UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

The performance of Maltese-speaking adults on the Boston Naming Test

A thesis submitted to the University of Limerick in fulfilment of the requirements for the Degree of Doctor of Philosophy

Rita Anna Grima

Department of Speech and Language Therapy

Faculty of Education and Health Sciences

September 2011

Supervisor: **Professor Sue Franklin**

The performance of Maltese-speaking adults on the Boston Naming Test
Rita Anna Grima

ABSTRACT

The Boston Naming Test (Kaplan et al 1983) is standardized on different populations and it is the most widely used naming test worldwide in both research and clinical settings. For this reason, this test was used to investigate naming performance in Maltese-speaking adults in a bilingual context.

The first part of the research aimed to determine the demographic and word variables that affect naming performance in healthy Maltese people. It also aimed to propose an adaptation of the BNT to suit the linguistic and cultural characteristics of the Maltese population. The second part of the research, then, applied the findings obtained initially to establish whether the proposed test could be used clinically in assessing Maltese people with naming impairments. It sought to establish whether the proposed modified test was able to differentiate between people with impaired and unimpaired naming, and whether it could contribute to making a differential diagnosis.

In order to obtain a measure of name agreement the BNT was administered to 24 individuals prior to the main study. Naming performance data was then collected from 60 healthy people. The BNT pictures were presented via a laptop and naming accuracy and reaction times were recorded. Analysis of the variables which affect naming performance was carried out. Finally, the BNT and additional tests of comprehension and repetition were administered on 18 people with aphasia to determine the suitability of the test for clinical use.

Only 38 items of the BNT had at least 70% name agreement. Therefore, a reduced test of 38 items was proposed and two alternative scoring methods were described: strict (accepting only dominant responses and their cognates) and lenient scoring (accepting all possible responses, including English equivalents).

The research on healthy Maltese adults showed that education was the strongest predictor of naming accuracy. Age also remained significant in a stepwise regression analysis. The strongest predictor of naming latency was age, with evidence of a marked increase in reaction times after 66 years of age. Since older Maltese adults had fewer opportunities for education, cohort effects are likely.

Some of the controls obtained very low scores which overlapped with the scores obtained by the people with aphasia. This indicated that scores alone cannot be used to differentiate between impaired and unimpaired people. On the other hand, some types of errors were only produced by people with impaired naming, and did not appear at all in error profiles of unimpaired individuals. Although particular types of errors were associated with particular levels of breakdown in spoken word production, a degree of overlap between errors and levels of breakdown became apparent.

A reduced set of the BNT with lenient scoring was found to be the most suitable test for assessing Maltese people at risk of naming impairments. It is recommended that test standardization in the elderly should consider cohort effects and regular test updates should be carried out.

In order to differentiate between people with impaired and unimpaired naming it is necessary to look at error profiles, apart from the number of errors, as the presence of atypical errors is the best indicator of naming impairments. Further, it was concluded that, although error profiles may prove to be valuable in differential diagnosis, they cannot be used on their own, as they do not provide sufficient information. A variety of additional tasks of semantic and phonological integrity must also be utilized to make conclusive decisions regarding differential diagnosis.

DECLARATION

The performance of Maltese-speaking adults on the Boston Naming Test

"I hereby declare that this is entirely my own work and has not been submitted to any other university or higher education institution, or for any academic award in this University. Where use has been made of the work of other people, it has been fully acknowledged and referenced."

Signed ______________

Date ______________

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to all those who, in some way, contributed to the fulfilment of this work.

First and foremost I thank you Professor Sue Franklin for your guidance, help and encouragement throughout these past few years.

A special thank you goes to all the 142 participants who volunteered and accepted to give some of their precious time to meet me during the data collection. Thank you also goes to a long list of people who gave me permission to contact participants and to use their premises for collecting data.

To Dr Liberato Camilleri, from the Department of Statistics and Operations Research at the University of Malta, for teaching me all I know about Statistical analysis.

To the translator Kenneth Grima who diligently researched every word to come up with all the possible responses and for answering all my questions.

To Stefan Scerri for his help in using Microsoft Office Excel.

To Myriam Galea for acting as inter-rater during the second part of the study.

To the ITSS staff at University of Limerick, especially Mai McGurk, for answering all my emails and for offering their technical assistance whenever I needed their help.

To Brian Jarvis for setting up the DmDx Voice Reaction Timer especially for me.

To all those researchers who responded to my emails and sent me articles and other information I requested.

To Dr Helen Grech for introducing me to my supervisor and for her encouragement.

To my funders, Malta Government Scholarship Scheme (MGSS grant MST 2008-20), and University of Limerick EU Fee Waiver for the Academic Year 2010/11. Without this financial assistance I could never have done this work.

To the publishers of the Boston Naming Test, Pro-Ed. Inc., for kindly providing me with a copy of the test for research purposes.

To all my friends and family for their continuous support, patience and encouragement, especially when accumulation of lack of sleep affected my mood.

To mum and dad. Thank you for your help with the kids whenever I was away in Ireland.

And last, but not least, to my husband, Kenneth, and my children, Matteo, Danjel and Katrina. Thank you Kenneth for your never-ending emotional and practical support. Thank you Matt, Dan and Kat for your patience whenever I told you to settle down and go to bed so that I could sit at my computer ….. .

To my family

Kenneth, Matteo, Danjel and Katrina

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AoA	Age of acquisition
BNT	Boston Naming Test (Kaplan et al 1983)
Cond d'appr	Conduite d'approche
IMG	Imageability
LEST	Language of the Elderly Screening Test (Delia et al, unpublished)
MBNT	Maltese Boston Naming Test
MW	Multiword
No Resp	No response
PA	Phonological assembly
Phon	Phonological
POL	Phonological output lexicon
PWA	People with aphasia
PwithoutA	People without aphasia
RIA	Restricted interaction account (Goldrick and Rapp 2002)
RT	Reaction time
Sem	Semantic
SLP	Speech Language Pathologist
TOT	Tip of the tongue

Chapter 1

Introduction

1.0 INTRODUCTION

The ability to retrieve words and to produce them are essential components for effective communication to take place (Gollan et al 2007). Without appropriate naming or lexical retrieval skills one encounters difficulties in talking fluently and conversing with others. In fact, a quarter of our adult vocabulary consists of names for objects (Levelt et al 1999). Therefore, any difficulty with this seemingly simple, yet complex skill, has the potential to disrupt the flow of communication and, hence, disrupt a conversation (Herbert et al 2008).

A difficulty with lexical retrieval is a phenomenon that commonly occurs in aphasia following stroke and other neurological conditions (Grossman et al 2004; Kavé 2005, Herbert et al 2008). It has also been shown time and time again that lexical retrieval difficulties increase with increasing age even in the absence of a specific neurological disorder (Worrall et al 1995; Mariën et al 1998; Connor et al 2004; Kavé 2005).

1.1 CONFRONTATIONAL NAMING

In research and clinical settings the tasks that are generally used to acquire information on lexical retrieval include: confrontation/picture naming, word repetition and word-picture matching. Picture naming is frequently used to study cognition, especially language and semantic memory (Mathuranath 2007) and it is very often used to make decisions about the neuropsychological status of individuals (Randolph et al 1999; Gollan et al 2007). Further, naming tests are said to be a reflection of how the human mind stores knowledge (Figueredo Balthazar et al 2008) and they are the most commonly used task in investigating lexical retrieval skills (Herbert et al 2008; Goral et al 2007).

Naming difficulties tend to increase with increasing age, with decline generally reported after 70 years of age, especially for individuals with less education (e.g. Neils et al 1995). They are also present in individuals with neurodegenerative conditions (Grossman et al 2004). Analysis of the type of retrieval difficulties can help in differentiating between cases of dementia and normal ageing processes and between different kinds of dementias (Cummings and Benson, 1984, 1992 cited in Allegri et al 1997; Williams et al 2007). Further, picture naming has the potential to identify the type and severity of one's disability as well as provide information on the possible effect on one's participation in conversation (Herbert et al 2008). However, it is essential to emphasise that an individual's performance on a picture naming test may be influenced by a number of factors other than his/her age and neuropsychological state. There are a number of variables that may affect the results of a picture-naming task.

1.2 VARIABLES THAT MAY AFFECT PICTURE NAMING

Ageing, level of education, gender, language, culture and bilingualism have all been studied in relation to their potential effects on lexical access abilities, and particularly on picture-naming skills. These variables, which will be referred to as *demographic variables,* reflect the specific characteristics of each individual and the effect that they may or may not have on the individual's ability to retrieve words.

Other variables must also be taken into account, however. These are variables that arise specifically from the material being used (the stimulus-related variables or *word variables*) and include word frequency, the age of acquisition, imageability, familiarity, word length and degree of concreteness of the particular words or pictures being used (D'Amico et al 2001; Alario et al 2004). Thus, the set of pictures or items used in assessment procedures or research studies must also be taken into account when analysing performance on picture-naming tasks. Culture plays an important role here, as different cultures use different food, clothing, tools, life style, religious practices and language – all of which may determine the extent to which one is familiar with a particular word, when it was learned, how often it is used, what it is called and which image is conjured up in one's mind when the word is heard (Mathuranath 2007).

Research has focused on the possible effects of each of the above variables on naming accuracy and on naming latency, as well as on the correlations that may exist between the different variables themselves.

1.3 THE PRESENT STUDY

The present study focuses on the performance of Maltese-speaking adults (aged between 20 and 85 years) on a confrontational naming task, namely the Boston Naming Test (Kaplan et al 1983). The research is divided into two parts: the first part focuses on data collected from 60 people without aphasia, and the second part focuses on research carried out on 18 people with aphasia. The effect of demographic and word variables on naming performance (naming accuracy and naming latency) will be investigated during the first part of the research. Preliminary normative data will be produced from the results obtained from the people without difficulties. The data collected from the people with aphasia will then be utilised to determine whether the BNT is an appropriate tool for the assessment of word finding skills in Maltese adults.

The Boston Naming Test (BNT) was chosen for the purpose of this study as it is the most widely used naming test (Allegri et al 1997; Kavé 2005). Although it was first introduced in 1983, the test is still considered to be a highly sensitive tool if it is standardised on the particular population and it is still widely used in both clinical and research settings (Tallberg 2005).

1.4 A NAMING TEST FOR MALTESE-SPEAKING ADULTS

In Malta, neuropsychological assessments that are standardised and normed on the Maltese population are still very scarce. This poses a problem for clinicians, such as Speech and Language Pathologists, who require accurate information regarding the true skills of their clients. This information may only be derived from suitable and reliable assessment measures. Therefore, this study should shed light on the performance of

Maltese speakers of different ages and educational levels on the Boston Naming Test, as well as provide information regarding the effects, if any, of several variables on their performance. Furthermore, it will help in identifying the type of material that would be appropriate for assessing Maltese-speakers with word-retrieval deficits, and hence, lead to the development of culturally and linguistically appropriate assessment material, with preliminary norms for the Maltese population.

A naming test may be a basic instrument to be used in clinical settings, as it is easy and quick to administer, and because it taps on a variety of linguistic and cognitive processes (Tallberg 2005; Gollan et al 2007) – thus, providing valuable information for the clinician working with adults with suspected lexical retrieval difficulties. Considering the widespread use of the Boston Naming Test in a variety of linguistic and cultural populations and considering the limited standardised resources we possess for clinical and research purposes locally, it follows that a study of the performance of Maltese-speaking adults on the BNT will provide essential information for producing a valuable research and clinical tool. It has been emphasised that the materials used must always be appropriate for the particular populations of different cultures, ages, etc. (Yoon et al 2006). Ultimately, it is hoped that a standardised naming test for Maltese-speaking adults will be produced following the information that emerges from the present study.

Chapter 2

Literature Review

2.0 INTRODUCTION TO LEXICAL ACCESS

The production of words involves the efficient access of those particular words from a mental store of 50,000 – 100,000 words, at a rate of two to four words per second (Levelt 2001; Nickels 2001). Considering the fact that speaking often takes place while carrying out other activities simultaneously, the skill of lexical retrieval can be viewed as an amazing and highly efficient process, which enables individuals to make their needs, ideas and dreams known to others, and in turn, leads them to form relationships and to socialise with whomever they desire. Thus, it is no wonder that any difficulty with the ability to speak is of great concern for individuals affected by conditions such as aphasia (Nickels 2002; Herbert et al 2008).

The study of normal and impaired lexical retrieval is generally understood within a framework that relies on the cognitive neuropsychological model (Wilshire and Coslett 2000). This approach has evolved (over the past 25 years or so, since 1984) from the traditional *box and arrow* diagrams to finer grained theories that centre around spreading activation theories of lexical access. Cognitive neuropsychology provides the researcher and/or the clinician with an excellent assessment tool as it enables us to distinguish between various stages of spoken word production (Howard 2000). It allows us to identify the processes which are intact, the processes which are impaired, and finally it allows us to describe how these processes interact together to produce the particular difficulties and abilities (Whitworth et al 2005). However, different researchers provide different and conflicting views regarding how activation of words takes place and how words are organized in our mental lexicon, for example. Most of the research has focused on the production of single words (e.g. Levelt et al 1999). However, in recent years, attention has also been given to the retrieval and production of multi-word utterances (e.g. Alario et al 2002; Herbert et al 2008).

This study focuses on single word production and, specifically, on naming words associated with a pictorial representation. The Boston Naming Test (Kaplan et al 1983) will be used throughout this research. The first part of the study will look into naming ability in healthy Maltese speakers aged between 20 and 85 years, and will investigate how this skill changes with age. The influence of demographic and word variables on naming ability will be analyzed and preliminary normative data for Maltese speakers will be produced. The second part of the study will investigate naming in Maltese people with aphasia and the factors which may affect it.

The following sections describe the processes involved in picture naming, as well as the theories that have been proposed to explain single word retrieval.

2.1 PROCESSES INVOLVED IN PICTURE NAMING

> Articulating the name that corresponds to the picture of a common object is a fast, efficient and relatively effortless cognitive skill.
>
> (Alario et al 2004, p.140)

Although picture naming appears to be such a simple task, it is a complex skill that involves different, but interacting mental representations and cognitive processes that must occur rapidly and efficiently (D'Amico et al 2001; Tsang and Lee 2003; Grossman et al 2004; Alario et al 2004; Gollan et al 2007; De Leon et al 2007). When naming a picture:

(i) the visual stimulus must be visually perceived and recognised;
(ii) the meaning (semantic concept) that corresponds to the picture must be identified and accessed;
(iii) the phonological word that corresponds to the picture (and the meaning) must be retrieved;
(iv) articulation must be programmed and planned; and
(v) the movement sequences involved in articulation must be executed.

Every one of these distinct processes must function effectively as any difficulty with any level or with the interaction between these levels will affect the production of the particular word (De Leon et al 2007).

Naming involves a complex network of brain regions in the left frontal, temporal and parietal cortex (Mesulam 2008). Thus, it is no wonder that brain injuries and cerebral infarcts affecting these areas will disrupt the normal processes involved in language, including the ability to produce words. In fact, word retrieval is probably the most common language difficulty faced by people with aphasia (e.g. Grossman et al 2004; Herbert et al 2008).

Different types of naming difficulties will be manifested in different individuals with aphasia, with varying levels of breakdown even in individuals who may appear to have similar aphasia syndromes (Howard and Gatehouse 2006). These individual variations in naming abilities are possibly the result of a concoction of individual features including the site of the neurological lesion, the time post onset of the neurological damage, individual cognitive processes (Budd et al 2010) and premorbid abilities. This highlights the importance of carrying out in-depth assessments on each individual in order to determine the exact level of breakdown in the process of naming (Howard and Gatehouse 2006). The section that follows describes the stages involved in word production and the factors that have been found to affect it in both healthy individuals and in people with aphasia.

2.2 THEORY OF LEXICAL ACCESS

2.2.1 The basic theory of lexical access

In general, most models agree on three levels in spoken word production: a lexical semantic, a lexical phonological and a segmental (phonological) level (Knobel et al 2008). Nickels (2001) explains that a conceptual semantic level precedes the lexical semantic level. While the former is concerned with non linguistic reasoning and conceptual knowledge, the lexical semantic level relies on linguistic aspects of meaning.

Some models of spoken word production (such as Butterworth, 1989, 1992, cited in Nickels 2001) omit the lexical phonological level and include a semantic stage and a phonological stage only, such that the phonological form is directly accessed from the semantic level. This type of model is unable to account for the performance of individuals who have intact semantic and phonological skills, despite their naming difficulties (Nickels 2001). Levelt et al (1999) refers to this level between semantics and phonology as the lemma level, allowing access to unimpaired phonological representations from unimpaired semantic representations.

A set of phonological processing stages follow the lexical phonological stage of word production (Rapp and Goldrick 2006). These post lexical stages include phonological encoding and buffering, articulatory planning and motoric execution. Therefore, the cognitive neuropsychological approach provides a framework for understanding the stages involved in lexical access.

Howard and Gatehouse (2006) state that word-finding errors in people with aphasia may arise from impairment to the semantic representations (lexical semantics level), from impairment in accessing the word from the phonological output lexicon, or from impairment to the actual lexical representation (phonological output lexicon). Whitworth et al (2005) and Howard and Gatehouse (2006) emphasise three important sources of information during the assessment process that will enable the clinician to identify the level of breakdown in terms of an information processing model of cognitive neuropsychology. The three sources of information include: (i) information from the effect of different word variables; (ii) information from the nature of the errors; and (iii) information from performance on different tasks that reflect common processing components (e.g. tasks of auditory and written comprehension), as well as investigating the effect of phonemic cues and miscues. The sections below break down lexical access into five essential components.

2.2.1.1 Conceptual level (object concepts)

The conceptual level (or semantic memory) holds one's knowledge about the world, including knowledge of concepts and words (Allegri et al 1997; Figueredo Balthazar et al 2008; Kuipers & La Heij 2009). It is concerned with *preverbal* concepts and involves *non-linguistic* reasoning abilities. Impairment at this level is generally associated with the type of difficulties evident in dementia (Nickels 2001), and includes poor word comprehension and semantic errors in naming, as well as difficulties with picture semantics judgements. Word frequency and familiarity mostly affect naming accuracy at this level (Howard 2000).

2.2.1.2 Lexical semantics level

The lexical semantics level constitutes a store of word meanings. During this stage the appropriate item must be selected and activated from the mental lexicon. This means that the concept must then be converted into a lexical representation, which is sometimes referred to as lemma retrieval (Levelt 1989). Levelt (1999) later changed the definition of lemma, to refer to syntactic properties only and removed its semantic functions.

In people with aphasia, difficulties at this level may involve poor word comprehension, as well as omissions and semantic errors in naming, as found in impairment at the conceptual semantics level. Picture semantics are not affected, however. There may also be evidence of an effect of imageability and age of acquisition on naming ability, implying that words which have a high imageability rating and/or words which are learned earlier in life are easier to name than those with low imageability and those which are learned later in life, respectively (Howard 2000; Whitworth et al 2005; Howard and Gatehouse 2006).

Semantic errors at this level of spoken word production arise as a result of conflict between semantically related competitors. Semantic competitors which are higher in frequency than the target name tend to be activated more, resulting in a semantic error. When there are fewer semantic competitors to choose from, errors tend to be omissions (Bormann et al 2008).

The result of providing phonemic cues to people with aphasia is a good indicator of the level of impairment in lexical retrieval. In fact, the effect of providing phonemic *miscues* (i.e. providing the phonemic cue of a semantically related word) is particularly important in determining whether this level is impaired or not. While phonemic *cues* may sometimes help elicit the correct response if there is a difficulty with lexical semantics, phonemic *miscues* usually elicit semantic associations. For example, the phonemic miscue ‘t’ for ‘lion' may elicit the word ‘tiger’ (ibid). This is presumably because auditory comprehension is also impaired and the person may not know that the miscued item is incorrect.

Word comprehension has been shown to be impaired if there is breakdown at this level (see above), although it is worth noting that individuals may perform well on tasks of spoken word to picture matching[1] (Howard and Gatehouse 2006; Nickels 2004), despite having difficulties at the semantic level. Nickels (2004) argues that spoken word production entails more complex semantics than that assessed in a simple word-picture matching task, and other tasks related to auditory comprehension and semantics should be used for in-depth assessment (e.g. synonym judgement tasks).

2.2.1.3 Lexical phonological level (phonological output lexicon)

The lexical phonological level includes a store of spoken word forms and it provides access to them. An impairment at this level may be demonstrated by tip of the tongue states, circumlocutions, phonological errors and also semantic errors. However, spoken and written comprehension remains intact (Howard 2000; Nickels 2000; Whitworth et al 2005). Effects of word frequency may be evident, while an imageability effect may also arise if there is involvement of access to the semantic system (Whitworth et al 2005).

As noted above, both phonological and semantic errors are said to occur if this level of spoken word production is impaired. Phonological errors arise at this level since only part of the word form is available or there is difficulty with accessing the phonological word form. On the other hand, the connection between the lexical semantic and the lexical phonological levels may be impaired and this may be the reason for the occurrence of both semantic and phonological errors (Cox 2008). Caramazza and Hillis (1990) describe two people with aphasia who produced semantic errors in naming (and reading aloud) despite being unimpaired in other modalities. These patients did not experience difficulties in writing tasks or in comprehension of spoken and written words (using the same stimuli used in naming). The researchers concluded that their

[1] In this task an individual may be requested to identify the correct picture from a choice of 4 pictures, which include a semantic distractor, a phonological distractor and an unrelated distractor.

patients' semantic errors were not arising from the lexical semantic level (as writing and comprehension were adequate), but from impairment to the phonological output lexicon.

Individuals who have difficulties with access to the phonological output lexicon have been shown to benefit from phonemic cues, while not being affected by phonemic miscues (Howard and Gatehouse 2006). While word frequency, familiarity and age of acquisition effects have been shown to affect people who fail at this level, word length appears to have no effect. Reverse length effects may be seen as a result of the higher incidence of phonological neighbours among short words. Repetition of words is not impaired.

2.2.1.4 Phonological assembly level

The phonological assembly level assembles a string of phonemes together in preparation for production of the words. Impairment at the phonological assembly level typically results in phonological errors in naming and impaired repetition. Thus, similar errors are produced in naming, reading and repetition of words, while comprehension is good (Howard 2000; Nickels 2000). Word length generally affects naming accuracy. Typical errors of impairment at this level include phonological errors, neologisms and *conduite d'approche* (i.e. "repeated attempts at the target word that often result in a closer approximation to the target word" (Whitworth et al 2005, p. 49).

Difficulties at this level are related to the selection and ordering of phonemes after the correct phonological word form has been retrieved. The errors that occur at this stage are not related to previous levels of production. They are post-lexical phonological errors which occur after semantic and lexical information has been correctly retrieved (Levelt 1999; Goldrick and Rapp 2002).

2.2.1.5 Articulation programming and execution level (phonological encoding)

Finally, this level involves the transfer of phonemes into neuromuscular activity (Whitworth et al 2005). Impairment at this level presents as articulation and/or phonological difficulties and is classified as dysarthria or apraxia of speech. Effects of word length and word complexity (e.g presence of consonant clusters) are generally seen (Howard 2000; Nickels 2000).

2.2.2 Debates regarding theory of lexical access

All the stages described above are important in order to understand the processes involved in lexical access. In fact, they may be viewed as a network of interconnecting processes that need to be activated at the right time. Theories of spoken word production enhance our understanding of the different stages involved. While earlier models of lexical access were concerned about the type of linguistic representations involved, more recent models focus on the interactions that exist between the different stages of spoken word production. Further, research is not simply concerned with *which* system may be impaired, but also with *how* it may be impaired (Wilshire 2008). However, as noted earlier (section 2.0), conflicting views regarding how these processes interact are put forward (see Nickels 2002; Goldrick and Rapp 2002; Wilshire 2008; Biedermann et al 2008). The two main controversies are between theories known as *discrete, modular* or *componential* theories (e.g. Level et al 1999), and *interactive* or *connectionist* theories (e.g. Dell et al 1997). An intermediate position has also been proposed by Rapp and Goldrick (2000), namely the *restricted interaction account* (RIA). The following provides a brief description of debates concerning the theories of lexical access.

