults: Intervention in production and comprehension. *Applied Psycholinguistics*, *35*(06), 1021-1053.

Contemori, C., & Marinis, T. (2014). The impact of number mismatch and passives on the real-time processing of relative clauses. *Journal of child language*, *41*(03), 658-689.

Corbett, G. G. (2006). *Agreement* (Vol. 109). Cambridge University Press.

Costa, J., Friedmann, N., Silva, C., & Yachini, M. (2014). The boy that the chef cooked: Acquisition of PP relatives in European Portuguese and Hebrew. *Lingua*, *150*, 386-409.

Costa, J., Lobo, M., & Silva, C. (2011). Subject–object asymmetries in the acquisition of Portuguese relative clauses: Adults vs. children. *Lingua*, *121*(6), 1083-1100.

Croot, K., Hodges, J. R., Xuereb, J., & Patterson, K. (2000). Phonological and articulatory impairment in Alzheimer's disease: a case series. *Brain and language*, *75*(2), 277-309.

Cross, K., Smith, E. E., & Grossman, M. (2008). Knowledge of natural kinds in semantic dementia and Alzheimer’s disease. *Brain and language*, *105*(1), 32-40.

Cuetos, F., Arce, N., Martínez, C., & Ellis, A. W. (2015). Word recognition in Alzheimer's disease: Effects of semantic degeneration. *Journal of neuropsychology*.

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.

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.

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.

Cummings, J. L., Benson, D. F., Hill, M. A., & Read, S. (1985). Aphasia in dementia of the Alzheimer type. *Neurology*, *35*(3), 394-394.

D'Alessandro, R., & Roberts, I. (2010). Past participle agreement in Abruzzese: split auxiliary selection and the null-subject parameter. *Natural Language & Linguistic Theory*, *28*(1), 41-72.

De Bleser, R., Bayer, J., & Luzzatti, C. (1996). Linguistic theory and morphosyntactic impairments in German and Italian aphasics. *Journal of Neurolinguistics*, *9*(3), 175-185.

De Martino, M., Bracco, G., & Laudanna, A. (2011). The activation of grammatical gender information in processing Italian nouns. *Language and Cognitive Processes*, *26*(4-6), 745-776.

De Vincenzi, M. (1991a). *Syntactic parsing strategies in Italian: The minimal chain principle* (Vol. 12). Springer Science & Business Media.

De Vincenzi, M. (1991b). Filler-gap dependencies in a null subject language: Referential and nonreferential WHs. *Journal of Psycholinguistic Research*, *20*(3), 197-213.

De Vincenzi, M. (1996). Syntactic analysis in sentence comprehension: Effects of dependency types and grammatical constraints. *Journal of Psycholinguistic Research*, *25*(1), 117-133.

De Vincenzi, M., Arduino, L. S., Ciccarelli, L., & Job, R. (1999). Parsing strategies in children comprehension of interrogative sentences. In *Proceedings of European Conference on Cognitive Science. Istituto di Psicologia del CNR, Rome* (pp. 301-308).

De Vincenzi, M., & Di Domenico, E. (1999). A distinction among features: the role of gender and number in the retrieval of pronoun antecendents. *Rivista di lingüística*, *11*(1), 41-74.

De Vries, M. (2002). *The syntax of relativization*. Netherlands Graduate School of Linguistics.

Di Domenico, E. (1997). Per una teoria del genere grammaticale. *Padova, Unipress*.

Diesfeldt, H. F. A. (1985). Verbal fluency in senile dementia: an analysis of search and knowledge. *Archives of gerontology and geriatrics*, *4*(3), 231-239.

Duong, A., Giroux, F., Tardif, A., & Ska, B. (2005). The heterogeneity of picture-supported narratives in Alzheimer's disease. *Brain and Language*, *93*(2), 173-184.

Drai, D., Grodzinsky, Y., & Zurif, E. (2001). Broca's aphasia is associated with a single pattern of comprehension performance: A reply. *Brain and Language*, *76*(2), 185-192.

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.

Forbes-McKay, K. E., & Venneri, A. (2005). Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task. *Neurological sciences*, *26*(4), 243-254.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of verbal learning and verbal behavior*, *22*(2), 203-218.

Fox, B. A., & Thompson, S. A. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language*, 297-316.

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender* (Vol. 38). John Benjamins Publishing.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition*, *101*(1), 173-216.

Franzon, F., Arcara, G., Peressotti, F., & Semenza, C. (2014). Gender agreement: a psycholinguistic and aphasia case study. In *Academy of Aphasia-52nd Annual Meeting. Frontiers in Psychology* (pp. 5-7).

Frazier, L., & Clifton Jr, C. (1989). Successive cyclicity in the grammar and the parser. *Language and cognitive processes*, *4*(2), 93-126.

Frazier, L., & d'Arcais, G. B. F. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of memory and language*, *28*(3), 331-344.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, *6*(4), 291-325.

Friedmann, N. (2002). Question production in agrammatism: The tree pruning hypothesis. *Brain and Language*, *80*(2), 160-187.

Friedmann, N. (2008). Traceless relatives: Agrammatic comprehension of relative clauses with resumptive pronouns. *Journal of Neurolinguistics*, *21*(2), 138-149.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*(1), 67-88.

Friedmann, N., & Biran, M. (2003). When is gender accessed? A study of paraphasias in Hebrew anomia. *Cortex*, *39*(3), 441-463.

Friedmann, N., Biran, M., & Dotan, D. (2013). Lexical retrieval and its breakdown in aphasia and developmental language impairment. *The Cambridge handbook of biolinguistics*, 350-374.

Friedmann, N., & Novogrodsky, R. (2002). BAMBI: Battery for assessment of syntactic abilities in children. *Tel Aviv: Tel Aviv University*.

Friedmann, N., & Szterman, R. (2011). The comprehension and production of wh questions in children with hearing impairment. *Journal of Deaf Studies and Deaf Education*, *16*(2), 212-235.

Fyndanis, V., Manouilidou, C., Koufou, E., Karampekios, S., & Tsapakis, E. M. (2013). Agrammatic patterns in Alzheimer's disease: Evidence from tense, agreement, and aspect. *Aphasiology*, *27*(2), 178-200.

Gainotti, G., Di Betta, A. M., & Silveri, M. C. (1996). The production of specific and generic associates of living and nonliving, high-and low-familiarity stimuli in Alzheimer's disease. *Brain and language*, *54*(2), 262-274.

Garraffa, M., Di Domenico, A., & Di Matteo, R. (2010). The syntactic domain of number agreement: attraction effects and pronominal classes. *Studies in Li*, 108.

Garraffa, M., & Grillo, N. (2008). Canonicity effects as grammatical phenomena. *Journal of Neurolinguistics*, *21*(2), 177-197.

Garrard, P., Maloney, L. M., Hodges, J. R., & Patterson, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain*, *128*(2), 250-260.

Gayraud, F., Lee, H. R., & Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical linguistics & phonetics*, *25*(3), 198-209.

Gehrke, B., & Grillo, N. (2009). How to become passive. *Explorations of phase theory: Features, arguments, and interpretation at the interfaces*, 231-268.

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition*, *111*(1), 1-23.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1-76.

Giusti, G. (1993). *La sintassi dei determinanti*. Padova, Unipress.

Giusti, G. (2009). Linguaggio e questioni di genere: alcune riflessioni introduttive. *Working Papers in Linguistics*, 19. Universitá Ca’ Foscari di Venezia.

Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of psycholinguistic research*, *30*(6), 627-651.

Gómez-Isla, T., Hollister, R., West, H., Mui, S., Growdon, J. H., Petersen, R. C., ... & Hyman, B. T. (1997). Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease. *Annals of neurology*, *41*(1), 17-24.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1411.

Goodglass, H. (2000). Grammatical gender is not always syntactic. In Grodzinsky, Y., Shapiro, L. P., & Swinney, D. (Eds.). *Language and the brain: Representation and processing*. Academic Press.

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S., & Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain*, *125*(4), 752-764.

Grillo, N. (2008). *Generalized minimality: Syntactic underspecification in Broca's aphasia*. LOT.

Grillo, N. (2009). Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism. *Lingua*, *119*(10), 1426-1443.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and language*, *27*(1), 135-159.

Grodzinsky, Y. (1989). Agrammatic comprehension of relative clauses. *Brain and Language*, *37*(3), 480-499.

Grodzinsky, Y. (1990). *Theoretical perspectives on language deficits*. MIT press.

Grossman, M., Mickanin, J., Onishi, K., & Hughes, E. (1996). Verb comprehension deficits in probable Alzheimer's disease. *Brain and Language*, *53*(3), 369-389.

Grossman, M., Murray, R., Koenig, P., Ash, S., Cross, K., Moore, P., & Troiani, V. (2007). Verb acquisition and representation in Alzheimer's disease. *Neuropsychologia*, *45*(11), 2508-2518.

Grossman, M., & Rhee, J. (2001). Cognitive resources during sentence processing in Alzheimer's disease. *Neuropsychologia*, *39*(13), 1419-1431.

Grossman, M., & White-Devine, T. (1998). Sentence comprehension in Alzheimer's disease. *Brain and language*, *62*(2), 186-201.

Grosu, A., & Landman, F. (1998). Strange relatives of the third kind. *Natural Language Semantics*, *6*(2), 125-170.

Guasti, M. T. (1996a). On the controversial status of Romance interrogatives. *Probus*, *8*(2), 161-180.