2.2.2.1 The direction of flow of activation

One of the debates that is often mentioned in the literature on theories of lexical access concerns the disagreement regarding the direction of flow of information during lexical retrieval. There are two main views (and two main groups of researchers). Levelt et al (1999) and Roelofs (2004) suggest that there is a strict sequence of activation from one stage of lexical retrieval to the other in a *feed-forward* processing manner (discrete/componential theories). On the contrary, Dell et al (1997) hold that there is no strict sequence of activation, and they propose that activation may be allowed to flow in both directions with *feedback* between the different stages or processes of lexical retrieval (interactive theories).

2.2.2.2 The continuity of flow of activation

The second debate concerns how activation flows between the stages of lexical retrieval. While Levelt et al (1999) suggest that activation is *gated*, Dell and his colleagues propose a *cascading* activation. The first approach holds that the levels involved in lexical retrieval are *independent* and that there must be semantic processing (lexical selection) first before moving on to phonological processing. This implies that processing in the lexical selection level must be complete before moving on to the phonological retrieval level. In this way, once the lexical item has been selected, only one lexical unit will activate its corresponding phonological units. On the other hand, cascading activation maintains that there is *overlap* between the levels of processing, and, therefore, a number of lexical units may activate their phonological units even before the lexical item has been chosen.

Levelt and Dell worked on a number of experiments to produce evidence in favour of their proposed theories of spoken word production. Nickels (2001) outlines different factors which have been used to support the two conflicting models in both healthy and impaired individuals. These will be described briefly below. Discrete feed forward models as proposed by Levelt and his colleagues found support in performance on word-picture interference tasks. Picture naming reaction times of healthy participants were measured while semantic, phonological and/or orthographic interference was presented. Schriefers et al (1990) found that only semantic interference affected naming by increasing reaction time when the interference was presented before the picture, and phonological interference had no effect. On the other hand, when the interference was presented after the picture, phonologically related words reduced reaction time while semantic interference had no effect. This was shown as evidence that semantic activation takes place first and a phonological level is activated later without overlap between levels (see also Levelt et al 1991). Dell and O'Seaghda (1992), however, showed that the above could also be simulated in an interactive model.

Nickels (2001) also explains how lexicality bias and mixed errors in spoken word production have been used to support the interactive model. There is a tendency for healthy individuals to produce speech sound errors which result in real words rather than non words (i.e. lexicality bias), implying that both semantic and phonological levels interact to produce real word errors. The production of mixed errors (i.e. errors which are both semantically and phonologically related to the target word) which are said to occur more frequently than is expected by chance, is another source of evidence for interactivity, as proposed in Dell et al (1997)'s model. As a result of the feedforward mechanism mixed errors are highly unlikely to be accounted for by discrete models of spoken word production. While mixed errors have been reported in both impaired and unimpaired naming, lexical bias is less conclusive. For example Nickels and Howard (1995) failed to identify lexical bias in their 15 people with aphasia.

It is evident that the discrete and interactive theories have very strongly opposing views. Goldrick and Rapp (2002) offer an intermediate position between the two theories. Their theory (restricted interaction account, RIA) allows interaction between levels of processing, but they selectively restrict this interaction between semantic and phonological processes. While feedback occurs from the phoneme level to the lexical level, there is no feedback from the lexical level to the semantic level. The feedback is also limited in strength. Phonological information influences selection at the lexical level, without influencing the semantic level. They argue that there is evidence that this position can account for both normal and aphasic naming.

Despite the differences between the various theories, the cognitive neuropsychological approach can prove very useful in understanding both normal and impaired spoken word production, as it sheds light on which levels may be impaired, and also how each one may be impaired (Rapp and Goldrick 2006; Wilshire 2008). It is now accepted that normal language production theories enhance our understanding of impaired language, while research on impaired language production increases our understanding of normal language processes (Nickels 2002). In fact, it is interesting to note that Levelt and his colleagues developed their theory using research on normal language production, whereas Dell and his colleagues studied the impaired language of individuals with aphasia (Nickels 2001). Thus, both areas of research are essentially useful in enhancing our general understanding of the complex processes involved (Howard 2000).

Howard (2000) highlights the benefits of relying on cognitive neuropsychological models in clinical situations. As clinicians are faced with various types of word retrieval difficulties, cognitive neuropsychology enables the clinician to identify the level/s of difficulty in each case. As impairments at different levels of naming are reflected in different types of difficulties, a cognitive neuropsychological approach to assessment provides good indications regarding the levels which are intact and those that are impaired, as well as what should be treated. Material used in the clinical setting must take into account all sources of evidence, as relying on just one (e.g. the nature of the errors) may be misleading and possibly result in reaching inaccurate conclusions about the area/s of difficulty.

2.3 The BOSTON NAMING TEST

The Boston Naming Test (BNT) (Kaplan et al 1983) will be used in the present study. The BNT is a picture-naming test that has been designed to assess visual confrontation naming abilities. It includes 60 black line drawings of objects which are presented sequentially, from easy to more difficult to name. The BNT, was developed by Kaplan, Goodglass and Weintraub in 1983[2] with normative data from the USA and it has been deemed useful in assessment of both children and adults with a variety of conditions (including assessment of dementia, aphasia, head injury etc.).

There is evidence that the BNT is a valuable tool in assessing object naming and that it provides a good reflection of lexical retrieval abilities (Connor et al 2004). A number of different naming tests have been used to study lexical retrieval skills and what influences this ability. However, the Boston Naming Test (Kaplan et al 1983) is the most widely used naming test (see section 1.3, in Chapter 1). It is still currently used in research concerning spoken word production and it is still being studied and normed on a variety of populations to date. Examples include Dutch-speaking Belgian elderly (Mariën et al 1998), Australian elderly (Cruice et al 2000), Dutch-speaking Belgian

[2] An earlier version of the BNT included 85 items. The BNT used throughout this study is the 60-item standard form.

children (Storms et al 2004), Swedish (Tallberg 2005), Brazilian Portuguese (Mansur et al 2006), Greek (Patricacou et al 2007).

Despite its widespread use, the BNT cannot simply be used across populations without being aware of the diverse linguistic and cultural characteristics of each population. Performance on naming tests is influenced by a number of demographic and stimulus-related variables which may lead to different results in different populations. In fact, simply administering the BNT does not take into account the word variables that may play a part in naming accuracy and naming latency. Thus, adequate translation and adaptation is necessary. Each population must have its specific normative data in order to reach conclusions regarding the naming ability of the individuals concerned (Tallberg 2005; Zec et al 2007b). It is necessary to know which variables affect BNT performance in different populations, and in what way performance is affected. For example, age, education and gender may all contribute to acquiring a different score on the test, while word length and item familiarity may additionally affect an individual's score.

2.3.1 Research on the BNT

Research has used the BNT to investigate various issues related to the ability to name objects and retrieve single words. While some studies have focused on producing normative data for diverse populations (both monolingual and bilingual populations), others have focused on analyzing the impact of particular variables on naming ability. More specific research on the BNT is described in section 2.4 which looks at the influence of demographic and word variables on naming.

As a result of their efforts to standardize the BNT on different populations, researchers produced new culturally-appropriate versions of the test. For example, Mariën et al (1998) produced BNT norms for Dutch-speaking Belgian elderly while using the original list of the BNT items and providing Dutch target words. Kim and Na (1999), Cruice et al (2000) and Patricacou et al (2007) totally replaced a few of the BNT items with more culturally appropriate ones. Other studies recommend a new sequence of presentation of the 60 items of the BNT (e.g. Allegri et al 1997; Tallberg 2005) such that the order of presentation reflects the frequency of correct responses for that particular population.

Further research has focused on studying the usefulness of shorter versions of the BNT. Despite its benefit of being an *easy-to-administer* neuropsychological test, the 60-item BNT may be considered too long for individuals with particular difficulties, e.g. the elderly and people with head injuries, who may exhibit limited attention skills (Fisher et al 1999; Graves et al 2004). The validity of shortened versions of the BNT (e.g. 15 items only; 30 items only) has been investigated. The 30-item (odd or even) versions of the BNT have been found to be highly correlated with the complete set of the BNT and have been shown to have high internal consistency (e.g. Mack et al 1992; Williams et al 1989). Norms for these short versions have also been produced (e.g. Fisher et al 1999; Elkadi et al 2006).

2.3.2 Normative studies on the BNT

As noted above, various studies have produced normative data for specific populations. These norms are derived from different cultures and from different samples of varying sizes and characteristics. Thus, it is essential to be aware that normative data reflects the particular sample which is used to derive the data. Furthermore, the information that accompanies normative data depends on the specific findings of the particular studies, and on the area(s) that are selected as the focus of the study (e.g. bilingualism).

Norms are generally presented in terms of means and standard deviations (SDs) of scores, and are often stratified by age and education (e.g. Allegri et al 1997; Welch et al 1996). On the other hand, Mariën et al (1998) found that gender also played an important role in BNT performance and, therefore, provide means, SDs, range, median and a cut-off score (of - 2 SDs) stratified by age, education and gender.

Following their study, on a Spanish-English bilingual population (from Buenos Aires), Kohnert et al (1998) present separate norms with means, SDs and 95% confidence intervals for English responses, Spanish responses and a composite score (i.e. adding both English and Spanish correct responses together). Tombaugh and Hubley (1997) studied a group of 219 Canadians and produced normative data stratified by age and education for three different BNT measures, namely (i) scores for spontaneous correct responses only, (ii) scores for spontaneous correct responses and responses that follow a stimulus cue, and (iii) scores for responses that follow a phonemic cue in addition to (ii). Norms are presented as percentile scores. In the Hebrew naming test, Kavé (2005) also provided norms for these three measures. However, she not only provided means, SDs and percentiles stratified by age and education, but also stratified by country of birth, in order to reflect the cultural characteristics of the population in question.

2.3.3 Response patterns

Not only is it necessary to look at BNT performance in terms of scores achieved, but also in terms of the response/error patterns. An individual may achieve a score which is within the *normal* range for his/her age and education, while showing evidence of language pathology in his/her error profile (Storms et al 2004). It is also very useful to know which response is expected for each particular picture of the BNT, and possibly which other response(s) may be regarded as acceptable alternative responses. This information may only be available following research on a *healthy* sample of individuals from the specific population. Further, the availability of normal response patterns for different individuals would provide essential information that may be used in clinical situations to make decisions regarding typical and atypical error profiles (Tallberg 2005), as particular types of errors may be associated with different neurological impairments (Tombaugh and Hubley 1997) and may also provide essential information regarding the identification of the level of breakdown in word retrieval (see section 2.2.1). Thus, a typical error profile or response pattern may provide valuable additional information to normative data on scores.

Although naming errors are experienced by most individuals who have suffered aphasia following stroke or other neurological impairments, there is extensive variation in their error profiles. This variation is a result of individual differences between the individuals' type and severity of impairment as well as the distinct impairment of cognitive processes (Budd et al 2010). Research on naming errors has focused on various aspects. Some studies investigate the possible level of impairment/s associated with particular types of errors (e.g. Howard and Gatehouse 2006). Other studies analyse the influence of word variable effects on the type of errors produced (as reported by

Nickels and Howard 1995), while some have focused on the typical error types associated with particular impairments, such as the types commonly found in semantic dementia and stroke (Budd et al 2010).

Naming errors produced by people with aphasia are generally categorized into different types of errors. Cuetos et al (2002) classified errors into the nine types described by Dell et al (1997): semantic errors, formal errors (phonologically related real words), mixed semantic and formal errors (words that are both semantically and phonologically related to the target word), visual errors, neologisms, unrelated errors, no-response errors, circumlocutions and other errors.

Interactive theories of spoken word production tend to put forward a continuity thesis that holds that the naming errors produced by people with aphasia differ in quantity, rather than quality, from the error patterns produced by healthy individuals (Dell et al 1997). Schwartz et al (2006) explain that although some types of errors are more commonly produced in normal speakers (such as semantic and mixed errors), and non word, formal and unrelated errors are more likely to be produced by impaired individuals, there is a continuum from low to high severity of difficulty, i.e. both normal and impaired naming lie on this continuum. The authors do not make a definite differentiation between errors which may be typically produced by people with unimpaired and impaired naming.

As shown in section 2.2.2.2, a phenomenon that is consistent with an interactive approach to naming is the production of mixed errors. The production of a mixture of different types of errors by individuals with aphasia may also be accounted for by an interactive model (Dell et al 1997). Since this model assumes forward and backward feedback during activation, all levels of word production may influence the target word to produce different types of errors. Consequently, semantic errors may occur at any stage of word production in interactive models (Cox 2008). On the other hand, some people with aphasia tend to produce semantic errors only, or phonological errors only (see Cuetos et al 2000). These patterns of error are not consistent with interactive theories, but may be more readily accounted for by means of discrete models. The latter propose serial activation as there is no feedback from the activated form to the previously selected lemma, implying that semantic errors are more likely to occur at the semantic level, while phonological errors are more likely to occur at the phonological level (Level et al 1999). Thus, Levelt et al (1999) argue that mixed errors may occur only by chance and cannot be used to support the idea of interactivity (ibid.).

2.4 FACTORS THAT INFLUENCE NAMING ACCURACY AND LATENCY

There are numerous studies on lexical retrieval and the possible factors that may influence naming accuracy and naming latency. Both picture-naming accuracy and the time taken to name the pictures (reaction time or naming latency) have been studied to provide valuable insights into the process of naming and what affects it. As seen in the introduction (section 1.2, in Chapter 1) several demographic variables have been associated with performance on naming tests, in terms of naming accuracy and latency. Other variables are related to the material that is used and are called item/word variables or stimulus-related variables.

There are debates regarding which variable or variables exert the strongest effect on word retrieval. The following sections provide an overview of research findings in terms of the factors that have been associated with naming performance. Several proposals will be mentioned in relation to possible findings about the naming performance of the healthy individuals in this study.

2.4.1 Demographic variables that may influence naming

The following is a description of the different demographic variables and the effect that each may have on naming performance on the Boston Naming Test (Kaplan et al 1983). The relationship between naming and ageing will be discussed in detail in section 2.5 and will not be included here.

2.4.1.1 Language

Different languages use different words depending on their cultural setting. Yet, some populations with different languages may use similar words but of different frequencies, familiarity, age of acquisition and imageability. These variables are very much associated with one's culture (see section 2.4.1.6). Phonological complexity and word length also differ in different languages. All of these variables have a potential effect on naming accuracy and reaction time.

Further, different populations experience language in different ways. While some populations may simply be exposed to (and use) only one language (such as English) throughout their life, other populations may be exposed to two or three languages simultaneously since birth, or from a very young age. The issue of bilingualism and language performance, and specifically naming performance, has attracted numerous studies. These studies highlight the advantages and disadvantages of being bilingual. This area will be discussed in depth in section 2.6 with special attention being given to the language situation in Malta.

2.4.1.2 Level of education

The number of years of education has also been found to be related to BNT performance, with a number of studies showing that those with higher levels of education tend to perform better on naming tests (Worral et al 1995; Ross et al 1995; Welch et al 1996; Randolph et al 1999; Tallberg 2005; Patricacou et al 2007; Zec et al 2007a). In fact, a number of researchers have identified level of education as the best predictor of BNT results (e.g. Neils et al 1995; Allegri et al 1997; Tombaugh and Hubley 1997; Mansur et al 2006). Goral et al (2007) found that individuals with higher levels of education did not demonstrate ageing effects. On the other hand, Tsang and Lee (2003) found no education effect in their study using the Chinese Naming Test. However, this result has been attributed to the possible stimulus-related effects of familiarity and age of acquisition, as the test items were purposely chosen to be high in familiarity and low in age of acquisition. Nicholas et al (1985), Van Gorp et al (1986), Kaplan et al (1983) and Farmer (1990) also did not find education to be a significant factor. It has been argued that the reason for not finding an effect of education may lie in the fact that particular studies may have been biased towards more educated samples of people (as discussed by Allegri et al 1997; and Patricacou et al 2007). In fact, the studies (e.g. Van Gorp et al 1986) which included individuals with high education levels tended to achieve a high overall mean score (i.e. a mean of above 50 correct out of 60 BNT responses). Patricacou et al (2007) argue that this is one possible reason for achieving a lower mean score (of 42.86) in their study. In fact, a quarter of the participants of Patricacou et al (2007)'s study had just 0 to 6 years of education. On the other hand, the same study found that the mean scores of the younger, more educated participants were similar to the results of other studies and achieved mean scores above 50 correct. Thus, the studies which found no effect of education have generally been said to include samples of participants with higher levels of education overall.

Apart from finding lower mean scores in individuals with lower levels of education, Welch et al (1996) and Zec et al (2007a) also found a greater variability in BNT score in those with lower levels of education. When naming tests are administered on individuals of lower levels of education, more stimulus and phonemic cues are necessary (Mansur et al 2006), there is a tendency to omit responses and to admit that they do not know an item, in contrast to individuals with higher levels of education, who tend to suggest a response even if they are unsure of its accuracy (Tallberg 2005).

Based on the evidence presented above, it is proposed that the present study will also find an effect of education. Participants with higher levels of education are expected to score higher than participants with lower levels of education and presumably will have shorter naming latencies.

In addition to formal education, others have demonstrated the benefits of reading habits on naming performance (Padakannaya et al 2002 cited in Mansur et al 2006). Thus, it is also hypothesized that individuals who read more will produce more correct responses with shorter naming latencies.

2.4.1.3 Vocabulary

Hawkins and Bender (2002) showed that vocabulary size may be a better predictor of BNT performance than education. Therefore, when interpreting naming test results it is also necessary to assess participants' vocabulary skills. A low score on a naming test may actually be a result of limited vocabulary rather than problems with naming ability. This is particularly salient in a bilingual community such as Malta, as bilingual individuals have been shown to have smaller vocabularies in each language (Bialystok and Feng 2009; Gollan et al 2008). See section 2.6.4.

2.4.1.4 Gender

Although gender has been given less importance in the literature on naming performance, some differences have been reported. Welch et al (1996), Mariën et al (1998) and Randolph et al (1999) found that men scored higher than women on the BNT, and that men responded significantly better in 18 of the 60 test items of the test (Randolph et al 1999). This gender effect remained stable even when age and education were controlled. However, other researchers failed to identify any association between gender and confrontation naming (e.g. Ross et al 1995; Tsang and Lee 2003; Tallberg 2005). Tallberg (2005) reported, however, that women tended to admit that they did not know a response more than the male participants did.

It is important to note that some cross-sectional studies have highlighted the fact that the particular items included in the BNT may be biased towards better performance for men (Randolph et al 1999). Further, women of an older age may be more disadvantaged on the test as a result of educational, occupational and sociocultural factors. Fewer opportunities for further education, occupation and general exposure to particular areas of life had been experienced by women who are ageing now. Cultural and socioeconomic factors play an important part here (Cruice et al 2000; Patricacou et al 2007). In fact, Patricacou et al (2007) did not identify any main effects of gender in their study. However, they did find a significant interaction between education and gender on naming performance. This emphasises the importance of looking at sociocultural issues when attempting to understand the combination of variables that may affect naming performance. It is necessary to bear in mind the particular changes that occur in society over a period of years. For example, the availability of education for both males and females of different ages is significantly different. Further, men and women of different ages may play different roles in society – both in the family context and in the work place. Thus, the interactions of *age and education*, and *age, gender and* *education* with naming performance must be understood in the context of the particular age group and population that is being studied.

It is hypothesized that the present study will identify a gender effect in the ageing population. This proposal is based on the common knowledge that the young Maltese females (who are now ageing) of some 70 years ago, very often stayed at home to help with family chores instead of going to school, as formal education was not obligatory by law until 1947 (Zammit Mangion 2000). Thus, an interaction effect between gender, ageing and education is expected to appear in this study.

2.4.1.5 Global cognitive ability

Tallberg (2005) found that cognitive ability was significantly correlated with BNT performance, showing that better performance on an IQ test was associated with better performance on this naming test. Further evidence of the impact of cognition on naming was outlined by Van Gorp et al (1986) who reported that high functioning older adults demonstrated only a mild decline with age on the BNT. This may be explained by the cognitive reserve hypothesis which states that cognitively active adults may experience less cognitive decline (section 2.5.1).

2.4.1.6 Culture

A considerable difference between norms of different populations for the same naming tests is highly evident (Barker-Collo 2001; Casas et al 2008). These differences are explained by language and cultural differences across the populations (Storms et al 2004). Although the BNT has been used worldwide for clinical and research purposes, Cruice et al (2000, p.152) stressed that the "BNT is susceptible to the cultural effects of word frequency and familiarity" and word length. The authors reached their conclusions following their study on Australian participants, indicating that this confrontational naming test may not be universal after all. Changes in items in the test were necessary to accommodate the cultural differences (e.g. *pretzel* and *beaver* were replaced by *pizza* and *platypus* respectively). On the other hand, for the Swedish adaptation of the BNT, Tallberg (2005) decided that no pictures needed to be replaced due to cultural differences.

2.4.2 Stimulus-related variables that may influence naming

Stimulus-related variables (or word variables) are variables that arise from the material that is used in studies or in assessment. Thus, different tests may be affected by different stimulus-related variables depending on the pictures or words used in the particular test. The following is a description of some of these variables.

2.4.2.1 Name agreement

Name agreement is one of the variables that is given importance in picture-naming studies. It refers to the percentage of people who produce the dominant name for each picture (Székely et al 2003). If a picture is said to have low name agreement it is said to elicit a number of different names, while high name agreement refers to the case where different people highly agree on the name of the picture. This affects naming latency, as the higher the name agreement is, the quicker it is named (Alario et al 2004, citing numerous studies). When a picture may be named correctly by means of a number of possible alternative responses, naming latency is said to increase as a result of competition between the alternative names and the time taken for selection of the preferred name to take place (Ellis and Morrison 1998). Mitchell (1989, cited in Hodgson and Ellis, 1998) compared naming latencies of young (19-32) and older (63-80) groups of people and found that name agreement produced a global effect with equal strength in both groups of participants, indicating that there was no evidence of an interaction between age and name agreement.

It is proposed that, in the present study, items with high name agreement will be named quicker than items with low name agreement, and items with high name agreement will have a higher number of correct responses as competition between different names will be minimised.

2.4.2.2 Word frequency

Word frequency refers to the extent of use of a given word and it is related to one's experience, occupation and culture (Mansur et al 2006). Word frequency is often based on counts of written corpora (Alario et al 2004). However, recent research on frequency measures has shown that film subtitle frequencies may be more reliable than text, since the "situations depicted on the screen may be more representative of everyday life" (Cuetos et al 2011: 141).

Word frequency has been found to be significantly correlated to naming ability (Randolph et al 1999) in both healthy people and people with aphasia (e.g. Nickels and Howard 1995). It has been regarded as the most critical factor in naming studies and it is generally reported to be inversely related to naming latency (Alario et al 2002). Thomas, Fozard and Waugh (1977 cited by Hodgson and Ellis, 1998) did not report any difference in the effect of word frequency on naming latency in young and older participants.

Since the 1970's the effect of age of acquisition (AoA) has gained importance with a tendency to challenge frequency as the most important predictor of naming latency (NL) (Carroll and White 1973; Székely et al 2003). Ellis and Morrison (1998) argue that what may appear to be frequency effects may in fact be the effect of age of acquisition. In fact, Hodgson and Ellis (1997) reported age of acquisition to be the best predictor of naming accuracy when participants were allowed either 5 or 15 seconds to respond. On the other hand, Levelt (2002) proposes that *speed of recognition,* rather than AoA, may explain word frequency effects on NL. However, Alario et al (2002) state the contrary emphasising that low frequency items are accessed more slowly as it takes longer to access the particular item and not because it takes longer to recognise it. Almeida et al (2007) confirm the latter as they found that NL differs in low and high frequency words primarily due to lexical processes rather than picture recognition (familiarity with item) and articulatory processes.

In their study on bilinguals, Kiran and Tuchtenhagen (2005) also showed that naming was better correlated with frequency than with AoA, which contrasts with the findings of Alario et al (2004) on monolinguals, who found that both frequency and AoA produce separate effects on naming times. Rated word familiarity is another variable that has been correlated with word frequency (e.g. Gernsbacher, 1984 cited in Whitworth et al 2005). Moreover, frequency, AoA, familiarity and imageability are also said to be correlated with each other (Cortese and Khanna 2007; Feyereisen et al 1988). For example, word frequency is highly correlated to rated familiarity as the latter is rated according to how frequently a word is experienced. This has led to confounding effects which have not always been taken into account by different researchers, possibly leading to the false importance given to word frequency in lexical retrieval (Nickels and Howard 1995).