Guasti, M. T. (1996b). Acquisition of Italian interrogatives. *Language Acquisition and Language Disorders*, *14*, 241-270.

Guasti, M. T., Branchini, C., & Arosio, F. (2012). Agreement in the production of Italian subject and object wh-questions. *Studies in Linguistics*, 6.

Guasti, M. T., & Cardinaletti, A. (2003). Relative clause formation in Romance child's production. *Probus*, *15*(1), 47-89.

Guasti, M. T., & Rizzi, L. (2002). Agreement and tense as distinct syntactic positions: Evidence from acquisition. *The structure of DP and IP: The cartography of syntactic structures*, *1*, 167-194.

Guasti, M. T., Stavrakaki, S., & Arosio, F. (2008). Number and Case in the comprehension of relative clauses: Evidence from Italian and Greek. *Studies in Linguistics*, 102.

Guasti, M. T., Thornton, R., & Wexler, K. (1995). Negation in children's questions: The case of English. In *Proceedings of the 19th Annual Boston University Conference on Language Development* (pp. 228-239).

Haendler, Y., Kliegl, R., & Adani, F. (2015). Discourse accessibility constraints in children's processing of object relative clauses. *Frontiers in psychology*, *6*.

Harley, H., & Ritter, E. (2002). Person and number in pronouns: A feature-geometric analysis. *Language*, *78*(3), 482-526.

Harnish, S. M., & Neils-Strunjas, J. (2008). In search of meaning: reading and writing in Alzheimer's disease. In *Seminars in Speech and Language* (Vol. 29, No. 1, pp. 44-59).

Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. *Language Acquisition and Language Disorders*, *32*, 175-206.

Håkansson, G. (1995). Syntax and morphology in language attrition: A study of five bilingual expatriate Swedes. *International Journal of Applied Linguistics*, *5*(2), 153-169.

Heim, S. (2008). Syntactic gender processing in the human brain: A review and a model. *Brain and language*, *106*(1), 55-64.

Heim, S., Alter, K., & Friederici, A. D. (2005). A dual-route account for access to grammatical gender: Evidence from functional MRI. *Anatomy and embryology*, *210*(5-6), 473-483.

Heim, S., Eickhoff, S. B., Opitz, B., & Friederici, A. D. (2006). BA 44 in Broca's area supports syntactic gender decisions in language production. *Neuroreport*, *17*(11), 1097-1101.

Hernandez, A. E., Kotz, S. A., Hofmann, J., Valentin, V. V., Dapretto, M., & Bookheimer, S. Y. (2004). The neural correlates of grammatical gender decisions in Spanish. *Neuroreport*, *15*(5), 863-866.

Hickok, G., & Avrutin, S. (1996). Comprehension of wh-questions in two Broca's aphasics. *Brain and language*, *52*(2), 314-327.

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.

Hodges, J. R., Salmon, D. P., & Butters, N. (1991). The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain*, *114*(4), 1547-1558.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, *90*(1), 3-27.

Hu, S. (2014). *Intervention effects and the acquisition of relativization and topicalization in Chinese*. PhD Dissertation. Universitat Autònoma de Barcelona/Università degli Studi di Milano-Bicocca. (available at http://filcat.uab.cat/clt/publicacions/tesis/pdf/Shenai_2014.pdf, 08.04.2016)

Hu, S., Gavarró, A., Guasti, M.T. (2016). Children's production of head-final relative clauses: The case of Mandarin. *Applied Psycholinguistics*, 37, 323-346.

Huff, F. J., Corkin, S., & Growdon, J. H. (1986). Semantic impairment and anomia in Alzheimer's disease. *Brain and language*, *28*(2), 235-249.

Hyman, B. T. (2006). Anatomical changes underlying dementia in Alzheimer's disease. In *Alzheimer: 100 Years and Beyond* (pp. 89-94). Springer Berlin Heidelberg.

Hyams, N., & Snyder, W. (2005, November). Young children never smuggle: Reflexive clitics and the universal freezing hypothesis. In *30th Annual Boston University Conference on Language Development* (Vol. 5).

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1), 101-144.

Irigaray, L. (1967). Approche psycho-linguistique du langage des dements. *Neuropsychologia*, *5*(1), 25-52.

Irigaray, L. (1973). *Le langage des déments* (Vol. 24). Walter de Gruyter.

Jenden de López, K. Sundahl Olsen, L., & Chondrogianni, V. (2014). Annoying Danish relatives: Comprehension and production of relative clauses by Danish children with and without SLI. *Journal of child language*, *41*(01), 51-83.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). The Boston naming test, 2nd. *Philadelphia: Lea & Febiger*.

Kavé, G., & Levy, Y. (2003a). Sensitivity to gender, person, and tense inflection by persons with Alzheimer's disease. *Brain and language*, *87*(2), 267-277.

Kavé, G., & Levy, Y. (2003b). Morphology in picture descriptions provided by persons with Alzheimer's disease. *Journal of speech, language, and hearing research*, *46*(2), 341-352.

Kayne, R. S. (1975). *French syntax: The transformational cycle*.

Kayne, R. S. (1991). Romance clitics, verb movement, and PRO. *Linguistic inquiry*, 647-686.

Kayne, R. S. (1994). *The antisymmetry of syntax* (No. 25). Mit Press.

Kayne, R. S. (2014). Why Isn't This a complementizer?. *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic inquiry*, 63-99.

Kemper, S., LaBarge, E., Ferraro, F. R., Cheung, H., Cheung, H., & Storandt, M. (1993). On the preservation of syntax in Alzheimer's disease: Evidence from written sentences. *Archives of Neurology*, *50*(1), 81-86.

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.

Kempler, D., Curtiss, S., & Jackson, C. (1987). Syntactic preservation in Alzheimer's disease. *Journal of Speech, Language, and Hearing Research*, *30*(3), 343-350.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860-897.

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.

Kupisch, T., Müller, N., & Cantone, K. F. (2002). Gender in monolingual and bilingual first language acquisition: Comparing Italian and French. *Lingue e linguaggio*, *1*(1), 107-150.

Laka, I. (2012). Merging from the temporal input: On subject-object asymmetries and an ergative language. *Rich Languages From Poor Inputs*, 127.

Laisney, M., Bon, L., Guiziou, C., Daluzeau, N., Eustache, F., & Desgranges, B. (2011). Theory of mind in Alzheimer's disease. *Alzheimer's & Dementia*, *7*(4), S275.

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

Lin, Y., & Garnsey, S. M. (2010). *Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin* (pp. 241-275). Springer Netherlands.

Longobardi, G. (1994). Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic inquiry*, 609-665.

Luzzatti, C., & De Bleser, R. (1996). Morphological processing in Italian agrammatic speakers: Eight experiments in lexical morphology. *Brain and Language*, *54*(1), 26-74.

Luzzatti, C., Laiacona, M., & Agazzi, D. (2003). Multiple patterns of writing disorders in dementia of the Alzheimer type and their evolution. *Neuropsychologia*, *41*(7), 759-772.

Lyons, K., Kemper, S., LaBarge, E., Ferraro, F. R., Balota, D., & Storandt, M. (1994). Oral language and Alzheimer's disease: A reduction in syntactic complexity. *Aging, Neuropsychology, and Cognition*, *1*(4), 271-281.

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.

MacWhinney, B. (1987). The competition model. *Mechanisms of language acquisition*, 249-308.

MacWhinney, B., & Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, *29*(2), 95-141.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, *47*(1), 50-68.

Manenti, R., Repetto, C., Bentrovato, S., Marcone, A., Bates, E., & Cappa, S. F. (2004). The effects of ageing and Alzheimer's disease on semantic and gender priming. *Brain*, *127*(10), 2299-2306.

Manetti, C., Belletti, A. (2015). Causatives and the Acquisition of the Italian Passive. In: Hamann, C., Ruigendijk, E. (Eds.), *Language acquisition and development: proceedings of GALA 2013*. New Castle upon Tyne: Cambridge Scholars Publishing.

Manouilidou, C., de Almeida, R. G., Schwartz, G., & Nair, N. P. V. (2009). Thematic roles in Alzheimer's disease: Hierarchy violations in psychological predicates. *Journal of Neurolinguistics*, *22*(2), 167-186.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological review*, *101*(1), 3.

Matteini, S. (2014). L2 multiple nominal feature agreement in German: the acquisition of grammatical gender in a formal environment. *International Review of Applied Linguistics in Language Teaching*, *52*(4), 313-355.

Matthews, P. H. (1997). *The concise Oxford dictionary of linguistics*. Oxford University Press.

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: Past studies and new evidence. *Cortex*, *45*(6), 738-758.

McCloskey, J. (2001). The morphosyntax of wh-extraction in Irish. *Journal of Linguistics*, *37*(01), 67-100.

McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. *Derivation and explanation in the minimalist program*, *5*, 184-226.

McKhann, G., 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.

Meinzer, M., Lahiri, A., Flaisch, T., Hannemann, R., & Eulitz, C. (2009). Opaque for the reader but transparent for the brain: neural signatures of morphological complexity. *Neuropsychologia*, *47*(8), 1964-1971.

Miceli, G., Turriziani, P., Caltagirone, C., Capasso, R., Tomaiuolo, F., & Caramazza, A. (2002). The neural correlates of grammatical gender: an fMRI investigation. *Journal of cognitive neuroscience*, *14*(4), 618-628.

Montanes, P, Goldblum, M. C., & Boller, F. (1996). Classification deficits in Alzheimer's disease with special reference to living and nonliving things. *Brain and language*, *54*(2), 335-358.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, *58*(3), 503-553.