Low frequency words tend to be named less accurately than high frequency names, especially by people with semantic dementia and aphasia (Knobel et al 2008, citing several studies) and they are more likely to result in semantic and phonological errors Kittredge et al (2008). The effects of word frequency on performance of individuals with language impairments on naming tests have been said to emerge from difficulties throughout the different levels of word production (Whitworth et al 2005; Knobel et al 2008). Knobel et al (2008) argue that frequency effects probably do not arise at the semantic level (since words are not represented at this level) and at the segmental level (since post-lexical information is held here). Their findings indicated frequency effects emerge in the lexical-segmental connections. However, they concluded that this is only one possibility and that frequency effects may arise throughout the levels of lexical access.

Unfortunately, reliable *spoken* word frequency data is not yet available for the Maltese language. The available Maltese *written* word frequency data is based on a corpus of about 9 million words from newspapers (news, opinion, editorials etc), web texts (blogs, etc), Maltese wikipedia, speeches, government and information leaflets, literary works and law. Not all the 60 items of the BNT are included in this vast data.

2.4.2.3 Age of Acquisition

Age of acquisition (AoA) refers to the age at which words are learned in spoken or written form. It is said to have a significant influence on language and memory processes (Stadthagen-Gonzalez and Davis 2006). AoA has been shown to be significantly correlated to naming (e.g. Carroll and White 1973; Ellis and Morrison 1998) and to any activity that necessitates access to stored information (Catling and Johnston 2008).

AoA, familiarity and imageability are very often measured through subjective ratings. Although subjective, this procedure has been proven to be valid and reliable for AoA as the results of subjective ratings were found to be highly correlated with objective measures of AoA (Carroll and White 1973; Morrison et al 1997; Iyer et al 2001).

When carrying out subjective ratings for AoA, participants are asked to rate a picture or a word for a particular variable on a scale. Different researchers have used different scales to rate variables. For example, Bird et al (2001), Cortese and Khanna (2007) and Stadthagen-Gonzalez and Davis (2006) used a Scale of 1 – 7, to acquire AoA ratings. This was originally employed by Gilhooly and Logie (1980). Participants are asked to rate each word depending on *when* they think they learned and/or understood it in spoken or written form, **even if** they did not use, read or write it at the time. They are asked to mark the appropriate box according to the age/scale provided. Words that have been learned at an early age should be assigned a low rating. On the other hand, a value of *7* is assigned to words that are learned later in life. Participants are sometimes instructed to write **X** for words that they do not know.

Different authors have reached different conclusions regarding the importance of AoA (see section 2.4.2.2) for naming accuracy and naming latency. While frequency and AoA are highly correlated with each other (Morrison et al 1997; Randolph et al 1999), objective AoA (Ellis and Morrison 1998) and/or subjective AoA ratings seem to be the best predictor of naming latencies – even better than frequency norms and familiarity ratings (Morrison et al 1992; Iyer et al 2001; D'Amico et al 2001). In fact, Cortese and Khanna (2007) found AoA to be the best predictor of naming and, particularly, lexical-decision latencies. On the other hand, Alario et al (2004) concluded that both frequency and AoA are important as they make "*independent* contributions to naming times" (p.148). Further, Bird et al (2001) found that AoA was highly correlated with imageability, word frequency, familiarity and length. There is conflicting evidence regarding the impact of different variables on naming ability since studies have not always controlled for the effect of different variables. Nickels and Howard (1995) conclude that despite the intercorrelations that may exist between variables, there are reports of independent effects of AoA and word length on naming in people with aphasia. The authors found that AoA effects in aphasic naming indicates breakdown at the semantic or the lexical level. Late AoA words were more likely to result in phonological errors (Kittredge et al 2008). However, Nickels and Howard (1995) reported that AoA effects are a good predictor of the occurrence of semantic errors, while familiarity and length effects predicted the presence of phonological errors. Hodgson and Ellis (1998) also identified AoA effects (as well as name agreement and length effects) in their elderly subjects, finding that words which are learned early are named more accurately than words which are learned later in life. They suggested that these effects are a result of a "reduction in flow of activation between pictorial, semantic, and phonological representations in the elderly brain" (p.159).

It has been argued that AoA effects actually reflect the cumulative frequency of the particular words (Zevin and Seinderberg 2002, cited in Stadthagen-Gonzalez and Davis 2006). However, in their study on word and picture naming Morrison et al (2000) found evidence of strong AoA effects in both young and older participants. Thus, the authors concluded that these effects do not arise from cumulative word frequency effects, but rather are a result of the age or the order in which the words were learned.

Schwartz et al (2006) argue that words which are high in frequency and early in AoA tend to be named more accurately as a result of "learning mechanisms that set connection weights in the network" (p. 231). The greater the experience of a particular word, the greater the connections within the language system and the greater the chance that it will be produced correctly.

Judging from previous research it is hypothesised that in the present study, NL will be found to be positively correlated with AoA, number correct will be negatively correlated with AoA and that AoA will be found to be the best predictor of NL and number correct, when different item variables are included in a multiple regression.

2.4.2.4 Familiarity

Familiarity is defined as the extent to which one comes in contact with or thinks about a particular item (Mathuranath 2007). Word familiarity (rather than item familiarity) will be used in the present research. To acquire subjective ratings for word familiarity, participants are instructed to rate each word for degree of familiarity, depending on how often they think they come in contact with or think about the particular word. Stadthagen-Gonzalez and Davis (2006) used a Scale of 1 – 7 (which was originally employed by Gilhooly and Logie 1980) where a rating of 1 is chosen for words that are very unfamiliar, i.e. for items that are never seen, heard or used; number 2 should be chosen for items that one rarely sees, hears or uses; and number 7 is chosen for items that are very familiar and that are encountered nearly every day; etc.. It is not necessary for the participants to understand the meaning of each word as they are simply requested to rate words depending on how often they come in contact with them. Furthermore, there may be some words that may be used or heard more often than they have been seen, or vice versa. In such cases, participants are instructed to give the highest rating of the three.

Familiarity ratings have been strongly correlated with both written and spoken word frequency measures (Stadthagen-Gonzalez and Davis 2006) giving an indication that subjective familiarity ratings reflect word frequency. Gernsbacher (1984 cited in Nickels and Howard 1995) showed that experiential familiarity, and not frequency, influences naming in healthy subjects.

In people with aphasia, familiarity (as well as length) has been found to predict phonological errors. Although, it is still unclear whether familiarity effects arise from impairment at the semantic level or the lexical level (Howard and Gatehouse 2006), Nickels (1995) showed that familiarity effects are related to the lexical level.

Familiarity ratings have been found to be highly correlated with AoA (ibid). In fact, Bird et al (2001) found familiarity, as well as imageability to be the best predictors of subjective AoA. Therefore, considering that familiarity has been found to be highly correlated with AoA, and that young AoA is associated with high familiarity ratings, it may be proposed that in this study NL will also be found to be negatively correlated with familiarity, and number correct will be positively correlated with familiarity.

2.4.2.5 Imageability

Imageability is a semantic variable that refers to the degree to which an item evokes mental images (Stadthagen-Gonzalez and Davis 2006). Words differ in their capacity to arouse mental images of things and events. Words that are said to arouse a sensory experience (mental picture or sound) very quickly and easily are said to be high in imageability. Other words may arouse a mental picture only with some difficulty (or after a long delay) and are said to be low in imageability (Cortese and Fugett 2004). Furthermore, the higher the imageability the greater the number of different images that are aroused when the word is heard (Alario et al 2004; Stadthagen-Gonzalez and Davis 2006). For example, the word *apple* would probably arouse a mental picture very easily, and a number of different images may be evoked; that is, it has high imageability.

For subjective ratings of imageability, participants are asked to rate each word depending on how easy and quickly it is for the particular word to arouse a mental picture. Alario and Ferrand (1999) and Alario et al (2004) used a 5-point scale. Participants are requested to mark the appropriate box according to the scale. A rating of 1 is chosen for words that are low in imageability and a rating of 5 is selected for words with high imageability (i.e. words that arouse a mental picture very quickly and easily). Participants are encouraged to use any value from 1 to 5 on the scale when rating the words.

Several authors have found correlations between imageability and naming performance and, in particular, naming latency, including Alario et al (2004), who found that imageability (as well as Familiarity, AoA and Name Agreement) produced an independent effect on NL. There is evidence, from people with aphasia, of better performance on words which are highly imageable (Nickels and Howard 1995; Bird et al 2000). Furthermore, in their study on accessing verbs and nouns in 4 patients with aphasia, Bird et al (2003) concluded that variables, such as imageability, may influence performance on lexical tasks rather than the grammatical category of the words. Together with familiarity, imageability has been shown to predict subjective AoA results (Bird et al 2001). On the other hand, Morrison and Ellis (2000) failed to find any effect of imageability and familiarity on NL. Cortese and Khanna (2007) also found that once AoA was included in a regression analysis, imageability did not remain significantly correlated to NL.

High imageability words have a larger number of semantic features (Plaut and Shallice 1993, cited in Nickels and Howard 1995) than low imageability words. For this reason, highly imageable words experience more activation, resulting in a tendency to resist errors, and to be named correctly (Schwartz et al 2006). In people with aphasia, imageability effects have been generally associated with semantic level impairment, as disturbances at this level may lead to selection and activation of less imageable competitors, resulting in the production of semantic errors (Nickels and Howard 1994; Nickels 1995). However, it is also possible to find imageability effects in post semantic levels of spoken word production. Franklin et al (1995) describe a patient who had difficulties with low imageability words despite having an unimpaired semantic level, which was demonstrated by intact written comprehension.

Since some studies have identified correlations between imageability and naming performance in both the normal population and in people with aphasia, it is proposed that naming performance would also be correlated with imageability in the present study as well.

2.4.2.6 Word length

Word length effects are generally explained by findings that state that word retrieval difficulty increases with increasing word length. A positive correlation between reaction time and word length has been reported in both people with aphasia and in people without aphasia (e.g. Nickels and Howard 1995; Morrison et al 1992 respectively). However, there is disagreement regarding whether it is phoneme length, syllable length or the number of clusters in the particular word that actually produces word length effects (Whitworth et al 2005). Nickels and Howard (2004) argue that phoneme length (and not syllable length or complexity) affects naming accuracy in people with aphasia. An effect of length is commonly seen in people with aphasia who produce phonological errors, and is indicative of a breakdown at the phonological assembly level of spoken word production. The longer the word, the greater the possibility of incorrect production, as there is a bigger opportunity for a phoneme to be accessed incorrectly. This may occur while the correct word form undergoes articulation (Nickels and Howard 1995; Schwartz et al 2006). On the other hand, reverse length effects may be the result of impaired access to the lexicon. (Nickels and Howard 1995; Howard and

Gatehouse 2006). These reverse length effects may be explained by neighbourhood density. Since shorter words may have more phonological neighbours than longer ones, there is a greater competition to access the correct item from several activated 'neighbours'.

Word length, word frequency and age of acquisition have been shown to be correlated with each other (Morrison et al 1997). Therefore, an apparent effect of word length may actually arise from word frequency or age of acquisition effects (Hodgson and Ellis 1998). However, when controlling for word frequency and age of acquisition, Hodgson and Ellis (1998) found a main effect of word length (in terms of phoneme length) on naming accuracy in an elderly group of participants. They found that shorter words were more accurately named than longer ones when 5 seconds were allowed for naming. However, this effect disappeared when 15 seconds were allowed. Damian et al (2010) did not find any effect of length on naming latency. However, their study consisted of only young students and only one-two syllable words were used. It is proposed that phoneme length may also produce an effect on naming accuracy and naming latency in the present study as well, that is test items will have more total correct responses as phoneme length decreases, and shorter naming latency as phoneme length decreases too.

2.4.2.7 Animacy

The predictor variable, animacy, is a semantic variable. Evidence of animacy effects have been shown by some researchers, such as Pilgrim et al (2005), who showed that non living items took longer to be named in a task of concept identification, rather than picture naming. On the other hand, Janyan (2007) found that when pictures were presented to the right visual field (left hemisphere) the opposite effect emerged (i.e. non living things were named faster). It is proposed that during naming of animate items there is a greater activation of the numerous shared characteristics of the living/animate item. This creates conflict or competition to select and retrieve the appropriate name and, therefore, this leads to a disadvantage for both naming accuracy and latency (Janyan 2007), especially in the left hemisphere as the retrieval conflict may be stronger in this hemisphere. On the other hand, studies on people with aphasia have produced varying results in terms of animacy. Cuetos et al (2002) found that animacy was not an independent significant predictor variable. However, two of the patients in their study were better at naming inanimate items as a result of different areas of breakdown in single word processing.

The number of variables that may have an effect on naming performance include ageing, education, gender, word frequency, word length and age of acquisition among others. This complicates studies as it may be somewhat difficult to study one variable alone while controlling for the others (Cortese and Khanna 2007). It has been noted that particular effects of one variable may in fact be attributable to another separate variable; one variable may actually have an effect on another one; and different variables may or may not be correlated to other variables. Regression analyses have been used by many researchers to target this difficulty and to determine the independent influence of each variable. As noted above, these variables have been the subject of much research and debate as different researchers have reached different conclusions regarding which variable may have the dominant influence on naming performance.

2.5 AGEING

Ageing has been shown to produce several changes in a variety of functions, including changes in sensory, musculoskeletal, immunological, cardiovascular, neurological and respiratory function (La Rue 1992 cited in Tsang and Lee 2003). Among the neurological changes that are said to occur in the elderly we find: loss of neurons and changes in dendritic arbors, increased neurofibrillary tangles and neuritic plaques, and reduced brain weight and volume (ibid.), particularly in the frontal areas of the brain (Cabeza et al 2005). It is evident that some changes, or decline, in intellectual function will follow. In fact, numerous studies have demonstrated that ageing produces particular changes in cognitive functioning and, consequently, in naming ability. The following is a description of the cognitive and neurological changes that have generally been associated with ageing.

2.5.1 Cognitive decline in normal ageing

Ageing individuals tend to assume that growing older will undoubtedly lead to memory deficits – an assumption that has been demonstrated in various studies (Cruice et al 2000; Wahl and Kruse 2003). Some degree of cognitive decline has been associated with ageing as a result of the neurological changes described above. However, rather than a loss of cognitive functions, normal ageing is said to cause a slowing of information processing (Verhaeghen and Salthouse 1997 cited in Lindeboom and Weinstein 2004). Specific cognitive changes have been frequently associated with ageing. These include changes in memory (e.g. episodic memory, possibly resulting from poor concentration, lexical retrieval and executive function) (Rugg and Morcom 2005); language (possibly including comprehension, expression and naming difficulties); visuospatial functions (e.g. perceptual speed, difficulties with spatial orientation, visualisation and construction); and executive functions (which may be demonstrated by apathy, loss of initiative or impulsivity) (Connor et al 2004; Lindeboom and Weinstein 2004).

A number of factors seem to affect the degree of cognitive decline, however. First and foremost, it must be emphasised that there is substantial variability in cognitive decline in different ageing individuals. The *cognitive reserve hypothesis* (Scarmeas and Stern 2003, cited in Connor et al 2004) provides one explanation for this variation between different individuals. This theory has been proposed to explain the idea that individuals who lead a cognitively active lifestyle, and who have high levels of education and intelligence may benefit from protection against cognitive decline. Bilingualism is also said to protect against cognitive decline in both healthy ageing and Alzheimer's Disease (Bialystok et al 2007). On the other hand, Christensen et al (1997) suggest that a higher level of education may help to compensate for the difficulties that are commonly associated with normal ageing rather than protect against them. Their conclusions were reached from findings that showed that individuals with different levels of education do differently on different cognitive tests, including naming tests.

Cabeza et al (2005) stress the importance of distinguishing between neurogenic and psychogenic effects (such as use or disuse of cognitive processes) on cognition in ageing individuals, as both types of effects may cause changes in cognitive function. In studies of healthy ageing most researchers emphasise the importance of excluding participants with known neurological conditions and other disturbances that might have an effect on cognitive performance. However, other factors that may possibly be associated with ageing (including cerebrovascular disease and diabetes) have been shown to adversely affect brain structure and function. Thus, commonly occurring conditions such as hypertension and diabetes mellitus, which are generally regarded as being of minor importance in case histories related to cognitive function assessment and research studies, may, in reality, have an influence on results of cognitive function (including naming and memory) tests in ageing (Raz 2005). In their recent study, Albert et al (1988) found that lexical retrieval was significantly influenced by hypertension but not by diabetes mellitus.

2.5.2 Ageing and language

The language of the elderly is characterized by an overall slowing in connected speech and decreased fluency (Bortfeld et al 2001, cited in Belke and Meyer 2007). When naming objects older adults tend to be slower and they pause more frequently between object names, suggesting a general slowing of speech (Belke and Meyer 2007). In fact, difficulties with word retrieval and Tip of the Tongue (TOT) states are common phenomena in the elderly.

2.5.3 Naming in ageing

As noted earlier, difficulties with naming are very common in various neurological conditions including Alzheimer's disease, frontotemporal dementia, corticobasal degeneration, head injury and in aphasia following strokes (Tsang and Lee 2003; Grossman et al 2004). However, a number of studies have shown that *normal* ageing leads to naming difficulties, especially over 65-70 years of age (Van Gorp et al 1986; Belke and Meyer 2007; Wierenga et al 2008; Figueredo Balthazar et al 2008). In fact, it is the most common and troubling complaint associated with ageing individuals (Maylor 1997 cited in Evrard 2002).

Many studies have found evidence of decline in lexical retrieval skills and an increase in response time with increasing age (Hodgson and Ellis 1998; Randolph et al 1999; Hawkins and Bender 2002; Connor et al 2004; MacKay et al 2005; Patricacou et al

2007) even after controlling for the effect of education (e.g. Tsang and Lee 2003). When studying the performance on the Boston Naming Test (BNT), Mariën et al (1998) and Patricacou et al (2007) found significant main effects of age. This has especially been found in those older than 70-80 years (e.g. Albert et al 1988; Ross et al 1995; Zec et al 2007b). While Welch et al (1996) found an interaction between age and education, Patricacou et al (2007) found an interaction between age, education and gender.

On the other hand, some studies found only a mild decline in naming with ageing (Van Gorp et al 1986) while others did not find a significant relationship between age and BNT performance at all (e.g. Farmer 1990; Allegri et al 1997; Tallberg 2005; Mansur et al 2006). Conflicting results may be due to the different kinds of samples and material used. Zec et al (2007a) argue that a possible reason for not finding a relationship between ageing and naming may be that the sample used was not large enough to be representative (e.g. including fewer than 200 participants). Further, most research on the effects of demographic variables on naming ability has used cross-sectional studies rather than longitudinal ones. This may lead to possible cohort effects as different people from different generations are naturally exposed to different experiences (including type of education, cultural experiences and nutrition) and, therefore, may perform differently on cognitive tests (Zec et al 2007a). Particular objects in picture naming tests may have become disused or changed in form over the years, implying that people of different ages may not recognise the same picture in the same way (Storms et al 2004).

Age and education are two variables which may very possibly be confounded as older people may have had fewer opportunities for education in most cultures. On the other hand, it must be noted that several studies have included samples of people with a generally higher level of education, which may not really be representative of the similarly aged population (e.g. Van Gorp et al 1986; Randolph et al 1999). In their attempt to look at the effect of ageing on naming over time, Cruice et al (2000) conducted a four-year longitudinal study. They reported only a weak effect of age when they looked for change in naming ability. However, it must be noted that they used a sample which was biased towards younger elderly, with an average education of 11.24

years. The authors stress that a longer study with a less-biased sample may have produced different results.

A greater variation in the performance on naming tests is evident in the older elderly population (Tsang and Lee 2003; Zec et al 2007a). This is probably a result of variation in cognitive abilities associated with normal ageing (Lindeboom and Weinstein 2004) and possibly a reflection of the variability in education and cultural experiences of the elderly of today (Zec et al 2007a).

2.5.3.1 Tip of the tongue states

Tip of the tongue states (TOTs) are fairly common in the language of the elderly (see Evrard 2002; Goral et al 2007). In a TOT state the individual is aware that he/she knows the name of a familiar person or object, has accessed its semantic information, but only partial phonological and orthographic information is available, resulting in a temporary inability to retrieve the name (Evrard 2002; Zec et al 2007a).

Evidence of increased occurrence of TOTs with increasing age may be explained by the *Transmission Deficit Hypothesis* which suggests that failure to achieve full phonological retrieval results from weakening connections between levels of lexical access (Goral et al 2007). Evrard (2002) found that the elderly experienced more TOTs overall, and the number of TOTs were found to increase with age for retrieving proper names but not for retrieving common names. Tsang and Lee (2003) attribute these common naming difficulties to particular changes that occur in the brain - changes which lead to decline in cognitive function, and hence naming ability, including picture naming.

Although naming difficulties exist in both normal ageing and in neurological conditions, one must emphasise that there are distinct types of difficulties associated with them. While normal elderly retain good vocabulary skills (Huntley and Helfer 1995 cited in Cruice et al 2000) and are said to experience lexical access and word retrieval errors (with evidence from enhanced naming ability with cueing, and the presence of frequent circumlocutions (Nicholas et al 1985), individuals with

Alzheimer's Disease produce naming errors that arise from problems in semantic knowledge, lexical retrieval and visuoperceptual impairments (Zec et al 2007b).

2.5.3.2 Neurological changes associated with naming

Naming requires visual-perceptual abilities together with adequate semantic and phonological processing. The skill of naming is governed by a large-scale neural network (Grossman et al 2004), involving prefrontal, frontal and temporal regions for memory, retrieval and executive functions (Tsang and Lee 2003; Wierenga et al 2008). Over the age of 65 there is evidence of a decrease in gross brain volume in the prefrontal region (Tsang and Lee 2003). Wierenga et al (2008) investigated neural activity during word retrieval and found that, compared to young adults, older adults demonstrated a greater activity in the left hemisphere frontal lobes (i.e. some of the areas usually associated with language). Their findings also included similar accuracy rates between young and old adults, but slower word retrieval for the older adults, hence concluding that healthy older adults may have problems in the executive functions of selecting, retrieving and manipulating lexical-semantic information rather than in lexical-semantic knowledge or semantic memory (Wierenga et al 2008).

Furthermore, in normal ageing inhibitory processes in working memory are said to weaken, peripheral perception and motor performance may decline, and attention skills may deteriorate. These changes contribute to the decline in naming performance and naming latency. An increase in naming latency with increasing age is potentially related to a slowed peripheral sensory and motor process, and neuronal loss, which lead to difficulties in perception, lexical processing and executive functions (Tsang and Lee 2003).

The processing of visual information (as in picture-naming tasks) becomes increasingly slowed as a result of ageing sensory functions. When an ageing individual is expected to name a visual stimulus that has been presented for a very limited time, the slowed processing effect is intensified even more (Madden et al 2005).

De Leon et al (2007) provide an overview of affected areas in the brain and the different corresponding difficulties that occur in picture naming. Different neural regions have also been associated with naming different word categories. For example, the left posterior lateral temporal and occipital regions are associated with retrieval of names for tools and utensils, while the left dorsolateral frontal cortex is associated with retrieving action words (Zec et al 2007a). Furthermore, Sörös et al (2003) found differences between healthy individuals and anomic patients with regards to verb and noun retrieval. While healthy individuals showed a similar neural network for verb and noun retrieval, the anomic patients demonstrated distinct cortical correlates for these two grammatical categories.

Despite the neuroanatomical and neurochemical changes that are negatively associated with ageing (Li 2005), different ageing individuals exhibit different levels and different rates of cognitive decline. One possible reason may lie in the interaction between neurogenic and psychogenic factors, implying that individuals who use particular cognitive processes may demonstrate less cerebral atrophy than individuals who do not use them (Cabeza et al 2005). There is also evidence of improvement in particular cognitive functions, such as memory, when these are targeted in therapy, implying that brain plasticity is still somewhat preserved in ageing (Ball et al 2002, cited in Nyberg et al 2005).