Morris, R. G. (1994). Working memory in Alzheimer-type dementia. *Neuropsychology*, *8*(4), 544.

Morris, R. G., & Baddeley, A. D. (1988). Primary and working memory functioning in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology*, *10*(2), 279-296.

Murdoch, B. E., Chenery, H. J., Wilks, V., & Boyle, R. S. (1987). Language disorders in dementia of the Alzheimer type. *Brain and language*, *31*(1), 122-137.

Murdoch, I. (1995). *Jackson's Dilemma*. Chatto & Windus.

Nebes, R. D., & Brady, C. B. (1990). Preserved organization of semantic attributes in Alzheimer's disease. *Psychology and aging*, *5*(4), 574.

Nebes, R. D., & Halligan, E. M. (1996). Sentence context influences the interpretation of word meaning by Alzheimer patients. *Brain and language*, *54*(2), 233-245.

Nebes, R. D., Martin, D. C., & Horn, L. C. (1984). Sparing of semantic memory in Alzheimer's disease. *Journal of abnormal psychology*, *93*(3), 321.

Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech, Language, and Hearing Research*, *28*(3), 405-410.

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.

Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *International Journal of Speech-Language Pathology*, *8*(4), 364-375.

O'Grady, W. (1997). *Syntactic Development*. Chicago: University of Chicago Press.
O'Grady, W., Lee, M., & Choo, M. (2003). A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition*, *25*(03), 433-448.

Padovani, R., & Cacciari, C. (2003). Il ruolo della trasparenza morfologica nel riconoscimento di parole in Italiano. *Giornale italiano di psicologia*, *30*(4), 749-772.

Padovani, R., Calandra-Buonaura, G., Cacciari, C., Benuzzi, F., & Nichelli, P. (2005). Grammatical gender in the brain: evidence from an fMRI study on Italian. *Brain research bulletin*, *65*(4), 301-308.

Paganelli, F., Vigliocco, G., Vinson, D., Siri, S., & Cappa, S. (2003). An investigation of semantic errors in unimpaired and Alzheimer's speakers of Italian. *Cortex*, *39*(3), 419-439.

Patel, P. G., & Satz, P. (1994). The language production system and senile dementia of Alzheimer's type: Neuropathological implications. *Aphasiology*, *8*(1), 1-18.

Patterson, K. E., Graham, N., & Hodges, J. R. (1994). Reading in dementia of the Alzheimer type: A preserved ability?. *Neuropsychology*, *8*(3), 395.

Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. *The representation of (in) definiteness*, *98*, 98-129.

Picallo, M. C. (1991). Nominals and nominalizations in Catalan. *Probus*, *3*(3), 279-316.

Picallo, M. C. (2008). Gender and number in Romance. *Lingue e linguaggio*, *7*(1), 47-66.

Platel, H., Lambert, J., Eustache, F., Cadet, B., Dary, M., Viader, F., & Lechevalier, B. (1993). Characteristics and evolution of writing impairmant in Alzheimer's disease. *Neuropsychologia*, *31*(11), 1147-1158.

Poletto, C. (2000). *The higher functional field*. Oxford University Press.

Poletto, C. (2006). Doubling as economy. *Working Papers in Linguistics*, *16*, 211-235.

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.

Prince, M., Comas-Herrera, A., Knapp, M., Guerchet, M., & Karagiannidou, M. (2016). World Alzheimer Report 2016. Improving healthcare for people living with dementia: coverage, quality and costs now and in the future. *Alzheimer's Disease International, London*.

Ralph, M. A. L., Patterson, K., & Hodges, J. R. (1997). The relationship between naming and semantic knowledge for different categories in dementia of Alzheimer's type. *Neuropsychologia*, *35*(9), 1251-1260.

Rizzi, L. (1982). *Issues in Italian syntax* (Vol. 11). Walter de Gruyter.

Rizzi, L. (1990). *Relativized minimality*. The MIT Press.

Rizzi, L. (1996). Residual verb second and the wh-criterion. *Parameters and functional heads*, 63-90.

Rizzi, L. (2000). *Comparative syntax and language acquisition*. London: Routledge.

Rizzi, L. (2001). On the position "Int (errogative)" in the left periphery of the clause. *Current studies in Italian syntax*, *14*, 267-296.

Rizzi, L. (2004). Locality and left periphery. *Structures and beyond: The cartography of syntactic structures, 3,* 223-251.

Rizzi, L. (2006). On the form of chains: Criterial positions and ECP effects. In Cheng, L. L. S., & Corver, N. (Eds.). *Wh-movement: Moving on* (Vol. 42). The MIT Press.