Based on the research findings presented above regarding the increasing difficulties with perceptual abilities, lexical processing and executive functions that are associated with ageing, it is hypothesized that naming performance will deteriorate with increasing age in the present study too. Both an increase in naming latency and a decrease in score with ageing are proposed hypotheses as a result of both psychogenic and neurogenic factors. There is evidence that both neuroanatomical and neurophysiological changes occur in ageing resulting in a subsequent cognitive decline. Furthermore, since the first part of the present study will use timed picture naming, and involves the presentation of a visual stimulus for only a short time, it is proposed that this method will intensify any naming difficulties that may be present, as visual information processing has been shown to slow down. The slowing of the processes involved in lexical access (including lexical semantics, lexical phonological, phonological assembly and articulation programming and execution levels) may also contribute to a slower retrieval of names. On the other hand, it is hypothesized that the older adults with higher education will exhibit less deterioration in naming ability in this study (as may be explained by the cognitive reserve hypothesis). Furthermore, exposure to a bilingual context may also provide further protection or compensation for the deterioration in language associated with normal ageing (Bialystok et al 2004; 2006).

2.6 A BILINGUAL CONTEXT

The term *bilingual* refers to people who can use and/or understand two languages (Brutt-Griffler and Varghese 2004). This is a broad term and does not tell us much about the extent of the bilingual person's ability to use and understand each language. There is much variation in proficiency of a second language and different authors define *balanced bilinguals* and *bilingualism* in different ways (Christoffels et al 2007). Definitions vary from those that describe individuals who acquire a second language at a later stage in life, to those who acquire both languages simultaneously from an early age and who are considered to be highly balanced bilinguals (Rivera Mindt et al 2008). Ellul (1978) describes two types of bilingualism: *co-ordinate bilingualism*, where two languages function independently and are often acquired in two different contexts; and *compound bilingualism*, where both languages are used and acquired within the same contexts. "The majority of Maltese in Malta are compound bilinguals …" (Ellul, 1978, p. 2) implying that code-switching is a very frequent phenomenon. It is, therefore, necessary to explain the bilingual context that is experienced in Malta, in order to understand the possible effects of using more than one language on naming performance.

2.6.1 Languages used in Malta

Both Maltese and English are the official languages of Malta, while the national language is Maltese. As a result of a history of nearly 200 years of British rule, the English language came to play an important role in the languages used in Malta. The 2005 Census (National Statistics Office 2006) found that the Maltese language is spoken by 97.9% of the Maltese population, while English is spoken by 87.9% of the population. However, these numbers do not necessarily imply high proficiency of each language. Further, it should be noted that many Maltese can also speak the Italian language. This is mainly a result of the close geographical proximity of the two countries and also from exposure to Italian TV. In fact, 56.7% admitted that they can speak Italian too.

2.6.1.1 The language used at home

In their research, Sciriha and Vassallo (2006) found that the majority of the Maltese population (96.2%) speak the Maltese language to meet their daily needs, even if they are bilingual. Switching between languages (sometimes within the same sentence, as in *Me* below) is also common place in Malta. Ellul (1978) classified the Maltese by using the following taxonomy to describe language use in Malta. People in Malta may use:

(i) Maltese only (M)
(ii) Maltese but with some English words (Me)
(iii) English only (E)
(iv) English but with some Maltese words (Em), or
(v) Roughly equal degrees of Maltese and English (ME).

Most Maltese people are exposed to some degree of English from a very young age, even if the language used in their family context is predominantly Maltese. This exposure to the English language comes from a variety of sources, including TV shows and adverts, family and/or friends who use English, books and other reading material (including children's books), signs and adverts in the streets, official letters from government or private institutions, etc. Some people also tend to use English key words in an otherwise Maltese sentence when addressing young children.

2.6.1.2 The language used in schools

The language used to address children in schools depends on the type of school attended. In general, most private schools and Church schools use mainly English as their language of instruction, while most state schools use Maltese with some English in particular lessons. On the other hand, higher education institutions tend to use English as their main language of instruction. Written material, such as textbooks and notes, is generally only available in English at this higher level of education. Thus, even if a person is not exposed to English at all within the family setting, he/she is bound to receive some degree of exposure to the English language as all children in Malta experience either simultaneous or sequential bilingualism (Grech and Dodd 2008). The degree and timing of exposure depends on the language choices made by the family, the schools attended, the area of living[3] and the place of occupation.

2.6.2 Studies on naming in bilingual populations

There is agreement that bilingualism is an influencing factor on picture naming (Gollan et al 2007). Among the effects mentioned in studies one finds that bilingual people tend to be slower in naming and also have more TOTs (e.g. Gollan and Acenas 2004; Bialystok et al 2008a). See below for a summary of the advantages and disadvantages of being bilingual (section 2.6.4).

Many studies that relate to naming performance and bilingualism tend to use Spanish – English bilinguals in America. Most researchers acknowledge the fact that Spanish has become the second most spoken language (after English) in the USA where Spanish-speaking immigrants are common (Kohnert et al 1998; Rivera Mindt et al 2008). Therefore, it provides ample opportunities to study the phenomenon of bilingualism. It has been noted that many studies on bilingual communities tend to include people who have acquired Spanish (for example) as their first language, and who use the Spanish language in their family contexts. However, most learn English from the time they go to school and acquire English as their dominant language. This phenomenon is called *switched dominance* (Ivanova and Costa 2008). Evidently, the Spanish-English bilingual situation mentioned here differs from the Maltese-English situation in Malta. In the Maltese islands, for the majority of the population, Maltese is both the first *and* the dominant language used (National Statistics Office 2006), even if a person is highly proficient in English too. The case of the Spanish-English bilingual situation mentioned here would be similar to the situation experienced by the large number of Maltese immigrants who emigrated to Canada, America and Australia some 40 years ago. In these cases, the Maltese used the Maltese language mainly within family settings and spoke English outside the family. This led young children of emigrant families to acquire Maltese as their first language and English as their dominant language (similarly to the Spanish-speaking emigrants mentioned above).

[3] While 10.7% of the population use English as their main spoken language at home in the Northern district, only 1.4% use English at home in the southern district (2005 Census, National Statistics Office 2006).

2.6.3 Lexical retrieval in bilinguals

> Bilingual speakers show remarkable flexibility in their ability to control their language output. They can restrict their speech to one language only, but also intentionally switch between languages in bilingual settings.
>
> (Christoffels et al 2007, p.192)

Bilingual speakers possess particular skills in accessing the words which they want to match to their conceptual representations. In fact, neurological differences have been found between monolinguals and bilinguals – the latter having a greater density in grey and white matter in the left inferior parietal cortex (Mechelli et al 2004, cited in Bialystok et al 2008b).

Research on bilingual and monolingual individuals has shown that there are both negative and positive differences associated with bilingualism. In a bilingual person, even simple lexical retrieval produces some conflict to select and retrieve the appropriate word. As both languages are said to be active all the time, it is essential to control the activation of one language while allowing the other (Costa 2005; Rivera Mindt et al 2008). There is less agreement regarding how this "cross-language" interaction influences language production (Finkbeiner et al 2006; Bialystok et al 2008b) and whether the phonological level of both languages is also activated at the same time (Christoffels et al 2007). The cognate facilitation effect (see below) may lead to the conclusion that the phonological forms of both languages are simultaneously activated in bilinguals (ibid). This conflict between the two languages leads to naming accuracy and latency effects (Roberts et al 2002; Gollan et al 2005). In fact, compared to monolinguals, bilinguals are said to be slower in lexical retrieval (Bialystok et al 2008a). Not only is lexical retrieval slower in bilinguals' non-dominant language, but it has also been shown to be slower in their more dominant language too (Gollan et al 2007). On the other hand, bilinguals are found to be faster and more accurate when naming pictures in their dominant language than in their non-dominant language (Kiran and Tuchtenhagen 2005; Christoffels et al 2006) – although this effect may be reversed in contexts of language switching (Costa et al 2006). It is important to note that the effort that is involved in resolving the conflict between the two languages may have the beneficial effect of boosting the executive function processes, which are governed by the frontal lobes (Bialystok et al 2008b). The executive functions governed by the frontal cortex include response switching (from one language to another) and response suppression (*deactivating* the undesired response) (Christoffels et al 2007). Interestingly, two (rather than one) frontal areas of the brain (left dorsolateral prefrontal cortex and the anterior cingulate cortex) are active when bilinguals perform a naming activity (Hernandez and Meschyan 2006, cited in Bialystok et al 2008a).

Another positive effect for bilingual naming may arise from the *cognate facilitation effect*. Some studies have investigated the effect of *cognate* words showing that the presence of cognates may facilitate retrieval of words. "*Cognates* are those translation words that have similar orthographic-phonological forms in the two languages of a bilingual ….; *noncognates* are those translations that only share their meaning in the two languages" (Costa et al 2000, p.1285). Thus, 'piramida' (Maltese) and *pyramid* (English) are cognates; while 'sodda' (Maltese) and *bed* (English) are noncognates. In fact, studies reveal that in languages with many cognates, bilinguals achieve a higher score on standardised tests (Gollan et al 2007), while monolinguals do not experience any cognate effects (Gollan and Acenas 2004). More cognate effects were found in the *non-dominant* language of bilinguals by Gollan et al (2007) and in high frequency words (Ivanova and Costa 2008). Further, faster naming latencies were recorded for cognates (Christoffels et al 2007) especially in the *dominant language* of bilinguals (Costa et al 2000). Cognates were named more accurately even by bilingual aphasic patients (Kohnert 2004).

The variables that are generally associated with monolingual naming (see section 2.4) may also affect bilingual naming. Bilingual naming has been found to be influenced by frequency of usage rather than by age of acquisition (Kiran and Tuchtenhagen 2005; Abulatebi and Green 2007). This has been shown by findings of faster and better naming skills in the dominant language.

2.6.4 Advantages and disadvantages of being bilingual

Bilingualism offers both disadvantages and advantages for the speaker. Table 2.1 provides a summary of the bilingual advantages and disadvantages that are found in the literature. While the disadvantages of being bilingual mainly include difficulties with lexical retrieval and lower vocabulary scores, the advantages of being bilingual are especially in the areas of executive functioning (Bialystok 2006; Costa et al 2008).

There are different theories that attempt to explain the reasons for the lexical retrieval disadvantages experienced by bilinguals. Possible reasons include: word frequency effects (as bilinguals use each language less frequently than monolinguals, therefore resulting in weaker links); cross-language interference/competition (as words from both languages are activated and compete for selection); and AoA effects. In cases of switched-dominance bilinguals (see section 2.6.2), the AoA may be at a later stage for those whose dominant language is not their first language (Ivanova and Costa 2008). It is evident that when one aims to interpret the results of assessment material that is administered on bilingual individuals, one cannot simply use the normative data available if this is not specifically derived from a bilingual population (Grech and Dodd 2008; Kohnert et al 1998). Grech and Dodd (2008) found that the Maltese bilingual context possibly enhances phonological acquisition and places bilingual children at an advantage over monolinguals (in the Maltese context at least). On the other hand, bilinguals are said to perform worse and more slowly in lexical retrieval than monolinguals even when tested exclusively in their dominant or non-dominant languages (Kohnert et al 1998; Roberts et al 2002; Gollan and Acenas 2004; Ivanova and Costa 2008). In the case of bilinguals, Kohnert et al (1998) found that the "alternative scoring method [i.e. responses in either language are both scored as correct] was a much better indicator of this group's confrontation naming skills than the single language only testing" (p.435).

It is important to note that the lexical retrieval disadvantages associated with bilingualism may actually be a result of smaller vocabularies rather than weak lexical access (Bialystok et al 2008a). Therefore, it is also necessary to be aware of bilinguals' vocabulary skills when interpreting lexical retrieval results.

As the Maltese-speaking participants involved in the present study were all exposed to some degree of bilingualism, it is proposed that naming will be less accurate and slower than that reported by other studies on monolingual individuals, as bilinguals have been shown to name pictures less accurately (Roberts et al 2002) and more slowly (Gollan et

TABLE 2.1

Disadvantages	Advantages
- Children who acquire two languages simultaneously are said to be slower (up to school age) in language acquisition than those who acquire the second language sequentially (River Mindt et al 2008).	- A bilingual environment enhances phonological acquisition (Grech and Dodd 2008).
- Both bilingual children and bilingual adults have smaller vocabularies in each language (Bialystok and Feng 2009; Gollan et al 2008).	- Executive functioning skills develop earlier in bilingual children (Bialystok 2001).
- Bilinguals name pictures more slowly (Gollan et al 2005; 2008).	- Bilinguals have enhanced executive functioning skills (Bialystok 2006; Costa et al 2008) and better performance on fluency tasks if vocabulary is controlled (Bialystok et al 2008a).
- Bilinguals name pictures less accurately (Roberts et al 2002).	- Bilingualism possibly offers protection against cognitive decline in ageing (Bialystok et al 2004; 2006).
- There is evidence of more TOTs in bilinguals (Gollan and Silverberg 2001; Gollan and Acenas 2004).	- Monolinguals and bilinguals show similar performance on a forward spatial span task; but bilinguals are better on a backward span[4] (Bialystok et al 2008a).
- Bilingual elderly also name fewer pictures correctly than monolingual elderly (Gollan et al 2007).	- When vocabulary skills are controlled, it is possible that bilinguals perform similarly to monolinguals (Bialystok et al 2008a).
- Bilinguals score less than monolinguals on verbal fluency tests – especially on semantic rather than letter fluency tests (Gollan et al; 2002).	

Summary of bilingual advantages and disadvantages

al 2005; 2008), as well as having smaller vocabularies (Bialystok and Feng 2009; Gollan et al 2008). On the other hand, it is hypothesized that individuals who produce more English responses when naming the items of the BNT, will score better, and name faster, than those who use fewer English responses. This assumption is based on the fact that the BNT was originally created and normed on a culturally and linguistically different (English-speaking) population. As some items (eg. *hammock*) may not be commonly used in Malta it is expected that some responses may be more easily accessed in English rather than Maltese as the words may have been experienced/learned in an English-speaking context (e.g. through television or reading).

[4] A backward span involves executive control to hold in memory and then to recite in reverse (Bialystok et al 2007).

Based on the research findings described above, it is hypothesized that in the present study cognate words will be retrieved faster and with more correct responses than non-cognate words, as the competition between the two translation equivalents would be weaker.

2.7 THE CURRENT STUDY

This study is concerned with analysing the performance of Maltese speakers on the Boston Naming Test. The choice of the BNT was made predominantly because of the huge amount of research that has been done using this test on a variety of cultures and populations of different ages and languages, including monolingual and bilingual populations. Further, it has been widely used to study both *normal* and impaired lexical retrieval both in research and clinical settings. Considering that the BNT was originally developed on a monolingual community in USA, it is expected that some of the items in the BNT may not be culturally appropriate for the Maltese population, which is largely bilingual and culturally different from USA. Likewise, the findings of this study are expected to differ significantly from the original normative data of the BNT.

2.7.1 Aims of this study

The first part of the study aims to obtain information regarding the performance of a sample of healthy Maltese-speaking adults on the Boston Naming Test and to establish whether it is an appropriate tool for use with the Maltese-speaking population. Naming accuracy and naming latency will be measured for Maltese-speaking adults of different age groups and different educational levels. Preliminary normative data will be produced and the relationship between ageing and the skill of naming will be investigated with particular emphasis on the changes that may occur in naming as ageing takes place. The second part of the study will investigate the performance of a sample of Maltese people with aphasia on the Boston Naming Test. Naming accuracy, the type of errors produced, and the appropriateness of the BNT for assessing Maltese people with aphasia will be among the areas of analysis.

2.7.2 Questions that need to be answered from part 1 of the study

It is proposed that a number of interesting results will be obtained from the data collected during the study. Descriptive statistics for different variables will be worked out, and global means and standard deviations will be worked out for score, reaction time and number of English responses. Two different areas of analysis will be necessary. The first type of analysis will focus on the performance of the different respondents on the BNT. The possible influence of demographic variables (e.g. age, education, gender, language use) on naming ability will be analysed (see Chapter 4). The second area of study will look at the specific items (the 60 pictures) included in the BNT and the word variables that may affect naming accuracy, latency and the number of English responses for each item. Such variables include name agreement, age of acquisition, imageability, familiarity, word length, cognate status, animacy and word initial fricative/affricate. Finally, the interaction between demographic and word variables will be investigated in order to reach conclusions regarding which word variables influence naming ability in different age groups and educational levels, for example (see Chapter 5). The following chapter (Chapter 3) provides a detailed description of the samples and data available, as well as the procedures used for data collection and analysis. The results and analysis of part 1 of the study will be discussed in Chapter 6.

2.7.3 Questions that need to be answered from part 2 of the study

Descriptive statistics for different variables will be worked out, and global means and standard deviations will be worked out for score. The number correct will be calculated for each person with aphasia and for each item in the BNT. An error profile for each person will ensue and will be analyzed in terms of the variables that may affect naming and the level of breakdown in spoken word production. A separate methodology chapter (Chapter 7) will describe the methods used to acquire the data from 18 people with aphasia and the procedure used during the analysis. Chapter 8 will report the results of Part 2 and a second discussion chapter will follow (Chapter 9). The final two chapters will provide recommendations and conclusions to the study.

2.7.4 Possible outcomes

Age, education and bilingualism are among the possible factors that may influence the performance of the Maltese-speaking respondents in this study. As seen in other studies a decline in naming ability may be evident with increasing age. Further, educational level may affect naming scores as has been demonstrated by other studies. The effect of bilingualism may also influence naming as bilinguals have been found to perform less well on naming activities.

Considering the bilingual context of the Maltese islands, a high number of English responses are highly predicted, in both parts of the research, especially since language-switching is common in Malta. Further, an amount of English responses is predicted because of the possibility that the material being used may be culturally biased. Hence, participants may use English as they may not be familiar with the items on a firsthand basis, but may have learned about them through reading and watching television - activities which are often carried out in English, especially in more highly educated adults.

Therefore, there may need to be some modifications regarding administration and scoring of the test. As this is a translation, it is possible that some items may have more than one possible correct response as different cultures may use more than one word to refer to the same object. Adhering to the strict guidelines of scoring may produce lower scores, which do not necessarily reflect the true naming abilities of the individual. Worrall et al (1995) and Cruice et al (2000) stressed that synonym responses should not be penalised/scored as incorrect – when administering the BNT on populations other than it was normed on. In the Greek version of the BNT the authors adopted this (possible responses) scoring method for four of their items (Patricacou et al 2007). Further, the "alternative-scoring" method (i.e. either language being scored as correct) has been suggested to be used with bilingual populations in order to achieve a better reflection of true abilities (Kohnert et al 1998).

As seen in other studies, when translating or adapting the BNT to produce a new version for other populations, some items may need to be omitted or replaced. However, care should be taken to select words of similar word variables (eg. with regards to AoA, familiarity), where possible, and from the same semantic categories (Patricacou et al 2007). Furthermore, the sequence of the test presentation may also require modification. The items in the test may need to be re-ordered depending on the frequency of correct responses obtained during the current study (e.g. Allegri et al 1997; Mariën et al 1998; Tallberg 2005) rather than according to word frequency counts.

2.8 CONCLUSION

Every population has its specific linguistic and cultural differences, and these will surely affect the performance of different individuals on the same test. The findings obtained from this study should demonstrate the potential effects that different variables may have on different individuals' ability to name pictures.

The results obtained from this study should serve as the foundation for further research in the area of spoken word production in the Maltese population. This research should provide normative data for different age and education groups and should result in the production of a standardised naming test for use with Maltese-speaking adults. This would be a valuable tool in the clinical situation, providing clinicians with a quick and easy-to-administer test, useful particularly with people who have had strokes, head injuries and those with dementia. In combination with other assessments the test will enable clinicians to make decisions regarding *normal vs. impaired* word retrieval skills. The results will then help clinicians to design appropriate therapy programmes according to the type of error/response patterns that occur. In particular, it is hoped that this study will provide valuable information about the way in which naming deteriorates as we age and how the Maltese bilingual context may affect one's performance on a naming test.

Chapter 3

Methodology 1

People without aphasia

3.0 INTRODUCTION

This chapter describes the methodology used during the first part of the study on healthy Maltese speakers. The methods used in the second part of the study are described in a separate chapter (Chapter 7, Methodology 2).

The Boston Naming Test (Kaplan et al 1983), or the BNT, has been used extensively for both research and clinical purposes. There are numerous studies that utilise the BNT to investigate a whole range of areas related to naming and cognitive function. Since the BNT was originally developed with norms from the USA, researchers from different countries have collected norms for specific populations, such as Patricacou (2007) for the Greek population and Mariën et al (1998) for Dutch-speaking Belgian elderly.

The use of the BNT in research is not solely limited to investigating the suitability of the test and to producing normative data for a variety of populations. Rather, the BNT has been used in studies to determine the influence of different variables on the skill of lexical retrieval (e.g. Randolph et al 1999). Studies have focused on word finding ability in normal ageing, in bilingual populations and in the presence of neurological impairments. Other researchers have used the BNT to identify the effects of common medical conditions (such as hypertension and diabetes mellitus) on word finding (Albert et al 2009). The use and validity of shorter versions of the BNT (such as the odd/even 30-item short forms or the 15-item shorter forms) have also been studied in different populations (e.g. Fisher et al 1999; Hobson et al 2009).

Relatively similar methods are used in different studies. However, the samples utilised to collect the data on BNT performance differ in a number of ways. The ages and education levels of each group of participants vary. While some studies focus on elderly participants (e.g. Neils et al 1995), other studies use samples of varying age groups (e.g. Allegri et al 1997). The number of participants taking part in each study also varies from 20 participants (e.g. Barker-Collo 2007) to over one thousand participants (e.g. Zec et al 2007a, 2007b). On the other hand, while some researchers use the BNT in its entirety and adopt the method of administration that is suggested by the authors of the test, other researchers choose to adopt a different method in order to elicit the information necessary for the particular study. Furthermore, while some studies focus only on naming accuracy, others also measure naming latency to investigate naming performance.

The present study has adapted the research methodology, and the BNT administration and scoring methods from the procedures used by a variety of different researchers in the area of lexical retrieval, including Mariën et al (1998), Bird et al (2001), Székely et al (2003), Patricacou (2007), etc.. The following is a description of the aims of the study, the participants, the methods and the materials used in the present study.

3.1 THE AIMS OF THE FIRST PART OF THE STUDY

As described in the preceding chapters, the present study aimed to analyse the performance of Maltese-speaking adults, of varying ages and educational levels, on a widely used confrontational naming task, the Boston Naming Test (Kaplan et al 1983), and, particularly, to analyse naming ability in normal ageing. Both responses and reaction times were recorded in the main study and the results were analysed and correlated with a number of demographic variables (including age, education, gender) and word variables (such as age of acquisition, imageability, familiarity and word length). Recommendations for using the BNT on Maltese speakers will be made, with suggestions regarding modifications in test administration, test scoring and test interpretation. Finally, normative data will be presented in order to make the best interpretation of the modified version of the BNT for Maltese speakers.

3.2 DIFFERENT SAMPLES

Four different samples were required for the purpose of this part of the study. Since there was no existing data on the performance of Maltese individuals on the BNT, and no information about the variables relating to the material itself, it was necessary to collect information on different aspects of the test material in order to be able to analyse the performance of the Maltese respondents who participated in the main study.

Firstly, information on name agreement for each of the items in the test was collected. This data was gathered from Sample A and was also useful to obtain a target name for each of the 60 items of the BNT. Next, information about three important stimulus-related variables was obtained through subjective ratings from Sample B. These variables were age of acquisition, familiarity and imageability of the words included in the test. Sample C included 50 respondents aged between 30 and 85 years of age, of varying levels of education. The information provided by this sample was analysed both for naming accuracy and naming latency. Finally, a sample of 10 university students (Sample D), aged between 20 and 23 years, was added to the main study for comparison purposes. Both naming accuracy and naming latency were measured from the students' responses too. The data collected from Samples C and D (the main study) was grouped together in one file, such that the analysis on demographic variables included 60 respondents between 20 and 85 years of age.

The following sections describe the different phases of the first part of the research, with information about each sample of participants.

3.3 SAMPLE A

3.3.1 Aims of Sample A

The main aim of this part of the study was to provide preliminary information on the suitability of the chosen test (the BNT) for the Maltese population. Several questions needed to be answered. First and foremost, it was necessary to know whether the items included in the BNT were generally familiar to the Maltese sample. Further, it was also essential to find out whether the pictures used in the BNT were easily recognised and associated with the original concept intended by the picture.

Name agreement was another important aspect that required attention. As the original BNT (Kaplan et al 1983) was normed on a diverse linguistic and cultural population, it was necessary to acquire information about the name that is given to each test item by the Maltese participants, and to decide whether there was in fact name agreement for each item. Moreover, in the case where more than one name was given for a particular item by the different participants, it was necessary to identify the dominant word for each item. The dominant word referred to the word that was used by the majority of the participants in Sample A. This dominant word was then considered to be the *target word* for the main study (Sample C and Sample D). When items were given more than one possible response, it was also necessary to decide on what could be regarded as alternative, but acceptable, names for the target word.

3.3.2 Participants of Sample A

Sample A included 24 individuals aged between 20 and 60 years, and with 13 or more years of education. The mean age of the participants of Sample A was 37.21 years (SD = 11.29) and the mean years of education 17.12 years (SD = 2.35). 37.5% were males and 62.5% were females. All were native speakers of Maltese.

The participants of Sample A were all independent in their daily living skills and they reported no history of neurological or cognitive disorders. All of them were volunteers and they were recruited from acquaintances of the researcher and through word of mouth. The term *participant/s* was used to refer to individuals who formed part of Sample A in order to differentiate them from individuals who participated in the main study (*respondents*).

3.3.3 Material of Sample A

The material used in this part of the study was a hard copy of the Boston Naming Test (Kaplan et al 1983). The publishers of the test, Pro-Ed. Inc., kindly provided the researcher with a copy of the test for research purposes.

Prior to meeting with and administering the BNT on the participants it was necessary to translate each of the 60 English words of the test into Maltese. A professional translator carried out this work. The *Maltese-English dictionary (*Aquilina 1987-1990)*, English-Maltese dictionary* (Aquilina 1999-2000) and *Id-dizzjunarju Malti u teżawru ta' Malti mħaddem* (Serracino-Inglott 2000) were used in this process. A final translation of each English word was drawn, listing all the possible responses for each word, including all synonyms. The English equivalents were sometimes included as possible responses where it was felt necessary by the translator. For example, although the Maltese word for *mushroom* is 'faqqiegħ'*,* the Maltese tend to use the English word *mushroom* more frequently. Likewise, the Maltese tend to use the word *toothbrush* more frequently than the Maltese equivalent 'xkupilja tas-snien'. On the other hand, some words have more than one possible correct translation. For example, both 'moxt' and 'pettne' are correct translations for the English word *comb*, although the former is used more by the older generations. Other items in the test tended to be associated with a greater number of possible responses. The translation for *racquet* produced five different possible Maltese and English responses. A response form, including all the possible responses for each item, was drawn up to be used by the researcher while administering the test on each participant. Following the preliminary study on Sample A, this response form was then amended to differentiate between target responses and alternative responses. The amended response form was used during the Main Study with Samples C and D (see Appendix A).

3.3.4 Method of administration

The researcher tested each of the participants of Sample A individually. An explanation of the aims of the study as well as the method of administration was provided to each participant prior to the actual administration of the BNT. The participants were told that a set of 60 pictures would be shown to them one by one, and that they were expected to name each picture with a single word whenever possible. They were encouraged to say the first word that came to mind when they saw the picture, but they would be allowed to correct their response if they needed to do so.

A short questionnaire regarding personal information was filled in by the researcher following the explanation. Personal details that were noted down included age, gender, level of education, occupation, and history of neurological and/or cognitive disorder.

The original method of administration of the test was slightly altered for the purpose of the study. The authors of the BNT (Kaplan et al 1983) recommend starting at item 30 when the test is used with non-aphasic individuals. However, during this study all of the 60 items were used in the original sequence of the BNT. Each picture was shown to each participant and he or she was requested to produce a name for the item. No time limit was given, although the authors suggest a 20 second time limit. However, if a participant showed that he or she was having some difficulty with producing a name, a semantic cue was provided. In the cases when the participants still could not provide a name they were asked to choose one of the following reasons:

> I can recognise the item, but I have *forgotten* the name for it (*I* – ***i**nsejt* meaning "I forgot");
> I can recognise the item, but I don't know its name (*MN* – ***m**a **n**afx* meaning "I don't know";
> I cannot recognise the item in the picture (*MJ* – ***m**a **j**ingharafx* meaning "not clear").

3.3.5 Recording the data collected from Sample A

The researcher transcribed each response for each item orthographically together with any comments regarding the participants' ability to name particular items, whenever necessary. The codes *I, MN* and *MJ* were used as described above. The procedure was recorded on a digital audio recorder in order to have the possibility of checking responses whenever this was required.

When all the participants were tested, the data collected was transferred to a Microsoft Excel file to work out percentages for each possible response and to determine the dominant response for each item. The dominant response for each item was the response that was produced by the majority of the participants of Sample A. This response was then considered as the target response during the Main Study. Name agreement was also determined from the percentage of participants who produced the same response for each item. An item was said to have high name agreement if at least 70% of the participants agreed on the same response for that particular item. The data was then transferred to an SPSS file (Statistical Package for the Social Sciences, version 16, 2007) for further analysis and statistical procedures. The available data included both information about each participant (including scores), and data about each BNT item (including the number of correct responses for each item). The results of this part of the study are reported in Chapter 4 (section 4.1).

3.3.6 Strict and lenient correct

The responses which matched the dominant response were termed as strict correct, while those responses which matched one of the alternative/acceptable responses were termed as lenient correct. Thus, two different scores were worked out for each participant of Sample A: a strict score and a lenient score. A strict score refers to the number of dominant responses produced by the participant and reflects the level of name agreement. Only the dominant responses were scored as correct in strict scoring. On the other hand, a lenient score refers to the total number of alternative responses produced by the participant (including dominant responses too). Therefore, if a response matched the dominant response or was one of the alternative responses it was scored as correct in lenient scoring.

Spontaneous self-corrections were recorded as correct in Sample A. Semantic cues were given to Sample A participants only, and they were provided only if visual misinterpretation was suspected. Accurate responses that followed a semantic cue were also given credit as a correct response, following the procedure of the original BNT. The strict and lenient scores of each participant are presented in Chapter 4 (Table 4.1).

3.3.7 English responses

English responses were regarded as possible responses, except when the English response was considered to be a dominant response. Thus, when a participant produced an English word in the case of a Maltese dominant response, the item was scored as correct in lenient scoring, but not in strict scoring. On the other hand, if the English word was the dominant word, an English response was scored correct in both strict and lenient scoring. Cognate words[1] (such as *camel* and 'ġemel', and *pyramid* and 'piramida') were both considered as dominant/target responses and were both scored as correct in strict scoring. Cognate pairs were treated as strictly correct when working out item correct responses, when working out participant scores and when working out

[1] "*Cognates* are those translation words that have similar orthographic-phonological forms in the two languages of a bilingual ….; *noncognates* are those translations that only share their meaning in the two languages" (Costa et al 2000, p.1285).

name agreement percentages. This means that, for example, both the English word *pyramid* and the Maltese word 'piramida' were scored as correct in strict scoring.

3.3.8 Complete and reduced sets

When the BNT was administered on Sample A, a considerable number of the items had low Name Agreement. Low Name Agreement implied that the same item generated a number of alternative responses. This created a difficulty during analysis, as alternative items could not be included in the analysis of word variables. Alternative responses for the same item would have different age of acquisition, word length, and different accuracy and latency results and so on. Therefore, two different sets of data were produced. The first set consists of the whole set of 60 items (a complete set), and the second set includes only those items which achieved at least 70% name agreement (a reduced set of 38 items) from Sample A. The reduced set is presented in Chapter 4 (Table 4.2). For both the complete and the reduced sets of data two separate files were produced. One file consists of the data and scores for each *participant* and another file consists of data on each *item* of the BNT. A similar procedure was carried out when recording the data collected from Samples C and D.

3.3.9 Analysis of data collected from Sample A

The data collected from Sample A was used to identify the target responses for the items of the BNT and to work out the percentage of name agreement for each item. Microsoft Excel was used to work out percentages and target responses. This information was compared to the percentage of correct BNT responses found by other studies on different populations (see Chapter 4, Tables 4.4 and 4.5) and to the results obtained from the main study of this project (see section 4.3.2).

3.4 SAMPLE B

3.4.1 Aims of Sample B

Sample B was used to collect subjective ratings on three word variables, namely age of acquisition (AoA), familiarity and imageability (see section 2.4.2 in Chapter 2 for a definition of each variable). These variables are said to reflect the personal experiences of the individuals who use the language and, hence, they have been found to affect both reaction time and naming performance (Stadthagen-Gonzalez and Davis 2006). Since no data on these variables was available from Maltese speakers, it was necessary to collect subjective ratings too as part of the study. The means of each item for each variable would then be correlated with other data collected, including naming scores and reaction times.

3.4.2 Participants of Sample B

Subjective ratings were obtained from 20 undergraduate students reading courses at the University of Malta. The students were recruited from two faculties: the Communication Therapy Department of the Faculty of Health Sciences and from the Faculty of Education. The researcher sought permission from the Registrar of the University of Malta and from the respective Heads of departments for access to lists of student names. The students who volunteered to participate in the study were all reported to be native speakers of Maltese, or of Maltese and English. Their mean age was 20 years (*SD* = 0.97) and their mean years of education was 15 years (*SD* = 0.49). 75% of the students were females and 25% were males. The higher percentage of female students in sample B reflects the particular University courses that they were registered in. Traditionally more females are enrolled in health care and education courses.

3.4.3 Material of Sample B

All the possible names[2] (Maltese and English) for each of the 60 items of the BNT were listed and a form for the Subjective Ratings was produced. The possible names were translated by a professional translator (see section 3.3.3). The list consisted of 106 words, which were reordered randomly using Microsoft Excel in order to obtain the same list of words in a different random order for each variable.

Since all of the items in the test are pictures of objects, all of the names expected were concrete names. Thus, it was necessary to add a set of 20 abstract words to the list of

[2] Ratings are available for all possible responses except for 'mouth organ', which was the dominant response for 'harmonica'.

imageability ratings to avoid the possibility that students would rate all of the object names as highly imageable. The abstract words were selected from Cortese and Fugett (2004)[3] and were once again translated into Maltese by the translator. Table 3.1 lists the abstract words selected. The extra 20 words were added randomly to the list of imageability ratings only, resulting in a list of 126 words. Finally, a 14-page booklet was produced including forms for subjective ratings of the three variables and a brief explanation on rating each variable (see Appendix B).

TABLE 3.1

English word	Maltese translation	English word	Maltese translation
tact	tattika	curse	saħta
tempt	ħajjar	how	kif
cheap	rħis	where	fejn
cheat	qarraq	tease	nibex
clue	ħjiel	fact	fatt
creed	twemmin	from	minn
need	bżonn	know	jaf
my	tiegħi	boast	ftaħar
less	inqas	free	liberu
bluff	blaff	and	u

The extra 20 words that were added to the list of imageability ratings only

3.4.4 Procedure of Sample B

Since different researchers use somewhat different procedures for obtaining subjective ratings it is necessary to describe the method used in the current study.

For age of acquisition (AoA) the current study used a Scale of 1 – 7, as used by Bird et al (2001), Stadthagen-Gonzalez and Davis (2006) and Cortese and Khanna (2007). This scale was originally employed by Gilhooly and Logie (1980). Instructions for AoA ratings were adapted from Stadthagen-Gonzalez and Davis (2006) and Cortese and Khanna (2007).

[3] Available from //www.psychonomic.org/archive/uploads/Cortese-BRMIC-2004.zip (see article Cortese and Fugett 2004).

The present study focused on word familiarity rather than object familiarity or picture familiarity. Instructions were based on those used by Stadthagen-Gonzalez and Davis (2006). A scale of 1 – 5 was used for familiarity ratings.

Instructions for rating imageability were based on the works of Cortese and Fugett (2004) and Stadthagen-Gonzalez and Davis (2006), who define imageability in terms of the ease or difficulty with which words arouse mental images, and on the work of Alario et al (2004) who define this variable in terms of the number of images that are aroused (and not just in terms of how easy it is to arouse a mental image). This was necessary to ensure that the students were provided with enough information to understand the meaning of this variable. The Scale of 1 – 5 was adopted here, as used by Alario and Ferrand (1999) and Alario et al (2004).

The researcher met the students in 3 groups in a quiet room. An explanation of the general aims and procedure of obtaining subjective ratings was given to the students, and an outline of what was expected from them followed. The Subjective Ratings booklet (Appendix B) was then given out to each student. The booklet includes definitions and instructions for variables age of acquisition, familiarity and imageability. The students were given some time to fill in some personal details (such as age, area of living and name of course). Subsequently, the definition of age of acquisition and the instructions for rating this variable were read to the students. Details on rating this variable were provided, with examples to make sure that the students understood the procedure. Once the explanation was over, the students were asked whether they required further clarifications and then, they were instructed to start rating the list of words individually (without consulting their peers) by marking the appropriate boxes on the forms provided. This procedure was repeated for the familiarity and the imageability ratings, making sure that each student understood the meaning of the variable and how to rate each word before starting.

Emphasis on explanation was made throughout this part of the data collection as none of the students had ever participated in this type of exercise and none of them had any

TABLE 3.2

AGE OF ACQUISITION SUBJECTIVE RATINGS									
	English Word	**Maltese Target**	**Mean AoA rating**	**SD**		**English Word**	**Maltese Target**	**Mean AoA rating**	**SD**
1.	bed	sodda	1.5	0.61	31.	rhinoceros	*rhinoceros*	4.5	1.07
2.	tree	siġra	1.6	0.60	32.	acorn	*acorn*	4.6	2.15
3.	pencil	lapes	1.8	0.55	33.	igloo	iglù	3.8	1.59
4.	house	dar	1.6	0.60	34.	stilts	*stilts*	5.6	1.36
5.	whistle	suffara	2.6	0.76	35.	dominoes	*dominoes*	3.5	1.54
6.	scissors	mqass	2.5	0.89	36.	cactus	kaktus	3.7	1.13
7.	comb	pettne	2.8	1.16	37.	escalator	*escalator*	4.2	1.57
8.	flower	fjura	1.5	0.60	38.	harp	arpa	4.0	1.12
9.	saw	serrieq	3.7	0.99	39.	hammock	*hammock*	5.7	1.30
10.	toothbrush	*toothbrush*	2.1	0.76	40.	knocker	ħabbata	5.1	1.47
11.	helicopter	ħelikopter	2.4	0.88	41.	pelican	pelikan	4.8	1.20
12.	broom	xkupa	2.8	1.02	42.	stethoscope	*stethoscope*	5.7	1.66
13.	octopus	qarnita	2.8	0.91	43.	pyramid	piramida	3.7	1.03
14.	mushroom	*mushroom*	3.2	1.39	44.	muzzle	sarima	4.4	1.52
15.	hanger	spalliera	3.1	0.69	45.	unicorn	*unicorn*	4.3	1.74
16.	wheelchair	*wheelchair*	3.5	0.76	46.	funnel	lembut	4.0	1.30
17.	camel	ġemel	2.8	1.12	47.	accordian	*accordian*	4.6	1.73
18.	mask	maskra	2.4	0.68	48.	noose	ingassa	5.4	1.54
19.	pretzel	*pretzel*	5.5	1.90	49.	asparagus	*asparagus*	6.1	1.05
20.	bench	bank	3.0	0.86	50.	compass	kumpass	4.8	1.07
21.	racquet	*racquet*	4.9	1.41	51.	latch	*lock*	4.9	1.39
22.	snail	bebbuxu	1.6	0.68	52.	tripod	*tripod*	6.4	0.77
23.	volcano	vulkan	3.1	0.79	53.	scroll	skroll	5.2	1.55
24.	seahorse	*seahorse*	3.7	1.69	54.	tongs	*tongs*	6.0	1.19
25.	dart	dart	3.6	1.10	55.	sphinx	*sphinx*	5.2	1.54
26.	canoe	kenura	5.3	1.37	56.	yoke	madmad	6.6	0.65
27.	globe	globu	3.8	1.06	57.	trellis	kannizzata	5.1	1.29
28.	wreath	girlanda	4.7	1.61	58.	palette	*palette*	4.6	1.54
29.	beaver	*beaver*	4.3	1.48	59.	protractor	*protractor*	5.0	1.26
30.	harmonica	*mouth organ*	not available	not available	60.	abacus	*abacus*	4.3	1.81

SCALE	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	0-2 years	**2-4 years**	**4-6 years**	**6-8 years**	**8-10 years**	**10-12 years**	**13+ years**

TABLE 3.3

FAMILIARITY SUBJECTIVE RATINGS									
	English Word	**Maltese Target**	**Mean Familiarity rating**	**SD**		**English Word**	**Maltese Target**	**Mean Familiarity rating**	**SD**
1.	bed	sodda	5.0	0.00	31.	rhinoceros	*rhinoceros*	2.7	1.08
2.	tree	siġra	4.8	0.41	32.	Acorn	*acorn*	3.0	1.05
3.	pencil	lapes	5.0	0.00	33.	Igloo	iglù	3.4	1.04
4.	house	dar	5.0	0.00	34.	Stilts	*stilts*	2.1	1.29
5.	whistle	suffara	4.4	0.67	35.	dominoes	*dominoes*	3.9	1.29
6.	scissors	mqass	4.7	0.59	36.	Cactus	kaktus	4.4	0.75
7.	comb	pettne	4.3	0.73	37.	escalator	*escalator*	4.3	0.72
8.	flower	fjura	4.85	0.37	38.	Harp	arpa	3.2	1.14
9.	saw	serrieq	3.9	0.75	39.	hammock	*hammock*	2.8	1.21
10.	toothbrush	*toothbrush*	4.8	0.41	40.	knocker	ħabbata	2.8	1.01
11.	helicopter	ħelikopter	4.2	0.83	41.	Pelican	pelikan	3.0	1.03
12.	broom	xkupa	4.7	0.59	42.	stethoscope	*stethoscope*	3.5	1.10
13.	octopus	qarnita	4.3	0.72	43.	pyramid	piramida	3.8	0.97
14.	mushroom	*mushroom*	4.8	0.44	44.	Muzzle	sarima	1.5	0.94
15.	hanger	spalliera	4.7	0.47	45.	Unicorn	*unicorn*	2.3	0.64
16.	wheelchair	*wheelchair*	4.2	0.81	46.	Funnel	lembut	3.1	0.94
17.	camel	ġemel	4.0	0.92	47.	accordian	*accordian*	3.0	1.38
18.	mask	maskra	4.4	0.75	48.	Noose	ingassa	2.8	1.25
19.	pretzel	*pretzel*	2.0	1.28	49.	asparagus	*asparagus*	3.0	1.39
20.	bench	bank	4.6	0.69	50.	compass	kumpass	3.7	0.98
21.	racquet	*racquet*	2.9	1.41	51.	Latch	*lock*	2.8	1.15
22.	snail	bebbuxu	4.4	0.68	52.	Tripod	*tripod*	1.9	0.97
23.	volcano	vulkan	3.9	0.85	53.	Scroll	skroll	2.7	1.30
24.	seahorse	*seahorse*	3.5	0.95	54.	Tongs	*tongs*	3.6	1.10
25.	dart	dart	3.9	1.02	55.	Sphinx	*sphinx*	3.0	1.08
26.	canoe	kenura	2.6	0.83	56.	Yoke	madmad	1.6	0.88
27.	globe	globu	3.5	0.89	57.	Trellis	kannizzata	3.4	1.23
28.	wreath	girlanda	3.1	1.33	58.	Palette	*palette*	3.5	1.19
29.	beaver	*beaver*	3.1	1.19	59.	protractor	*protractor*	3.9	1.04
30.	harmonica	*mouth organ*	not available	not available	60.	Abacus	*abacus*	2.9	1.31

SCALE	**1**	**2**	**3**	**4**	**5**
	very unfamiliar	**unfamiliar**		**familiar**	**very familiar**

TABLE 3.4

IMAGEABILITY SUBJECTIVE RATINGS									
	English Word	**Maltese Target**	**Mean IMG rating**	**SD**		**English Word**	**Maltese Target**	**Mean IMG rating**	**SD**
1.	bed	sodda	4.9	0.31	31.	rhinoceros	*rhinoceros*	4.1	1.23
2.	tree	siġra	5.0	0.00	32.	acorn	*acorn*	3.0	1.43
3.	pencil	lapes	4.9	0.31	33.	igloo	iglù	4.2	0.88
4.	house	dar	5.0	0.00	34.	stilts	*stilts*	1.9	1.37
5.	whistle	suffara	4.3	0.91	35.	dominoes	*dominoes*	4.0	1.36
6.	scissors	mqass	4.6	0.69	36.	cactus	kaktus	4.5	0.83
7.	comb	pettne	4.2	0.95	37.	escalator	*escalator*	4.5	0.76
8.	flower	fjura	5.0	0.00	38.	harp	arpa	3.9	1.17
9.	saw	serrieq	4.2	1.01	39.	hammock	*hammock*	2.5	1.67
10.	toothbrush	*toothbrush*	4.9	0.31	40.	knocker	ħabbata	3.6	1.36
11.	helicopter	ħelikopter	4.7	0.59	41.	pelican	pelikan	3.4	1.23
12.	broom	xkupa	4.6	0.76	42.	stethoscope	*stethoscope*	3.5	1.67
13.	octopus	qarnita	4.1	1.17	43.	pyramid	piramida	4.5	0.83
14.	mushroom	*mushroom*	4.7	0.67	44.	muzzle	sarima	1.6	1.35
15.	hanger	spalliera	4.5	0.83	45.	unicorn	*unicorn*	4.0	0.97
16.	wheelchair	*wheelchair*	4.5	0.76	46.	funnel	lembut	3.6	1.27
17.	camel	ġemel	4.3	0.91	47.	accordian	*accordian*	2.8	1.36
18.	mask	maskra	4.4	0.82	48.	noose	ingassa	2.6	1.31
19.	pretzel	*pretzel*	2	1.38	49.	asparagus	*asparagus*	2.6	1.31
20.	bench	bank	4.7	0.67	50.	compass	kumpass	4.0	1.36
21.	racquet	*racquet*	2.5	1.54	51.	latch	*lock*	2.8	1.29
22.	snail	bebbuxu	4.7	0.8	52.	tripod	*tripod*	2.4	1.39
23.	volcano	vulkan	4.4	0.93	53.	scroll	skroll	3.0	1.49
24.	seahorse	*seahorse*	3.8	1.29	54.	tongs	*tongs*	3.6	1.39
25.	dart	dart	4.1	0.97	55.	sphinx	*sphinx*	2.8	1.51
26.	canoe	kenura	3.0	1.26	56.	yoke	madmad	1.8	1.37
27.	globe	globu	4.4	0.75	57.	trellis	kannizzata	2.9	1.68
28.	wreath	girlanda	3.3	1.52	58.	palette	*palette*	3.3	1.30
29.	beaver	*beaver*	3.2	1.36	59.	protractor	*protractor*	4.1	1.10
30.	harmonica	*mouth organ*	not available	not available	60.	abacus	*abacus*	2.6	1.50

	1	**2**	**3**	**4**	**5**
SCALE	**low imageability**				**high imageability**

theoretical experience about obtaining subjective ratings. Not only was it necessary to explain the procedure of acquiring subjective ratings, but also to ensure that the meaning of each variable was understood.

3.4.5 Recording and analysing the data collected from Sample B

The data collected during this part of the study was then transferred to a Microsoft Excel file, and average ratings and standard deviations were worked out. Minimum and maximum ratings for each item were also worked out. Tables 3.2, 3.3 and 3.4 present the mean ratings and standard deviations for each target response in a separate table for each word variable.

Sample B was only used to collect data on three stimulus variables, namely age of acquisition, familiarity and imageability. Other predictor variables will be used in the study. These consist of both stimulus variables (word length, word initial fricative/affricate, animacy and cognate status) and demographic variables (age, gender, education, occupation, language and reading habits). See Sections 3.5.5.2 and 3.6.