Rizzi, L., Shlonsky, U. (2007). Strategies of Subject extraction. In: Gärtner, H.-M., Sauerland, U. Eds.), *Interfaces+Recursion= Language? Chomsky's minimalism and the view from syntax-semantics.* Berlin: Mouton de Gruytier, 115-160.

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.

Roesch, A.D., Chondrogianni, V. (2015). The use of case in the comprehension of wh-questions in German-speaking children with and without SLI. In: Hamann, C., Ruigendijk, E., (Eds.), *Language acquisition and Development: Proceedings of GALA 2015*. Cambridge Scholars Publishing.

Rohlfs, G. (1968). *Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia.* Torino: Einaudi.

Sanfelici, E., Poletto, C. (2015). *What's in a C? On relative pronouns and complementizers*. Workshop "Variation in C" October, University of Venice, Ca' Foscari.

Sanfelici, E., Poletto, C. (in press). *On demonstratives as relative pronouns: New insight from Italian varieties*. In Coniglio, Schlachter, Veenstra (eds.). Demonstratives. Berlin: Mouton de Gruyter.

Sanfelici, E., Schulz, P., Trabandt, C., Thiel, A. (2014). *Intervention and its avoidance in children's production of relative clauses*. Ms Goethe Universität Frankfurt.

Scalise, S. (1994). *Le strutture del linguaggio. Morfologia*. Soc. Ed. Il Mulino.

Schachter, P. (1973). Focus and relativization. *Language*, 19-46.

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in the processing of locally ambiguous wh-questions in German. In *German sentence processing* (pp. 65-93). Springer Netherlands.

Schulz, P. (2013). Wer versteht wann was? Sprachverstehen im frühen Zweitspracherwerb des Deutschen am Beispiel der w-Fragen. In Deppermann, A. (Ed.), *Das Deutsch der Migranten*. Jahrbuch 2012 des Instituts für deutsche Sprache.

Schwartz, M. F., Marin, O. S., & Saffran, E. M. (1979). Dissociations of language function in dementia: A case study. *Brain and language*, *7*(3), 277-306.

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.

Small, J. A., Kemper, S., & Lyons, K. (2000). Sentence repetition and processing resources in Alzheimer's disease. *Brain and Language*, *75*(2), 232-258.

Small, J. A., & Sandhu, N. (2008). Episodic and semantic memory influences on picture naming in Alzheimer's disease. *Brain and language*, *104*(1), 1-9.

Snyder, W., & Hyams, N. (2015). Minimality effects in children's passives. In: Elisa Di Domenico, Cornelia Hamann and SimonaMatteini (Eds.). *Structures, Strategies and Beyond: Studies in honour of Adriana Belletti*. John Benjamins. (343-368).

Spinnler, H. (1996). Le demenze "corticali", in particolare la malattia di Alzheimer. In Denes, G., & Pizzamiglio, L., *Manuale di neuropsicologia*. Zanichelli, Bologna.

Stromswold, K. (1995). The acquisition of subject and object wh-questions. *Language Acquisition*, *4*(1-2), 5-48.

Tait, M. E., C. K., & Ballard, K. J. (1995). Subject-object asymmetries in agrammatic comprehension of four types of wh-questions. *Brain and Language*, *51*, 77-79.

Thompson, C. K., Tait, M. E., Ballard, K. J., & Fix, S. C. (1999). Agrammatic Aphasic Subjects' Comprehension of Subject and Object ExtractedWhQuestions. *Brain and Language*, *67*(3), 169-187.

Thornton, R. J. (1990). Adventures in long-distance moving: The acquisition of complex wh-questions.

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.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, *47*(1), 69-90.

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature reviews neuroscience*, *2*(10), 717-726.

Uriagereka, J. (1995). Aspects of the syntax of clitic placement in Western Romance. *Linguistic inquiry*, 79-123.

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.

Vergnaud, J. R. (1974). *French relative clauses* (Doctoral dissertation, Massachusetts Institute of technology).

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological science*, *8*(4), 314-317.

Walenski, M., Sosta, K., Cappa, S., & Ullman, M. T. (2009). Deficits on irregular verbal morphology in Italian-speaking Alzheimer's disease patients. *Neuropsychologia*, *47*(5), 1245-1255.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79-112.

Warrington, E. K. (1975). The selective impairment of semantic memory. *The Quarterly journal of experimental psychology*, *27*(4), 635-657.

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.

Whitaker, H. (1976). A case of the isolation of the Language Function. *Studies in neurolinguistics*, *2*, 1.