3.5 THE MAIN STUDY : SAMPLES C AND D

3.5.1 Aims of the Main Study

Sample C included respondents aged 30 to 85 years from the general population while Sample D consisted of a group of 10 University students in their twenties. For analysis purposes these two samples were grouped together as one sample and they were referred to as the Main Study. Samples C and D were regarded as the main study as they provided essential information about the typical naming/lexical retrieval abilities and reaction times of *healthy* Maltese-speaking adults in the age range of 20 to 85 years and of varying levels of education on the Boston Naming Test. Results from this part of the study were then correlated with a number of demographic variables (e.g. age and years of education) and with word variables (such as word length and imageability).

3.5.2 Criteria for inclusion into the main study

In order to acquire data on typical abilities associated with particular age groups and education levels, for example, it is essential to make sure that the data collected is derived from individuals who are free from any cognitive, neurological and/or psychiatric disorders that may influence performance on the task that is being studied. Thus, it was necessary for prospective respondents to satisfy a set of inclusion criteria. These criteria were drawn from a variety of different studies on naming tests.

During the early recruitment stage, two necessary criteria were kept in mind, namely that the respondents for the main study were to be *healthy* and *Maltese-speaking*. Furthermore, care was taken to recruit individuals from a wide cross-section of the Maltese population, in terms of age, gender and level of education. This was necessary in order to avoid any bias towards one gender, age group, etc..

3.5.2.1 Healthy Maltese adults

One of the criteria necessary for inclusion in the study was that the respondents were independent in daily living activities and lived in the community, rather than in homes or shelters for the elderly. Further, the respondents were asked a number of health questions (see section 3.5.2.3) and they were also screened for neurocognitive impairment (see section 3.5.2.4). These measures were taken to ensure that the respondents participating in the final study were healthy, and to limit the possibility of including any individual with signs of neurological, cognitive or psychiatric disorders that may affect naming ability.

3.5.2.2 Maltese-speaking adults

Another necessary prerequisite was that participants were to be native Maltese-language speakers and that they were to be predominantly Maltese speakers in their everyday life. Please refer to section 2.6.1, in Chapter 2, for a description of the language/s used in Malta.

From their in-depth study on the use of language in different environments, Sciriha & Vassallo (2006) concluded that overall, in Malta, Maltese is predominantly used in conversational settings, and that "…. Maltese is invariably the most widely used language with different people and in different situations" (Sciriha & Vassallo 2006, p.102). Therefore, for this reason the researcher decided to focus on Maltese-speaking Maltese adults only in the current study and to exclude the minority who use *English only* (E) and *English but with some Maltese words* (Em) (as described in section 2.6.1.1 of Chapter 2). Notwithstanding this, it was still necessary to acquire information regarding the pattern of language acquisition, as well as the extent of exposure to, and use of, the two official languages (Maltese and English), depending on family, social and occupational contexts (See Appendix C for Language History Questionnaire).

3.5.2.3 Questionnaire

Prior to administering the BNT the respondents completed a questionnaire with the researcher. The following is a description of this questionnaire.

A set of questions relating to personal, demographic details and health history were formulated. The questions were aimed at (1) ensuring inclusion into the current study, and at (2) gaining essential personal details that would eventually be correlated with the results obtained from the study (see Appendix D). As noted above, the questions were derived from the works of different authors, such as Mariën et al (1998), Tallberg (2005) and Patricacou et al (2007), and included questions such as age, gender, area of residence, level of education, past and present occupation, native language, handedness and reading habits[4]. The respondents were also asked about their medical history.

3.5.2.4 Screening measure

In the literature regarding the BNT researchers make different decisions as to which additional tasks or tests are performed apart from the BNT. These decisions may arise from the availability of standardised language and neuropsychological tests on the particular population, as well as the specific aims of the research at hand. Frequently these additional tests are used as screening measures, with a definite "cut-off" score, that enables the researcher to decide on who will be included in the final study. The most widely used test that has been used for this purpose is probably the MMSE

[4] Padakannaya et al (2002, cited in Mansur et al 2006) found that naming abilities were influenced by individuals' reading habits apart from their level of education.

(Folstein et al 1975). For example, Kavé (2005) and Patricacou et al (2007) used the Hebrew and the Greek version of the MMSE respectively for screening purposes.

Unfortunately, in Malta we are still very limited in the availability of neuropsychological tests that are standardised on the Maltese population and, in fact, a Maltese version of the MMSE is not yet available. An adapted version of the MMSE might have had additional difficulties in Malta. For example, the sections on reading and writing of the MMSE would not have been completed with all respondents since there is a high rate of illiteracy in Malta, especially in those over 60 and 70 years of age (National Statistics Office 2006)[5]. However, numbers are usually well recognized and written. Furthermore, the section on Orientation to Place may be somewhat difficult to translate as Malta, being such a small country, does not have specific names for states/provinces or county. Thus, for the purpose of the current research, it was necessary to identify a different screening measure that would *not* be culturally and/or linguistically biased when applied to different populations. The *Clock Drawing Test* (CDT) fits this purpose effectively. The following sections describe the CDT in detail.

3.5.3 The Clock Drawing Test

The Clock Drawing Test has long been utilized to screen for cognitive and neurological dysfunction and it "meets all the criteria necessary for a cognitive screening instrument" (Shulman, 2000, p.559). It has the benefits of being a simple, quick, easy to administer and non-threatening test (Agrell & Dehlin 1998; Burns et al 2004) and it is much less affected by linguistic, ethnic, cultural and educational factors than the MMSE (Shulman 2000; Lin-Kiat Yap et al 2007). In fact, in emails to the researcher, both Royall and Heinik confirmed that clock drawing is much less influenced by culture and language than other commonly used neuropsychological tests (Sept 2008). Moreover, the Clock Drawing Test has been highly correlated with the MMSE (Agrell & Dehlin 1998). In general, normal clock drawing reasonably excludes cognitive impairment (ibid.). Therefore, the Clock Drawing Test was identified as the screening measure to be used in the current study. When combined with the general questionnaire and with the

[5] According to the 2005 Census of Population and Housing in Malta there were 26,121 (7.2% of the population) illiterate individuals over 10 years of age.

inclusion criteria (see section 3.5.2), the possibility of including individuals with any sign/s of cognitive and/or neurological impairments in the study was greatly minimized.

3.5.3.1 Processes involved in clock drawing

The performance on the Clock Drawing Test (CDT) has been shown to demonstrate a number of neurocognitive processes, as it reflects frontal and tempero-parietal functioning (Brodaty & Moore 1997; Burns et al 2004). It has, therefore, been used to screen for cognitive impairment and dementia, as well as to gain insight on one's visuo-spatial functioning (Agrell & Dehlin 1998). Several authors list the processes that are reflected in clock drawing. These include: auditory comprehension, visual memory and reconstruction, spatially-coded knowledge, constructive skills (as constructional apraxia may be an early sign of Alzheimer's disease (Agrell and Dehlin 1998)), visual neglect, numerical knowledge, motor programming and executive function (CDT is a culturally unbiased measure that reliably identifies impaired executive control function, which may be a first sign of dementia (Royall et al 2004)), semantic instruction (abstract thinking), inhibition of distracting stimuli (or inhibition of the tendency to be 'pulled' by perceptual factors), concentration and frustration tolerance (Agrell & Dehlin 1998; Shulman 2000; Lin-Kiat Yap et al 2007). Shah (2001) also states that the CDT may be valuable in identifying and localising brain lesions. In fact, apart from scoring the CDT, by using a point system, it should be noted that qualitative analysis may yield valuable information in clinical use. On the other hand, a simple scoring method with a clear "cut-off" score is essential for research purposes. This was clarified in an email to the researcher from Shulman (Sept 2008). Other positive aspects of the CDT are that it is a cost-effective test and that it allows easy graphical reviews of progress and deterioration over time (Shah 2001). Shulman (2000) also emphasises good inter-rater and test-retest reliability.

3.5.3.2 Potential effects on clock drawing

Some authors have investigated the potential effects of several factors on Clock Drawing, including age, education and gender. Seigerschmidt et al (2002) and Von Gunten et al (2008) both highlight the possible influence of age and education. Although a possible decline on performance may be seen over 70 years of age (Kozara & Cullum 1994 cited in Strauss et al 2006), Cahn & Kaplan (1997 cited in Strauss et al

2006) state that clock drawing remains fairly preserved between 70 and 90 years of age. Furthermore, education effects are more apparent in the lower scores (Ainslie and Murden 1993, cited in Agrell & Dehlin 1998). Although some controversy remains about these possible effects, one should note that they are less pronounced than in the MMSE (Shulman 2000).

3.5.3.3 Method of administration and scoring procedure

There are numerous methods of administering and scoring the Clock Drawing Test. Various authors (including Shulman 1986, 2000; Manos 1998; Royall et al 1998; Heinik et al 2004, and many others) have come up with their own method and interpretation of the CDT. However, comparison of the different methods has shown that despite the variations, the different versions of the CDT result in similar conclusions (Brodaty & Moore 1997; Shulman 2000). The Shulman method seems to be the best for distinguishing between individuals with dementia and *normal* individuals and it appears to be one of the most sensitive procedures (Schramm et al 2002). Further, it has been recommended as a valuable screening test for mild to moderate dementia, especially when combined with the MMSE (Schramm et al 2002; Burns et al 2004; Strauss et al 2006). Brodaty & Moore (1997) found the Shulman scale easy to follow and reliable with the "highest discriminatory power" (p. 623). The Clock Drawing Test has also been shown to correlate positively with commonly used tests of cognitive ability and it is possibly superior to the MMSE in predicting degree of executive function and functional status (Brodaty & Moore, 1997; Royall et al 2003; Samton et al 2005). Therefore, the CDT (especially the method developed by Shulman) "appears to provide a reasonable measure of cognitive functioning" (Strauss et al 2006, p. 978). The Shulman Scale (Shulman 2000; Shulman & Feinstein 2003) was the method chosen to be used as a screening test in the present study for the reasons stated above.

3.5.3.4 Method of instruction of the Shulman scale (Shulman 2000)

In this method, the individual is presented with a pre-drawn circle of 4 inches or 10cm in diameter and he or she is given the following instructions:

> *"This circle represents a clock face. Please put in the numbers so that it looks like a clock and then set the time to 10 minutes past 11."*
>
> (Shulman 2000, p. 550)[6]

3.5.3.5 Scoring procedure of the Shulman scale (Shulman 2000)

The scoring procedure involves a simple 0 – 5 point system. Table 3.5 gives a brief description of each point. Shulman (1986: 136) provides a more detailed explanation of the above clock errors. However, the point-system is reversed, such that a score of V was given to an inability to make a reasonable attempt at a clock (0 points), while a score of I was given for minor visuo-spatial errors (4 points). For the purpose of this research a "cut-off" score of 3 was used. This "cut-off" was recommended by Shulman in emails to the researcher (Sept 2008), and it was agreed that any respondent who scored 3 points or less would be excluded from the study.

TABLE 3.5

5 points	'Perfect' clock
4 points	Minor visuo-spatial errors
3 points	Inaccurate representation of 10 after 11 when the visuo-spatial organization is well done
2 points	Moderate visuo-spatial disorganization of numbers such that accurate denotation of '10 after 11' is impossible
1 point	Severe level of visuo-spatial disorganization
0 points	Inability to make any reasonable representation of a clock

Shulman (2000: 554)

Scoring procedure for Clock Drawing Test

3.5.4 Participants of the Main Study

The term *respondents* was used throughout the study to refer to the volunteers who took part in this main study (Samples C and D) in order to differentiate them from Sample A and Sample B participants.

The respondents were recruited from different settings, including the University of Malta, various companies, schools and elderly day centres. Care was taken to select respondents evenly with regards to age group, gender and education. Letters were

[6] The use of setting the time to 11:10 is of importance as it involves both visual fields (Shulman, 2000; Strauss et al, 2006).

addressed to the respective heads of departments including a description of the aims and procedures involved in the study as well as a request for volunteers. Permission for the researcher to visit the mentioned settings and to meet the volunteers was granted. All the individuals who accepted to participate were provided with a description of the study and their role in it before actually meeting the researcher. The local Ethics Committee of the University of Malta approved the study.

The Main Study consisted of 60 respondents, aged between 20 and 85 years. Their mean age was 52.70 years (*SD* = 20.93) and their mean years of education was 12.65 years (*SD* = 4.39), with a range of 0 to 19 years of education. The respondents were grouped into 2 main education categories and 6 age groups. 48.3% (N = 29) of the group had 0 to 12 years of education and 51.7% (N = 31) had 13 or more years of education. The respondents who formed part of the lower education group had minimal education or obligatory education only, while the higher education group consisted of individuals who had post-secondary education including diplomas and University degrees. A similar number of individuals (N = 29 and 31 respectively) were involved in each education group. Age groups consisted of the following categories: 20 – 29 years, 30 – 45 years, 46 – 55 years, 56 – 65 years, 66 – 75 years and 76- 85 years. Each age group consisted of approximately 10 individuals. 51.7% (N = 31) of the respondents were males and 48.3% (N = 29) were females.

Table 3.6 tabulates the number of respondents in each gender, age group and education category. The main study sample consisted of a similar amount of participants in each age group, education group and gender category. However, age and education were not contrasted orthogonally as the older age groups tended to consist of individuals with lower levels of education. This sample is a reflection of the sociological reality that exists in the Maltese islands, since older people (especially females) had fewer opportunities for education than the younger people of today.

Most of the respondents had manual types of occupations (40%). A further 5% of the 60 respondents had clerical jobs, 8.3% had a supervisory position, 23.3% had professional jobs, 6.7% had managerial positions and 16.7% were students.

All the respondents were native speakers of Maltese and all were exposed to some degree of English. Table 3.7 demonstrates the respondents' language experience from birth up to the secondary school years. 58.3% of the respondents used Maltese only (M) on a daily basis, 23.3% used Maltese with some English words (Me) daily and 18.3% used both Maltese and English (ME) daily. Individuals who used English with some Maltese words (Em) or English only (E) were excluded from the study (see section 2.6.1.1, in Chapter 2).

TABLE 3.6

			0 – 12 years of Education	**13 or more years of Education**	**Total**
Males	**Age Group**	20 – 29 years	0	5	5
		30 – 45 years	2	3	5
		46 – 55 years	3	3	6
		56 – 65 years	3	3	6
		66 – 75 years	4	0	4
		76 – 85 years	2	3	5
		Total	14	17	31
Females	**Age Group**	20 – 29 years	0	5	5
		30 – 45 years	2	3	5
		46 – 55 years	2	2	4
		56 – 65 years	2	2	4
		66 – 75 years	4	1	5
		76 – 85 years	5	1	6
		Total	15	14	29

The number of respondents in each gender, age group and education category

More than half of the respondents reported reading everyday (60%), 15% said that they read several times a week, 13.3% read once a week, 3.3% rarely read and 8.3% never read. All the respondents reported normal or corrected-to-normal visual acuity. Any respondent with a history of stroke, head injury, cerebral tumour, learning disability, epilepsy, history of concussion or loss of consciousness, psychiatric illness and/or alcohol/substance abuse were subsequently excluded from the study. Only respondents who achieved a score of 4 or 5 on the Clock Drawing Test (Shulman 2000) were included in the main study. Three respondents achieved a lower score and were subsequently excluded from the study.

TABLE 3.7

Language experience	Percentage of respondents
Predominantly exposed to Maltese at HOME and attended Maltese-speaking schools (primary & secondary school) **M/m**	68.3%
Predominantly exposed to Maltese at HOME and attended English-speaking schools (primary & secondary school) **(M/e)**	10%
Predominantly exposed to Maltese at HOME and attended Maltese-speaking primary school & English speaking secondary school **(M/me)**	8.3%
Predominantly exposed to Maltese at HOME and attended English-speaking primary school & Maltese-speaking secondary school **(M/em)**	6.7%
Exposure to Maltese & English at HOME and attended Maltese-speaking schools (primary & secondary school) **(ME/m)**	3.3%
Exposure to Maltese & English at HOME; attended Maltese-speaking primary school & English speaking secondary school **(ME/me)**	3.3%

Maltese-English language experience of respondents in the main study

3.5.5 Materials and procedure

All of the pictures of the BNT were scanned onto a laptop in the original sequence of the test. This was carried out by Technical staff at the Faculty of Education and Health Sciences, University of Limerick, Limerick, Ireland. The pictures appeared individually on a 14-inch Fujitsu S7020 computer screen in the same size of the original BNT pictures. The respondents were seated in front of the screen and viewed the pictures from a distance of approximately 80cm. The picture remained on the screen for 3 seconds and an interval of 8 seconds (with a blank screen) was allowed before the next picture appeared. Thus, a total of 11 seconds was allowed before introducing the next picture. Each picture was seen only once by the respondent. The procedure used was similar to that used by Howard et al (2006) in their study on semantic inhibition in picture naming. Howard et al (2006) allowed a more restricted time limit, however (2 seconds for viewing the picture and an interval of 500 milliseconds before showing the next picture).

3.5.5.1 Administration of the BNT

The researcher met all the respondents individually in a quiet, dimly lit room with no distractions. One session of approximately 45 minutes was allowed for each respondent. The respondents were provided with information about the study and about their involvement in it. An informed consent form was signed by each respondent. The respondents were asked a number of questions regarding age, education, health and language, and the researcher wrote down the information provided on the prepared forms (see questionnaires in Appendix C and D). The administration of the Clock Drawing Test (Shulman 2000) followed the questionnaires. Finally, the researcher described the administration of the BNT to the respondents. Care was taken to provide the exact same instructions to each respondent. The respondents were asked to name each picture that would appear on the screen:

- **clearly and loudly enough into the microphone provided;**
- **with a *single* word;**
- **with the *first name* that comes to mind;**
- **as *quickly* and *accurately* as possible;**
- **before the next picture appeared;**
- **without emitting any other sounds;**
- **to avoid coughing and hesitations (e.g. "er", "erm"); and**
- **without using articles, descriptions or comments.**

3.5.5.2 Recording the data

The researcher wrote down all the responses of the respondents orthographically on a record form that included all the alternative responses and the target name (see Appendix A) in bold. Notes on naming were written down on the form whenever necessary.

All the responses of each respondent were inserted into a Microsoft Excel File and coded according to the categories outlined in Table 3.8. These error categories were derived from other researchers' work on naming (e.g. Mariën et al 1998). Percentages for each of the categories listed were worked out automatically by Microsoft Excel. Thus, the percentage results for each category for each item were calculated. For example, for item 14 ('mushroom') 90% of the respondents produced the Target name 'mushroom', 4% produced an alternative response *'faqqiegħ'*, and 6% required a Phonemic cue.

The data was then transferred into two distinct SPSS files (Statistical Package for the Social Sciences, version 16, 2007). The first file consisted of demographic variables while another file included word variables.

TABLE 3.8

Target name	The dominant response of each item as worked out from Sample A
Possible responses	Alternative responses for each item as worked out from Sample A (including all synonyms and English translation equivalents)
Self correction	Immediate self correction
Subordinate error	Responses which belong to a more specific semantic category than the Target name, e.g. *cottage* for *house*
Superordinate error	Responses which belong to a broader semantic category than the Target name, e.g. *animal* for *beaver*
Coordinate error	Responses which belong to the same semantic category of the Target name, e.g. *rat* for *beaver*
Visual association	Reponses which show that the item was not recognized, e.g. *tape* for *whistle*
Semantic association	Responses which are semantically related to the Target, e.g. *Egypt* for *pyramid*
Phonological association	Responses which are phonologically related to the Target, e.g. *glove* for *globe*
Circumlocutions	Responses that describe the target word
Don't know	When respondents admit that they do not know the name for the item
Other	A response that could not be classified into any of the above categories.

List of categories for responses produced by Sample C and D

The demographic variables SPSS file consisted of the following categorical predictor variables: age group, gender, education, occupation, pattern of language acquisition, pattern of language use and reading habits. Quantitative predictor variables included age, years of education and the Clock Drawing Test score. The following dependent variables were derived from the data collected: strict score, lenient score, total number of English responses produced and mean reaction time for all the 60 respondents. Both complete and reduced versions of the demographic SPSS file are available (please refer to sections 3.3.6 and 3.3.8 in Chapter 3 and section 4.2.1 in Chapter 4 for a description of the data available).

The word variables SPSS file also consisted of a number of quantitative predictor variables (imageability rating, age of acquisition rating, familiarity rating, phoneme and syllable length), and categorical predictor variables (word initial fricative/affricate, animacy and cognate status). The dependent variables derived from the data consisted of the strict number of correct responses for each item, the lenient number of correct responses for each item, the number of English words produced for each item and the mean reaction time for each item. The reaction times of the responses which were strictly correct only were retained for analysis purposes. The strict and lenient number of correct responses of Sample A were also included in this file for the possibility of comparing results. Only the reduced set of the word variables SPSS file was used during analysis on word variables (see section 5.1 in Chapter 5).

3.5.6 Reaction Time

The Reaction Time, or naming latency, of each response collected from Samples C and D was recorded (in milliseconds) by means of a DmDx Voice Reaction Timer that was installed into the computer. "DMDX is a Win32 program designed to precisely time the presentation of text, audio, graphical, and video material and to enable the measurement of reaction times (RTs) to these displays with millisecond accuracy" (Forster & Forster 2003, p.116). Each response and the Reaction Time (RT) associated with each response were recorded automatically by the Voice Reaction Timer and a list of RTs corresponding to each item of the BNT was automatically saved for each respondent. The RT data was, then, transferred into two different SPSS files. The demographic variables file included the mean RT for each respondent (i.e. the average of the RTs of all BNT responses for the particular respondent), while the word variables file included the mean RT for each BNT item (i.e. the average of the RTs of all respondent responses for each item).

Target or alternative responses that were uttered after the allowed time frame had elapsed were scored as correct target and alternative responses respectively. However, the Reaction Time in this case was taken as 8000 msecs, as this was the maximum time recorded by the Voice Reaction Timer.

Self corrections were noted down as SC and included in a separate column (in the SPSS file). The RTs of self corrections were calculated manually following the completion of the test. In order to calculate a RT manually it was necessary to look at the audacity file of the particular respondent, to find the particular item in the test, to mute all other item responses and to listen to the item selected. The start of the self correction was taken as the corrected RT.

The procedure of manual RT calculation was also necessary whenever the respondents coughed, had false starts or hesitations. From the researcher's notes on the record form, it was easy to identify which items required manual calculation of RT.

Phonemic cues were provided at the end of the test for items which appeared to have been visually misperceived. Target or alternative responses that followed a phonemic cue were not scored as correct in strict or lenient scoring. They were simply noted down as PhCue on the record sheet. The RTs of responses that were produced following a phonemic cue were eliminated from the data as the RT was considered to be invalid. Circumlocutions and *Don't Know* responses were also considered to have invalid RTs. Only valid RTs were retained for analysis. Thus, the mean RTs for each respondent and for each item were calculated from the valid RTs only. The valid RTs included the RTs of those responses which were strictly correct and the manually corrected RTs of SCs and of responses with false starts, hesitations and coughs. Only the reaction times of the responses which were strictly correct were retained for analysis purposes.

3.6 ANALYSIS AND RESULTS

Several statistical calculations were performed by means of SPSS (Statistical Package for the Social Sciences, version 16, 2007) including analysis of demographic data, simple linear correlations and regression analysis. The following two chapters present the results obtained from the analysis on demographic variables and word variables respectively.

The first Results chapter (Chapter 4) mainly focuses on presenting the analysis and results of the effect of demographic variables on naming accuracy and naming latency. Demographic variables include age, education and gender. The first sections of the chapter present the information obtained from the data collected from Sample A as well as the comparison of results obtained from using different sets of the BNT, different samples and different scoring methods. Hypotheses relating to demographic variables are then stated one by one and analyses using simple linear correlations and one-way

ANOVAs follow each hypothesis. Results are obtained from both the reduced and complete sets of data. Regression analysis is then carried out in order to identify the most significant predictor variables. Only the reduced set of data is used during regression analysis (see section 4.7 in Chapter 4).