APPENDIX

Task 1: Grammatical Gender Retrieval

Class	GG	NG	Marker	Nouns
I transparent	masculine	masculine	*-o*	*marito* *maschio* *diavolo* *toro* *genero*
I transparent	masculine	inanimate	*-o*	*libro* *cervello* *buco* *fosso* *detersivo*
II transparent	feminine	feminine	*-o*	*mamma* *femmina* *suora* *strega* *nuora*
II transparent	feminine	inanimate	*-a*	*musica* *valigia* *gonna* *torta* *carezza*
III opaque	masculine	masculine	*-e*	*padre* *militare* *prete* *conte* *frate*
III opaque	masculine	inanimate	*-e*	*cuore* *ponte* *bastone* *limone* *sedile*

Class	**GG**	**NG**	**Marker**	**Nouns**
III opaque	feminine	feminine	*-e*	*madre* *moglie* *vergine* *comare* *nubile*
III opaque	feminine	inanimate	*-e*	*legge* *voce* *cenere* *cambiale* *peste*
V irregular	masculine	masculine	*-a*	*papa* *pilota* *poeta* *duca* *profeta*
V irregular	masculine	inanimate	*-a*	*clima* *dramma* *pianeta* *panorama* *diploma*
II transparent	feminine	masculine	*-a*	*guardia* *sentinella* *vedetta* *recluta* *spia*
VII irregular	feminine	inanimate	*-a*	*mano* *radio* *dinamo* *foto* *moto*

Class	GG	NG	Marker	Nouns
III opaque (derived)	masculine	masculine	*-e*	*portiere* *cameriere* *pompiere* *barbiere* *giardiniere*
III opaque (derived)	masculine	inanimate	*-e*	*motore* *spessore* *malore* *contatore* *bollitore*
III opaque (derived)	masculine	Masculine	*-e*	*direttore* *genitore* *dottore* *produttore* *creatore*
III opaque (derived)	masculine	inanimate	*-e*	*quartiere* *cantiere* *mestiere* *braciere* *pallottoliere*
III opaque (derived)	feminine	feminine	*-e*	*scrittrice* *direttrice* *lavoratrice* *collaboratrice* *levatrice* *pittrice* *cooperatrice* *produttrice* *cantautrice* *sostenitrice*
III opaque (derived)	feminine	inanimate	*-e*	*stupidaggine* *testardaggine* *goffaggine* *sfacciataggine* *sbadataggine*
III opaque (derived)	feminine	inanimate	*-e*	*solitudine* *latitudine* *moltitudine* *vicissitudine* *similitudine*

TASK 2: Wh-question comprehension

WhoS

1. Chi mi sta fotografando?
   who I-ACC.1SG AUX-3PS photograph-PROG
   'Who is photographing me?'
2. Chi mi sta bagnando?
   'Who is wetting me?'
3. Chi mi sta baciando?
   'Who is kissing me? '
4. Chi mi sta coprendo?
   'Who is covering me? '
5. Chi mi sta spingendo?
   'Who is pushing me?
6. Chi mi sta asciugando?
   'Who is drying me? '
7. Chi mi sta salutando?
   'Who is greeting me? '
8. Chi mi sta pettinando?
   'Who is combing me? '

WhoO

9. Chi sto fotografando?
   Who AUX-1PS photograph-PROG
   'Who am I photographing?'
10. Chi sto bagnando?
    'Who am I wetting? '
11. Chi sto baciando?
    'Who am I kissing? '
12. Chi sto coprendo?
    'Who am I covering? '
13. Chi sto spingendo?
    'Who am I pushing? '

14. Chi sto asciugando?
'Who am I drying? '

15. Chi sto salutando?
'Who am I greeting? '

16. Chi sto pettinando?
'Who am I combing? '

WhichS

17. Quale dottore mi sta fotografando?
Which doctor I-ACC.1SG AUX-3PS photograph-PROG
'Which doctor is photographing me?'

18. Quale vigile mi sta bagnando?
'Which policeman is wetting me? '

19. Quale sposa mi sta baciando?
'Which bride is kissing me? '

20. Quale suora mi sta coprendo?
'Which nun is covering me? '

21. Quale bambino mi sta spingendo?
'Which boy is pushing me? '

22. Quale bambina mi sta asciugando?
'Which girl is drying me? '

23. Quale prete mi sta salutando?
'Which priest is greeting me? '

24. Quale infermiera mi sta pettinando?
'Which nurse is combing me? '

WhichO

25. Quale dottore sto fotografando?
Which doctor AUX-1PS photograph-PROG
'Which doctor am I photographing?'

26. Quale vigile sto bagnando?
'Which policeman am I wetting? '

27. Quale sposa sto baciando?
'Which bride am I kissing? '

28. Quale suora sto coprendo?
'Which nun am I covering? '

29. Quale bambino sto spingendo?
'Which child am I pushing? '

30. Quale bambina sto asciugando?
'Which child am I drying? '

31. Quale prete sto salutando?
'Which priest am I greeting? '

32. Quale infermiera sto pettinando?
'Which nurse am I combing? '

TASK 3: Relative clause comprehension

SVO

1. Il nonno bacia il bambino
   The granfather.SG kiss-PRS.3SG the child.SG
   'The grandfather kisses the child'
2. Il bambino bagna il papà
   'The child wets the father'
3. Il bambino pettina il re
   'The child combs the king'
4. La nonna copre la bambina
   'The grandmother covers the child'
5. La mamma fotografa la bambina
   'The mother photographs the child'
6. Il dottore disegna il soldato
   'The doctor draws the soldier'
7. La bambina disegna la mamma
   'The child draws the mother'
8. Il cane morde il gatto
   'The dog bites the cat'

SR

9. Mostra-mi il bambino che bacia il nonno
   Show-IMP.2SG I-OBL.1SG the child.SG Rel kiss-PRS.3SG the grandfather.SG
   'Show me the child that kisses the grandfather'
10. Mostrami il nipote che tira lo zio
    'Show me the nephew that pulls the uncle'
11. Mostrami il bambino che spinge il cane
    'Show me the child that pushes the dog'
12. Mostrami la bambina che asciuga la mamma
    'Show me the child that dries the mother'
13. Mostrami la bambina che bacia la nonna
    'Show me the child that kisses the grandmother'

14. Mostrami il soldato che disegna il dottore
'Show me the soldier that draws the doctor'

15. Mostrami la mamma che disegna la bambina
'Show me the mother that draws the child'

16. Mostrami il gatto che morde il cane
'Show me the cat that bites the dog'

OR

17. Mostrami il bambino che il nonno bacia
Show-IMP.2SG I-OBL.1SG the child.SG Rel the grandfather.SG kiss-PRS.3SG
'Show me the child that the grandfather kisses'

18. Mostrami il papà che il bambino bagna
'Show me the father that the child wets'

19. Mostrami il re che il bambino pettina
'Show me the king that the child combs'

20. Mostrami il gatto che il cane morde
'Show me the cat that the dog bites'

21. Mostrami la bambina che la nonna copre
'Show me the child that the grandmother covers'

22. Mostrami la bambina che la mamma fotografa
'Show me the child that the mother photographs'

23. Mostrami la mamma che la bambina disegna
'Show me the mother that the child draws'

24. Mostrami il soldato che il dottore disegna
'Show me the soldier that the doctor draws'

POR

25. Mostrami la bambina che è baciata dalla nonna
Show-IMP.2SG I-OBL.1SG the child.SG Rel AUX kiss-PTCP by-the grandmother.SG
'Show me the child that is kissed by the grandmother'

26. Mostrami il bambino che è bagnato dal papà
'Show me the child that is wetted by the father'

27. Mostrami lo zio che è tirato dal nipote
'Show me the uncle that is pulled by the nephew'

28. Mostrami il bambino che è pettinato dal re
'Show me the child that is combed by the king'

29. Mostrami il bambino che è spinto dal cane
'Show me the child that is pushed by the dog'

30. Mostrami la mamma che è asciugata dalla bambina
'Show me the mother that is dried by the child'

31. Mostrami la nonna che è coperta dalla bambina
'Show me the grandmother that is covered by the child'

32. Mostrami la mamma che è fotografata dalla bambina
'Show me the mother that is photographed by the child'

ClOR

33. Mostrami la nonna che la bambina la bacia
Show-IMP.2SG I-OBL.1SG the grandmother.SG Rel the child.SG she-ACC kiss-PRS.3SG
'Show me the grandmother that the child kisses (her)'

34. Mostrami il bambino che il papà lo bagna
'Show me the child that the father wets (him)'

35. Mostrami il nipote che lo zio lo tira
'Show me the nephew that the uncle pulls (him)'

36. Mostrami il bambino che il re lo pettina
'Show me the child that the king combs (him)'

37. Mostrami il cane che il bambino lo spinge
'Show me the dog that the child pushes (it)'

38. Mostrami la bambina che la mamma la asciuga
'Show me the child that the mother dries (her)'

39. Mostrami la nonna che la bambina la copre
'Show me the grandmother that the child covers (her)'

40. Mostrami la mamma che la bambina la fotografa
'Show me the mother that the child photographs (her)'

ORdem

41. Mostrami quella che la nonna bacia
Show-IMP.2SG I-OBL.1SG that-F Rel the grandmother.SG kiss-PRS.3SG

'Show me the one that the grandmother kisses'

42. Mostrami quello che lo zio tira
'Show me the one that the uncles pulls'

43. Mostrami quello che il cane spinge
'Show me the one that the dog pushes'

44. Mostrami quella che la mamma asciuga
'Show me the one that the mother dries'

45. Mostrami quello che il soldato disegna
'Show me the one that the soldier draws'

46. Mostrami quella che la mamma disegna
'Show me the one that the mother draws'

47. Mostrami quello che il bambino bacia
'Show me the one that the child kisses'

48. Mostrami quello che il gatto morde
'Show me the one that the cat bites'