The second Results chapter (Chapter 5) is concerned with word variables. The influence of word variables on naming performance is investigated. The word variables included word length, age of acquisition, imageability, cognate status and familiarity. Once again, hypotheses are stated one by one and simple linear correlations and one-way ANOVAs are worked out. Regression analysis is then used to determine the most significant predictor word variables for naming performance. Only the reduced set is utilised during the analysis on word variables. The final part of Chapter 5 looks at the most significant demographic and word predictor variables and the relationship between them.

The results obtained from the data collected from people without aphasia are discussed in Chapter 6. The second part of the research is introduced in Chapter 7 (Methodology 2). The results and analysis of the data obtained from people with aphasia are, then, reported in the subsequent chapters.

Chapter 4

Results

Demographic variables and people without aphasia

4.0 INTRODUCTION

This chapter is concerned with the analysis of demographic variables and their influence on naming accuracy and naming latency. The results related to demographic variables were mainly derived from the data collected from Sample C and Sample D (the main study). Before looking into these results, however, it is necessary to present the information derived from the first part of the study (Sample A / name agreement study). The data collected from Sample A produced useful preliminary information that was later utilised with Samples C and D, as it guided the methodology and scoring procedures used in the main study. The following sections describe the information derived from Sample A (see section 3.3 in Chapter 3 for a description of data collection from Sample A).

4.1 SAMPLE A

Sample A provided data regarding the suitability of the BNT for Maltese-speaking individuals. The responses expected for each item of the BNT were obtained from this sample and name agreement for each item was worked out. Since a number of the items had low name agreement, a reduced set of the BNT was also produced. This reduced set consisted of those items which had at least 70% name agreement. Name agreement for each item was calculated and compared to the main study data and to the results of other studies. Finally, the scores achieved by the participants of Sample A were worked out and compared to the scores obtained from the main study.

4.1.1 Dominant and alternative responses

The responses produced by the participants of Sample A were matched to the list of possible responses that had been initially drawn by the translator (see section 3.3.3 in Chapter 3). All the possible responses for each item were listed by the translator. The response which was produced by the majority of the participants of Sample A for each

BNT item was termed the dominant response. The other names that were used by the participants for each item were regarded as alternative responses if they were included in the translator's list of possible responses. All other responses were considered as incorrect. Table 4.1 provides the list of dominant and alternative responses for the full set of 60 items of the BNT. All English translation equivalents (in italics) were also regarded as alternative responses, when the dominant word was in Maltese, and vice versa. Therefore, alternative responses for each item were made up of synonyms and English/Maltese translation equivalents. The dominant responses were used as the target responses during the main study.

4.1.2 Name agreement

Name agreement refers to the percentage of people who produce the dominant name for each picture (Székely et al 2003). Table 4.1 also shows the percentage of Sample A participants who produced a dominant response as well as the percentage of those who used a possible response for each item. The possible responses include both the dominant response and the alternative responses. Therefore, the percentages listed in the final column of Table 4.1 show the percentage of participants who produced any acceptable response, and not only the percentage of participants who produced an alternative response. The items which had less than 70% name agreement (in the dominant responses column) were regarded as having low name agreement. There were 22 items with low name agreement and 38 items with at least 70% name agreement.

TABLE 4.1

No.	English Words	Dominant Response	Alternative Responses	Participants who produced a dominant response (%)[1]	Participants who produced a possible response[2] (%)
1	*bed*	sodda	*bed*	100	100
2	*tree*	siġra	*tree*	100	100
3	*pencil*	lapes	*pencil*	100	100
4	*house*	dar	*house*	100	100
5	*whistle*	suffara	*whistle*	100	100
6	*scissors*	mqass	*scissors*	100	100
7	*comb*	pettne	moxt, *comb*	87.5	100
8	*flower*	fjura	*flower*	100	100
9	*saw*	serrieq	*saw*	91.67	95.84
10	*toothbrush*	*toothbrush*	xkupilja tas-snien, xkupilja	83.33	100
11	*helicopter*	ħelikopter		100	100

[1] Strict number of correct responses for each item in percentage (from Sample A).
[2] Lenient number of correct responses for each item in percentage (from Sample A).

No.	English Words	Dominant Response	Alternative Responses	Participants who produced a dominant response (%)	Participants who produced a possible response (%)
12	*broom*	xkupa	mikinsa, *broom*	83.33	87.50
13	*octopus*	qarnita	*octopus*	100	100
14	*mushroom*	*mushroom*	faqqiegħ	79.17	100
15	*hanger*	spalliera	*hanger*	100	100
16	*wheelchair*	*wheelchair*	siġġu tar-roti	66.67	100
17	*camel*	ġemel	*camel*	100	100
18	*mask*	maskra (maskla)	*mask*	95.83	95.83
19	*pretzel*	*pretzel*		16.67	16.67
20	*bench*	bank	*bench*	100	100
21	*racquet*	racquet	*tennis racquet,* paletta tat-tennis, paletta, *racquet* tat-tennis	41.67	91.67
22	*snail*	bebbuxu	*snail*	100	100
23	*volcano*	vulkan	*volcano*	100	100
24	*seahorse*	*seahorse*	żiemel tal-bahar	58.33	95.83
25	*dart*	dart		79.17	79.17
26	*canoe*	*canoe*	kenura	54.17	54.17
27	*globe*	globu	id-dinja, *globe*	100	100
28	*wreath*	girlanda	*wreath,* kuruna	37.50	75
29	*beaver*	*beaver*	kastur	33.33	33.33
30	*harmonica*	*mouth organ*	orgni tal-ħalq, *harmonica*, organett, fisarmonika	45.83	75
31	*rhinoceros*	*rhinoceros*	rinoċeronte	70.83	70.83
32	*acorn*	*acorn*	ġandra, ċippitatu	45.83	54.16
33	*igloo*	iglù		91.67	91.67
34	*stilts*	*stilts*	strippi, trample	50	58.33
35	*dominoes*	*Dominoes*	domino	87.5	87.50
36	*cactus*	Kaktus		91.67	91.67
37	*escalator*	*escalator*		87.50	87.50
38	*harp*	arpa	*harp*	95.83	95.83
39	*hammock*	*hammock*	benniena, branda	33.33	50
40	*knocker*	ħabbata	*knocker*	75	75
41	*pelican*	pelican	*pelican*	54.17	54.17
42	*stethoscope*	*stethoscope*	stetoskopju	75	75
43	*pyramid*	piramida	*pyramid*	100	100
44	*muzzle*	sarima	*muzzle*	37.50	50
45	*unicorn*	*unicorn*	unikornu	91.67	91.67
46	*funnel*	lembut	*funnel*	87.50	100
47	*accordian*	*accordian*	akkordjin, armonju	50	58.33
48	*noose*	ingassa	*noose*	25	45.83
49	*asparagus*	*asparagus*	*asparagu*	25	25
50	*compass*	kumpass	*compass*	100	100
51	*latch*	*lock*	lukkett, sokra, serratura, firroll	37.50	79.17
52	*tripod*	*tripod*	tripied	75	75
53	*scroll*	skroll		50	50
54	*tongs*	*tongs*		33.33	33.33
55	*sphinx*	*sphinx*	sfingi	50	50
56	*yoke*	*yoke*	madmad	13	20.83
57	*trellis*	kannizzata	xeblieka, *trellis*	45.83	70.83
58	*palette*	*palette*	paletta, tavlozza	79.17	79.17
59	*protractor*	*protractor*		100	100
60	*abacus*	*abacus*	abaku	100	100

Dominant and alternative responses with name agreement percentages

4.1.3 A complete and a reduced set of the BNT

A reduced set of the BNT was created using the name agreement data. This set included only the 38 items which had at least 70% name agreement. A list of the reduced set of BNT items is shown in Table 4.2.

TABLE 4.2

Orig. BNT No.	No.	English Words	Orig. BNT No.	No.	English Words	Orig. BNT No.	No.	English Words
1	1	bed	14	14	mushroom	37	27	escalator
2	2	tree	15	15	hanger	38	28	harp
3	3	pencil	17	16	camel	40	29	knocker
4	4	house	18	17	mask	42	30	stethoscope
5	5	whistle	20	18	bench	43	31	pyramid
6	6	scissors	22	19	snail	45	32	unicorn
7	7	comb	23	20	volcano	46	33	funnel
8	8	flower	25	21	dart	50	34	compass
9	9	saw	27	22	globe	52	35	tripod
10	10	toothbrush	31	23	rhinoceros	58	36	palette
11	11	helicopter	33	24	igloo	59	37	protractor
12	12	broom	35	25	dominoes	60	38	abacus
13	13	octopus	36	26	cactus			

The reduced set of the BNT

4.1.4 Strict and lenient number of correct responses

Responses which matched the dominant response were regarded as strict correct responses. Pairs of cognate words (such as ‘piramida’ and *pyramid*) were both regarded as strict correct responses. On the other hand, responses which matched one of the possible responses (dominant and/or alternative responses) were regarded as lenient correct responses. Since the alternative responses were not incorrect responses *per se*, it was necessary to create a system that gives credit to those responses which were acceptable but not dominant (see Table 4.1 for the list of acceptable words, i.e. the dominant and the alternative responses). Thus, two sets of results were worked out for *each item* of the BNT: a strict number of correct responses, as well as a lenient number of correct responses.

A strict score and a lenient score were also worked out for *each participant*. A strict score was awarded for each response that matched the target/dominant word and a lenient score was awarded for each response that matched one of the possible responses. This same scoring procedure was then used during the main study – both in terms of the number of correct responses for each BNT item and also in terms of the score for each respondent.

4.1.5 Comparison of name agreement across studies

Several researchers report the percentage of correct responses obtained from their studies on different populations. Table 4.3 provides an overview of the data reported by some studies on the BNT, together with the results obtained from the present project (name agreement study and main study). The results are in percentages for comparison purposes. There are a number of differences between the samples who participated in the different studies presented. First and foremost, the populations differed in terms of language and culture, which undoubtedly affects word variables. Other sample differences include the age groups and education levels of the participants, the sample size, as well as different test administration and scoring procedures. Nevertheless, a comparison of the results of Sample A to other studies was carried out to determine whether there was a significant difference between the number of correct responses for each item across different studies. Both the complete set and the reduced set were used to compare results in the studies. When using the complete set, significant differences between the results of the Maltese study and the Swedish Study, the Canadian Study and the New Zealand Study became evident (see Table 4.4).

Higher means were achieved by the Swedish, Canadian and the New Zealand samples than those achieved by the Maltese Sample A study. On the other hand, there was no significant difference between the results of the Greek Study and the Maltese Study when Strict Scoring was used, $t(55) = .870$, $p = .388$. A significant difference from the Greek study was found when using Lenient Scoring, however, with higher means for the Maltese sample.

A very contrasting picture emerged when the reduced set of the BNT was used to compare the different studies. Only the Greek study was significantly different from the Maltese study (using Sample A), with the Maltese sample producing better results. All the other studies showed that there was no significant difference in the performance on items of the reduced set (see Table 4.5) when compared to Sample A (using strict and

TABLE 4.3
Name Agreement results of different studies (in percentages)

	English item	SampleA[3] Strict correct[4]	Sample A Lenient correct[5]	Greek study[6]	Swedish study[7]	New Zealand students[8]	Canadian study[9]	Main Study[10] Strict correct	Main Study Lenient correct
1	bed	100	100	99	100	100	100	100	100
2	tree	100	100	98	100	100	100	100	100
3	pencil	100	100	91	100	100	100	100	100
4	house	100	100	94	100	100	100	98.3	100
5	whistle	100	100	94	98	98.3	99.5	80	81.7
6	scissors	100	100	100	100	100	100	100	100
7	comb	87.5	100	100	100	100	100	83.3	98.3
8	flower	100	100	93	99	100	100	100	100
9	saw	91.67	95.84	100	100	100	100	96.7	98.3
10	toothbrush	83.33	100	88	99	100	100	85	91.7
11	helicopter	100	100	93	100	100	99.1	93.3	93.3
12	broom	83.33	87.50	96	98	100	100	78.3	83.3
13	octopus	100	100	99	93	89.7	90	75	95
14	mushroom	79.17	100	91	100	98.3	99.5	91.7	95
15	hanger	100	100	100	100	100	100	95	98.3
16	wheelchair	66.67	100	96	99	100	100	91.7	100
17	camel	100	100	93	100	98.3	99.1	96.7	96.7
18	mask	95.83	95.83	73	99	89.7	98.6	80	80
19	pretzel	16.67	16.67	replaced item	99	27.6	92.2	6.7	6.7
20	bench	100	100	85	94	98.3	99.5	100	100
21	racquet	41.67	91.67	85	100	100	100	21.7	65
22	snail	100	100	97	100	98.3	95.4	95	98.3
23	volcano	100	100	76	89	98.3	97.7	85	85
24	seahorse	58.33	95.83	54	84	84.5	84.9	41.7	66.7
25	dart	79.17	79.17	62	99	96.6	98.6	68.3	68.3

[3] Name Agreement sample (Sample A). 24 participants; average age : 37.2 years; age range : 20-60 years; average years of education : 17.75; educ range : 11-20 years.
[4] Percentage of participants who provided the dominant response. Self corrections and cognate words were marked as strict correct too in samples A, C and D for this part of the study.
[5] Percentage of participants who provided a possible response (i.e. either a dominant response, a synonym or else an English equivalent. It is the total of the sum of possible responses).
[6] Patricacou et al (2007). Study on native Greek speakers. 100 participants; average age : unknown; age range : 20 – 71+ years; average years of education : unknown; educ range: 0 – 13+ years. Synonyms counted as correct too in this study.
[7] Tallberg (2005). Study on native Swedish speakers. 111 participants; average age : 49.5 years; age range : 21 – 80; average years of education : 12.8; educ range : 6 – 22 years.
[8] Barker-Collo (2001). 58 participants; average age : 20 .2 years; age range : 17.5 - 25.25 years; 1st & 2nd year University students.
[9] Tombaugh & Hubley (1997). 219 participants; average age : 59 years; age range : 25-88 years; average years of education : 12.90; educ range : 9 - 21 years.
[10] Main Study (Samples C & D together). 60 participants; average age : 52.7 years; age range : 20- 85 years; average years of education : 12.65; educ range: 0- 19 years.

	English item	SampleA Strict correct	Sample A Lenient correct	Greek study	Swedish study	New Zealand students	Canadian study	Main Study Strict correct	Main Study Lenient correct
26	canoe	54.17	54.17	35	95	89.7	100	56.7	56.7
27	globe	100	100	70	96	77.6	96.8	88.3	95
28	wreath	37.50	75	100	100	93.1	99.5	18.3	40
29	beaver	33.33	33.33	28	90	31	97.5	20	20
30	harmonica	45.83	75	73	100	98.3	96.8	26.7	58.3
31	rhinoceros	70.83	70.83	79	99	94.8	90.4	61.7	61.7
32	acorn	45.83	54.16	74	88	79.3	93.6	18.3	25
33	igloo	91.67	91.67	39	95	91.4	99.1	53.3	53.3
34	stilts	50	58.33	55	95	91.4	95	20	23.3
35	dominoes	87.50	87.50	35	86	82.8	90.9	55	55
36	cactus	91.67	91.67	74	98	93.1	100	73.3	73.3
37	escalator	87.50	87.50	69	100	96.6	99.1	60	60
38	harp	95.83	95.83	62	88	98.3	97.3	91.7	91.7
39	hammock	33.33	50	54	95	93.1	94.1	28.3	65
40	knocker	75	75	replaced item	35	81	97.7	45	48.3
41	pelican	54.17	54.17	72	84	98.3	92.7	38.3	38.3
42	stethoscope	75	75	replaced item	91	93.1	95	41.7	41.7
43	pyramid	100	100	80	98	94.8	96.8	73.3	73.3
44	muzzle	37.50	50	74	60	91.4	92.7	20	30
45	unicorn	91.67	91.67	18	60	82.8	91.3	41.7	41.7
46	funnel	87.50	100	94	99	75.9	96.3	90	96.7
47	accordion	50	58.33	87	100	93.1	81.7	30	63.3
48	noose	25	45.83	72	72	89.7	91.3	35	45
49	asparagus	25	25	50	77	70.7	93.6	25	25
50	compass	100	100	74	77	60.3	69	85	85
51	latch	37.5	79.17	46	46	84.5	80.8	16.7	46.7
52	tripod	75	75	56	67	55.2	89.5	50	50
53	scroll	50	50	replaced item	50	93.1	92.7	30	30
54	tongs	33.33	33.33	60	85	89.7	84.5	30	30
55	sphinx	50	50	46	79	77.6	75.8	31.7	38.3
56	yoke	13	20.83	27	67	69	63	0	3.3
57	trellis	45.83	70.83	26	81	87.9	77.2	6.7	13.3
58	palette	79.17	79.17	37	82	77.6	69	43.3	43.3
59	protractor	100	100	58	71	27.6	39.7	48.3	48.3
60	abacus	100	100	38	76	65.5	57.5	45	45

lenient scoring). Thus, when using only the 38 items of the reduced set the different studies compared here (except for the Greek study) showed similar name agreement.

TABLE 4.4

Paired Differences									
					95% Confidence Interval of the Difference				
Pairs of samples compared		**Mean**	**SD**	**Std. Error Mean**	**Lower**	**Upper**	**t**	**df**	**P-value**
Sample A strict scoring	Greek Study	3.167	27.234	3.639	-4.127	10.460	.870	55	.388
Sample A lenient scoring	Greek Study	10.449	22.912	3.062	4.314	16.585	3.413	55	.001
Sample A strict scoring	Swedish Study	-15.317	26.524	3.424	-22.169	-8.465	-4.473	59	.000
Sample A lenient scoring	Swedish Study	-8.519	23.630	3.051	-14.624	-2.415	-2.793	59	.007
Sample A strict scoring	Canadian Study	-18.650	27.896	3.601	-25.856	-11.444	-5.179	59	.000
Sample A lenient scoring	Canadian Study	-11.853	24.760	3.196	-18.249	-5.457	-3.708	59	.000
Sample A strict scoring	New Zealand Study	-13.903	27.599	3.563	-21.033	-6.774	-3.902	59	.000
Sample A lenient scoring	New Zealand Study	-7.106	23.147	2.988	-13.086	-1.127	-2.378	59	.021

Comparison of Sample A name agreement results to different studies using the complete set

It should be noted that the same comparisons described above were carried out between the main study results and the other studies. In this case, the BNT main study results were significantly different from the results obtained by the Swedish, New Zealand and Canadian studies. This was evident in both the complete and the reduced sets of the BNT. On the other hand, there was no significant difference between the Greek study and the main study results.

4.1.6 The scores achieved by Sample A participants

The mean scores achieved by the participants of sample A were compared to those achieved by the respondents of the main study. The results are reported in section 4.3.2. Different sets of scores are presented and compared, namely the strict and lenient scores for both the reduced and the complete set of the BNT.

TABLE 4.5

Pairs of samples compared		Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference: Lower	95% Confidence Interval of the Difference: Upper	t	df	P-value
		Paired Differences							
Sample A strict scoring	Greek Study	15.371	21.754	3.626	8.010	22.731	4.239	35	.000
Sample A lenient scoring	Greek Study	17.338	20.005	3.334	10.570	24.107	5.200	35	.000
Sample A strict scoring	Swedish Study	.588	14.533	2.358	-4.189	5.365	.249	37	.804
Sample A lenient scoring	Swedish Study	2.452	13.231	2.146	-1.897	6.801	1.142	37	.261
Sample A strict scoring	Canadian Study	-1.159	16.476	2.673	-6.575	4.256	-.434	37	.667
Sample A lenient scoring	Canadian Study	.705	15.589	2.529	-4.419	5.829	.279	37	.782
Sample A strict scoring	New Zealand Study	2.477	17.685	2.869	-3.336	8.290	.864	37	.393
Sample A lenient scoring	New Zealand Study	4.342	16.897	2.741	-1.212	9.896	1.584	37	.122

Comparison of Sample A name agreement results to different studies using the reduced set

4.2 THE MAIN STUDY

A number of hypotheses have been proposed to identify the possible effects of particular demographic variables on naming performance. These hypotheses are based on previous research findings on other populations, as well as on the characteristics of the Maltese population. Chapter 2 combines an overview of the most salient research findings on the topic of spoken word production (in the context of confrontation naming), with related proposals about possible findings in the present study and including the rationale for the particular proposals. This chapter, on the other hand, aims to state the hypotheses that are related to demographic variables, to present the findings that emerged from this study and to analyse the relevant results. Prior to commencing with result presentation, it is essential to provide a brief description of the data available.

4.2.1 The available data

The demographic predictor (or independent) variables that are available in the present study include both qualitative (categorical) and quantitative variables. The categorical variables include age group, gender, education category, pattern of language acquisition, pattern of daily language use and reading habits, while the quantitative variables include age and years of education. Chapter 3 provides a detailed description of the sample available, with information regarding each demographic independent variable (see section 3.5.4 in Chapter 3).

Three different dependent variables emerge from the present data. These include the *score* (i.e. the total number of correct responses for each respondent), the *mean reaction time* for each respondent (RT), and the total *number of English responses* produced by each respondent. Each of these dependent variables will be correlated with particular predictor variables.

As described in Chapter 3, data files containing strict and lenient scores for both complete and reduced BNT sets are available. Thus, four different sets of data are available. These include:

- a complete set of 60 BNT items using strict scores only;
- a complete set of 60 BNT items using lenient scores;
- a reduced set of 38 BNT items using strict scores only;
- a reduced set of 38 BNT items using lenient scores.

Different analyses using these four different sets of data were initially performed in order to demonstrate possible differences in results between sets of data, and to identify the best set that may be used in further analyses. The following sections present the results related to the three dependent variables (Score, RT, English responses). Results from each set of data (see the 4 sets listed above) will be presented for comparison purposes.

4.3 GLOBAL MEAN SCORE

Global mean scores, and standard deviations, derived from the four different sets of data are presented in Table 4.6. Percentages of the mean scores are included for comparison purposes, as the scores of the complete set are out of 60, while those of the reduced set are out of a maximum score of 38. The upper section of the table presents the results derived from the main study sample[11] while the lower section of the table shows the global mean scores derived from Sample A[12], which was the sample used to provide preliminary information about name agreement and target responses.

4.3.1 Comparing complete and reduced sets

The complete and the reduced set of the BNT were compared in terms of the mean scores achieved by the respondents. A strong positive correlation emerged between the strict score results of the complete set and the strict score results of the reduced set ($r = .952$, $p < .001$), showing that the respondents who scored high on the complete set also scored high on the reduced set and vice versa. This implies that whichever set is used similar results would be achieved.

TABLE 4.6

MAIN STUDY Complete Set (score out of 60 items)				**MAIN STUDY Reduced Set** (score out of 38 items)			
Mean Strict Score	**SD**	**Mean Lenient Score**	**SD**	**Mean Strict Score**	**SD**	**Mean Lenient Score**	**SD**
34.6	9.52	**39.2**	9.72	**28.9**	6.17	**30.3**	6.20
57.6%		**63.3%**		**76.1%**		**78.1%**	
SAMPLE A Complete Set (score out of 60 items)				**SAMPLE A Reduced Set** (score out of 38 items)			
Mean Strict Score	**SD**	**Mean Lenient Score**	**SD**	**Mean Strict Score**	**SD**	**Mean Lenient Score**	**SD**
44.8	6.05	**48.6**	6.10	**35.1**	2.60	**35.9**	2.42
74.7%		**80.9%**		**92.4%**		**94.4%**	

Global Mean Scores (with scores and percentages)

[11] The main study sample consisted of 60 respondents, mean age of 52.7 years and mean education of 12.6 years.
[12] Sample A consisted of 24 participants, with a mean age of 37.2 years and a mean education of 17.1 years.

The strict score (percentage score) obtained in the main study using the complete set (57.6%) is lower than that obtained from using the reduced Set (76.1%). The difference between the strict percentage score of the two sets is 18.5%. A paired samples t-test indicated that this is a significant difference t (59) = -28.45, $p < .02$. Likewise, there is a significant difference between the strict score of the complete and the reduced sets in Sample A (t (23) = -17.83, $p < .02$). This discrepancy may be explained by the fact that the reduced set includes only the items with names that were known by the majority of the Maltese participants, as all the items with less than 70% name agreement were eliminated in the reduced set. A similar discrepancy in results can also be seen when lenient scores are compared between complete and reduced sets in the main study (t (59) = -20.98, $p < .02$) and for sample A (t (23) = -12.51, $p < .02$). Thus, there is a significant difference between the percentage score of the complete and the reduced sets.

4.3.2 Comparing Sample A to the main study sample

A further difference may be seen between the results obtained from the name agreement sample and those obtained from the main study. In this case, the mean scores obtained from the name agreement sample were markedly higher than those obtained in the main study for all sets of data (see Table 4.6). Possible reasons for this difference may lie in the fact that the name agreement sample involved different demographic characteristics from those in the main study. Fewer participants took part in the name agreement study and they had a mean age of 37.2 years and a mean education of 17.1 years (SD = 2.35), while the respondents of the main study had a mean age of 52.7 years and a mean education of 12.6 years (SD = 4.39). Another possible reason for obtaining higher scores in the name agreement sample was the context of data collection. While the name agreement study involved showing a hard copy of pictures without any time pressure, the respondents in the main study were shown each picture on a screen for a brief period of time, were allowed a restricted amount of time to respond, and they were also told that their reaction time was being recorded. This time pressure may be a major cause of the lower amount of correct responses obtained in the main study.

In order to reach conclusions regarding the impact of timed picture naming on the performance of the respondents of the main study, the performance of a set of participants from Sample A was compared to that of a matched set of respondents from the main study. The matched set consisted of participants and respondents aged between 30 and 45 years and having 13 or more years of education. The number of correct responses for each item of the BNT was compared between the two groups. A t-test showed that there was no significant difference between the performance of the Sample A participants and the Main Study respondents when a matched set of individuals was compared. Both strict scoring ($t(37) = -.812$, $p = .422$) and lenient scoring ($t(37) = -.750$, $p = .458$) showed that there was not a significant difference between the two groups. Thus, this provides evidence that the method of data collection (including the time pressure to respond as quickly and as accurately as possible) used during the main study did not negatively influence the performance of the respondents. Rather, the significant difference found between the performance of Sample A and that of the main study is more likely to be attributed to the effects of age and education. It should be noted that a matched set of older participants was not available. Therefore, no true conclusions can be reached regarding the performance of older respondents under time pressure. The impact on the older age groups may have been stronger as a result of slower processing of visual information, which becomes increasingly difficult for ageing individuals in contexts of time pressure (Madden et al 2005). (See section 2.5.3.2 in Chapter 2).

4.3.3 Comparing scoring methods

The type of scoring procedure used also plays a very important role in the results achieved. A strong positive correlation emerged between the strict score results and the lenient score results of the complete ($r = .982$, $p < .001$) and the reduced ($r = .984$, $p < .001$) sets, showing that the respondents who scored high on strict scoring methods also scored high on the lenient scoring methods. While strict scoring produced a low mean score, lenient scoring methods produced higher mean scores, as expected. A paired-sample t-test showed a significant difference of 5.7% between the scores achieved when using strict and lenient scoring ($t\ (59) = -13.63$, $p < .02$) in the complete set, and a smaller significant difference of 2% also when using the reduced set ($t\ (59) = -5.35$, $p < .02$).

Thus, it is evident that the different sets of data are correlated with each other, that is, participants who did well on the complete set were expected to do well on the reduced set too, and those who did well when strict scoring was used also did well on lenient scoring too. However, it is evident that the highest scores in the main study were achieved when lenient scoring was used in the reduced set.

4.3.4 Comparing different studies

The mean scores obtained from the Maltese individuals taking part in this study are lower than those generally reported on different populations. A mean strict score of 34.6, and a mean lenient score of 39.2 were obtained from the main study using the complete set (out of 60). Most studies have found a mean score above 50 correct out of 60 items (e.g. mean scores of 54.0 (Van Gorp et al 1986), 54.5 (Nicholas et al. 1989) and 52.8 (Worrall et al 1995)). This trend was also shown when the number of correct responses for each item was compared across studies (see section 4.1.5). However, it has been argued that the higher scores obtained in other studies may have been related to the higher education bias in their samples. For example, Patricacou et al (2007) found a lower mean score of 42.86 in their study which also included participants with low levels of education. A number of reasons may explain the lower mean scores achieved by the Maltese participants. The varied education backgrounds of the Maltese participants may be one explanation as the education of the respondents ranged from 0 to 19 years.

The linguistic and cultural differences between studies may also have played a significant role in achieving lower scores here. In fact, another possible reason for the lower mean scores obtained in the present study may be the bilingual circumstances of the Maltese population. Other studies on bilingual populations have also found lower mean scores (Roberts et al 2002; Kohnert et al 1998). The possible influence of bilingualism in explaining the lower mean scores obtained from the Maltese sample is discussed in Chapter 6 (Section 6.2.4.2).

4.4 GLOBAL MEAN REACTION TIME

The global mean reaction time (RT) and standard deviations for both complete and reduced sets are presented in Table 4.7. Only the responses which were strictly correct were used to calculate the mean reaction times. Therefore, lenient scoring is not included in the RT analysis. Only the mean RTs of the responses which were strictly correct could be kept for analysis in order to ensure uniformity among RT results of different respondents, and to avoid possible effects of item variables on the mean RT of each respondent. If any lenient (alternative) response had been included in analysing the mean RT of different individuals, different lenient responses would have had different AoA, different phoneme length, etc.. It is also worth remembering that the reaction time was not recorded in the preliminary study using Sample A.

TABLE 4.7

MAIN STUDY Complete Set		**MAIN STUDY Reduced Set**	
Mean Reaction Time	**SD**	**Mean Reaction Time**	**SD**
1771.8	372.54	**1700.2**	381.88

Global Mean Reaction Time (in milliseconds)

As expected, there is a faster mean reaction time in the reduced set (which includes only items with high name agreement), showing that the reaction time is faster on the item names which were more accurately produced by Sample A. The difference between the mean RTs of the complete set and the reduced set is of 71.6 milliseconds. This seemingly small difference is a significant difference, as shown by a paired samples test t (59) = 7.09, $p < .02$.

4.5 GLOBAL MEAN NUMBER OF ENGLISH RESPONSES

It was predicted that an amount of English responses would be used by the respondents in the study because of the bilingual context of the Maltese islands, as well as because of the possibility that the material being used may be culturally biased. As the material was originally developed on an English-speaking population with different linguistic and cultural characteristics from the current study, it was predicted that the respondents would probably know the English words of some items more than their Maltese equivalents, as the words may have been primarily experienced or learned in an English-speaking context (e.g. through television or reading).

TABLE 4.8

MAIN STUDY Complete Set (out of 60 possible responses)		**MAIN STUDY Reduced Set (out of 38 possible responses)**	
Mean number of English responses	**SD**	**Mean number of English responses**	**SD**
11.63	7.68	6.45	4.05

Global Mean number of English responses

In fact, respondents produced a mean number of 11.63 English responses (SD = 7.68) in the complete set and a mean number of 6.45 English responses (SD = 4.05) in the reduced set (see Table 4.8).

4.6 HYPOTHESES TESTING

This section aims to state each hypothesis and to present the findings related to effects of demographic variables. Simple linear correlations and one-way ANOVAs were worked out first. Multiple regressions were performed to analyse possible interactions between variables and to identify the most important predictor/s for each dependent variable (see section 4.7). The predictors that are being used in the analysis described below include age/age group, years of education/education category and gender. The researcher is aware of the possible difficulties associated with small cell sizes during the analysis. Nonetheless, statistical investigations were still carried out in order to test predictions and to illustrate potential differences between variables.

4.6.1 Hypothesis 1

Hypothesis 1a

The respondents will produce fewer correct responses as age increases.

Hypothesis 1b

The respondents will have a longer mean reaction time as age increases.

Rationale

Previous studies have shown that ageing affects naming accuracy and naming latency as a result of neurological changes that occur to varying degrees in different individuals (see section 2.5 in Chapter 2). Thus, it is hypothesised that the respondents in the current study will also have a lower mean score (i.e. they will have fewer correct responses) with increasing age and will have a longer mean reaction time with increasing age.

4.6.1.1 Correlations between age and score

Simple linear correlations were performed in order to identify the effect of age on score. One-way ANOVAs with age group as the categorical variable were also performed. A Kolmogorov-Smirnov test confirmed that the dependent variable, score, had a normal distribution as the p-value exceeds the 0.05 level of significance in strict score and lenient score of both complete and reduced sets (see Table 4.9).

TABLE 4.9

MAIN STUDY	Strict Score Complete Set	Leninet Score Complete Set	Strict Score Reduced Set	Lenient Score Reduced Set
Kolmogorov-Smirnov Z	.649	.691	.945	1.061
P-value (2-tailed)	.793	.727	.333	.210

One-Sample Kolmogorov-Smirnov Test (score)

Simple linear correlations with age as the independent variable showed that there is a significant negative correlation between age and score. All the four sets of data described above (section 4.2.1) showed this negative correlation (see Table 4.10).

Thus, while complete sets produced a moderate negative correlation for both strict ($r = -.487$, $p < .01$) and lenient scores ($r = -.478$, $p < .01$), the reduced sets produced a strong negative correlation for both strict ($r = -.602$, $p < .01$) and lenient scores ($r = -.621$, $p < .01$). This demonstrates that when taken on its own ageing has a negative effect on naming ability with fewer correct responses with increasing age, as proposed in hypothesis 1a.

TABLE 4.10

		Strict Score Complete Set	Leninet Score Complete Set	Strict Score Reduced Set	Lenient Score Reduced Set
Age	Pearson Correlation	-.487**	-.478**	-.602**	-.621**
	P - value	.000	.000	.000	.000
	Sample size	60	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between age and score

A one-way ANOVA was also performed to make it possible to compare scores across the six age groups available in the present study. A significant decrease in the mean score became apparent with increasing age. There was a particularly striking mean score decrease in the two older age groups (i.e. 66 - 75 years and 76 – 85 years) when compared to the younger age groups. Tables 4.11 and 4.12 show the mean scores and the standard deviations for each age group, for complete and reduced sets respectively. The box plots which follow the tables demonstrate the striking dip in naming accuracy after 66 years for all sets described.

TABLE 4.11

		Strict Score Complete Set		Lenient Score Complete Set	
Age Group	**N**	**Mean**	**SD**	**Mean**	**SD**
20-29 years	10	39.40	5.835	42.70	6.584
30-45 years	10	38.10	6.967	41.20	8.011
46-55 years	10	37.10	5.547	41.00	6.164
56-65 years	10	37.90	9.723	41.80	9.864
66-75 years	9	28.11	10.018	31.56	9.262
76-85 years	11	27.09	10.737	30.27	10.574
Total	**60**	**34.60**	**9.523**	**39.20**	**9.723**

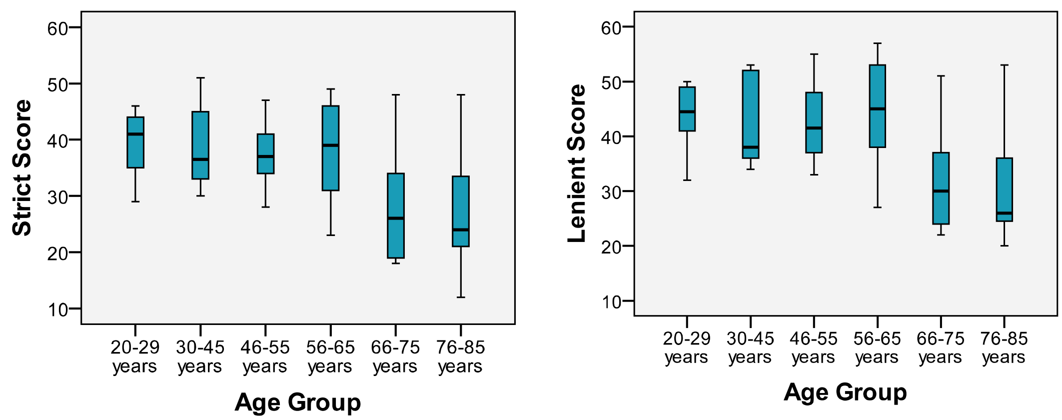
Mean scores for each age group using strict and lenient scoring in the complete set

TABLE 4.12

		Strict Score Reduced Set		Lenient Score Reduced Set	
Age Group	**N**	**Mean**	**SD**	**Mean**	**SD**
20-29 years	10	32.80	3.706	33.90	4.228
30-45 years	10	32.20	2.573	33.00	3.018
46-55 years	10	31.20	3.259	31.80	3.259
56-65 years	10	29.70	5.599	30.30	5.458
66-75 years	9	24.56	6.579	25.67	6.225
76-85 years	11	23.18	6.997	23.64	6.845
Total	**60**	**28.92**	**6.171**	**30.30**	**6.201**

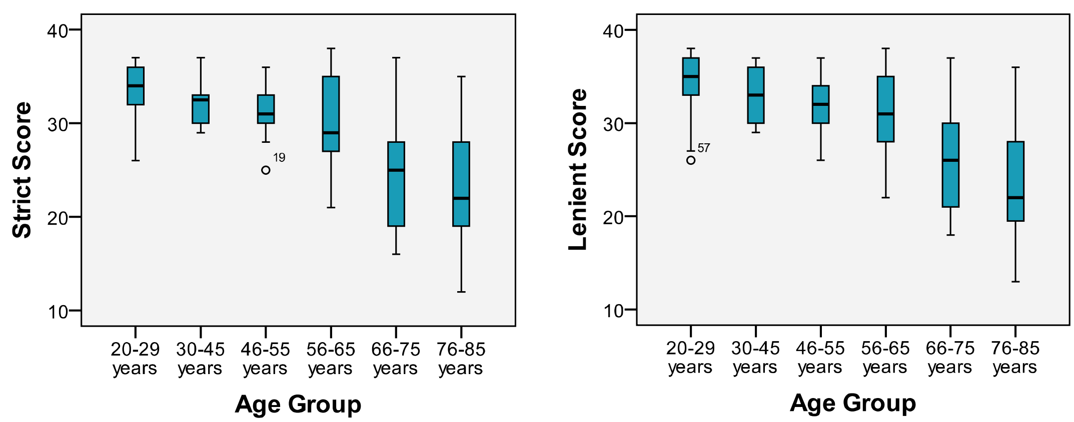
Mean scores for each age group using strict and lenient scoring in the reduced set

FIGURE 4.1



Mean scores with increasing age when using the complete set

FIGURE 4.2



Mean scores with increasing age when using the reduced set

Older age groups exhibit greater variation in performance. This replicates the findings of other researchers (e.g. Tsang and Lee 2003; Zec et al. 2007a). The greater variation is demonstrated by the higher standard deviation values of the older age groups and it is also evident in the box plots. Therefore, not only is the mean score decreasing with age, but there is also a greater variation in performance in older individuals. This may be a reflection of the greater variation in the cognitive abilities of different ageing individuals (e.g. Lindeboom and Weinstein 2004) and/or a reflection of the variation of education opportunities experienced by different ageing individuals (Zec et al. 2007a).

Post hoc comparisons using the Tukey test were carried out in order to decide whether there was an overall significant difference between the mean scores of each age group. The different sets of data produced different results when age groups were compared. Table 4.13 presents the significant differences in mean scores.

TABLE 4.13

Age Group	Age Groups	Mean Difference	Std. Error	P-value	95% Confidence Interval	
					Lower Bound	Upper Bound
Multiple Comparisons using Strict Scoring in the Complete Set						
76-85 years	20-29 years	-12.31*	3.678	.018	-23.18	-1.44
	30-45 years	-11.01*	3.678	.045	-21.88	-.14
	46-55 years	-10.01	3.678	.087	-20.88	.86
	56-65 years	-10.81	3.678	.052	-21.68	.06
	66-75 years	-1.02	3.784	1.000	-12.20	10.16
Multiple Comparisons using Lenient Scoring in the Complete Set						
76-85 years	20-29 years	-12.43*	3.756	.020	-23.52	-1.33
	30-45 years	-10.93	3.756	.056	-22.02	.17
	46-55 years	-10.73	3.756	.064	-21.82	.37
	56-65 years	-11.53*	3.756	.037	-22.62	-.43
	66-75 years	-1.28	3.863	.999	-12.70	10.13
Multiple Comparisons using Strict Scoring in the Reduced Set						
76-85 years	20-29 years	-9.62*	2.222	.001	-16.18	-3.05
	30-45 years	-9.02*	2.222	.002	-15.58	-2.45
	46-55 years	-8.02*	2.222	.008	-14.58	-1.45
	56-65 years	-6.52	2.222	.053	-13.08	.05
	66-75 years	-1.37	2.286	.991	-8.13	5.38
66-75 years	20-29 years	-8.24*	2.337	.011	-15.15	-1.34
	30-45 years	-7.64*	2.337	.022	-14.55	-.74
	46-55 years	-6.64	2.337	.066	-13.55	.26
	56-65 years	-5.14	2.337	.254	-12.05	1.76
	76-85 years	1.37	2.286	.991	-5.38	8.13
Multiple Comparisons using Lenient Scoring in the Reduced Set						
76-85 years	20-29 years	-10.26*	2.213	.000	-16.80	-3.73
	30-45 years	-9.36*	2.213	.001	-15.90	-2.83
	46-55 years	-8.16*	2.213	.007	-14.70	-1.63
	56-65 years	-6.66*	2.213	.043	-13.20	-.13
	66-75 years	-2.03	2.276	.947	-8.76	4.70
66-75 years	20-29 years	-8.23*	2.327	.010	-15.11	-1.36
	30-45 years	-7.33*	2.327	.030	-14.21	-.46
	46-55 years	-6.13	2.327	.106	-13.01	.74
	56-65 years	-4.63	2.327	.361	-11.51	2.24
	76-85 years	2.03	2.276	.947	-4.70	8.76

*. The mean difference is significant at the 0.05 level.

Mean score Post Hoc multiple comparisons using different sets of data

When using strict scoring in the complete set post hoc tests only found a significant difference between the mean scores of the youngest age groups 1 (20 – 29 years) and 2 (30 – 45 years) and the mean scores of the eldest age group 6 (76 – 85 years). When using lenient scoring in the complete set post hoc tests found a significant difference between the mean scores of the youngest age group 1 (20 – 29 years) and the mean scores of eldest age group 6 (76 – 85 years). A significant difference was also noted between age group 5 (66 – 75 years) and the eldest age group 6 (76 – 85 years).

When using the reduced set of data there was a significant difference in mean scores between the older age groups and various age groups. When using strict scoring in the reduced set, post hoc comparisons identified significant differences in mean scores between the eldest age group (76 – 85 years) and age groups 1 (20 – 29 years), 2 (30 – 45 years) and 3 (46 – 55 years). The mean score of age group 5 (66 – 75 years) was also found to be significantly different from the youngest two age groups 1 (20 – 29 years) and 2 (30 – 45 years). Similar results were found when lenient scoring in the reduced set was used. The only difference was that age group 4 (56 – 65 years) was also found to be significantly different from the eldest age group 6 (see Table 4.13).

4.6.1.2 Correlations between age and reaction time

Simple linear correlations were performed in order to identify the effect of age on RT. One-way ANOVAs with age group as the categorical variable were also performed. A Kolmogorov-Smirnov test confirmed that the dependent variable, reaction time, had a normal distribution as the p-value exceeds the 0.05 level of significance in both sets of data.

TABLE 4.14

	Reaction Time Complete Set	**Reaction Time Reduced Set**
Kolmogorov-Smirnov Z	1.159	1.280
P-value (2-tailed)	.136	.076

One-Sample Kolmogorov-Smirnov Test (reaction time)

Simple linear correlations with age as the independent variable showed that there is a significant positive correlation between age and reaction time. Two sets of data were analysed. These included only the Strict Scores in both complete and reduced sets (Table 4.15).

TABLE 4.15

		Mean RT Strict Score Complete set	Mean RT Strict Score Reduced set
Age	Pearson Correlation	.447**	.475**
	P - value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between age and RT

Thus, both complete and reduced sets produced a moderate positive correlation between age and RT when using strict scores (r = .447, p < .01, and r = .475, p < .01, respectively). This demonstrates that when taken on its own ageing has a positive effect on naming latency implying that reaction time becomes slower with increasing age, as proposed in hypothesis 1b.

One-way ANOVAs were performed to make it possible to compare reaction times across the six age groups available in the present study. Table 4.16 shows the mean RTs (in milliseconds) for each age group, and Figure 4.3 shows a graphical representation of the findings. An increase in the mean reaction time is apparent with increasing age and the two older age groups show a marked increase in mean RT. This result is similar to the marked decrease in mean score for these two older age groups (see sect. 4.6.1.1). However, despite the moderate positive correlations found between age and mean RT, it is worth noting that the shortest mean RT was achieved by the third age group (i.e. 46 – 55 years) who got a mean RT of 1561.70 msec (*SD* = 244.44) in the complete set and a mean RT of 1495.40 msec (*SD* = 205.48) in the reduced set.

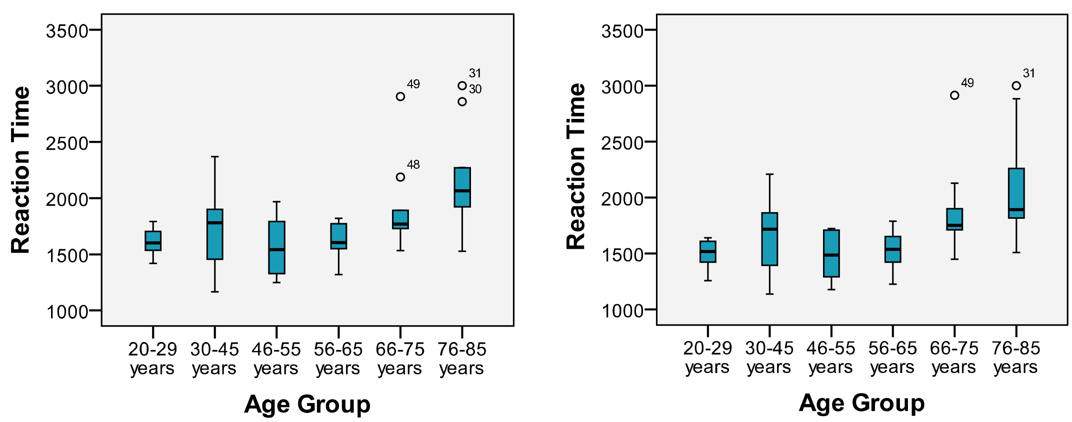
The youngest age group (20 – 29 years), which included only University Students, did not achieve the shortest mean RT as expected. The shortest mean RT had been expected from the 20 – 29 years age group as this included the youngest respondents, and also because all the respondents were enrolled in University courses at the time of the data collection. Their mean RT (1612.90; *SD* = 116.70 in the complete set, and 1504.90; *SD* = 123.78 in the reduced set) was longer than the mean of the 46-55 year age group. Further, the second youngest age group (i.e. 30-45 years) also achieved a mean RT (1720.90; *SD* = 369.31 in the complete set, and 1655.10; *SD* = 335.70 in the reduced set) which was longer than that of both the 46 – 55 year age group and of the 56 – 65 year age group for both complete and reduced sets. Thus, although an increase in RT is expected with increasing age, the younger age groups will not necessarily be the quickest in naming. On the other hand, it is possible that this finding may be the result of the small number of individuals included in each age group.

TABLE 4.16

		Mean RT Complete Set using Strict scoring only		Mean RT Reduced Set using Strict scoring only	
Age Group	**N**	**Mean**	**SD**	**Mean**	**SD**
20-29 years	10	1612.90	116.70	1504.90	123.78
30-45 years	10	1720.90	369.31	1655.10	335.70
46-55 years	10	1561.70	244.24	1495.40	205.48
56-65 years	10	1626.20	153.02	1534.30	167.95
66-75 years	9	1921.56	407.59	1896.11	423.58
76-85 years	11	2163.55	440.46	2095.27	471.58
Total	**60**	**1771.83**	**372.54**	**1700.17**	**381.88**

Mean Reaction Time (in milliseconds) for each age group in the complete and reduced sets

FIGURE 4.3



Complete Set **Reduced Set**

Mean Reaction Time with increasing age

The 6 age groups were once again compared using post hoc comparisons (Tukey test) in order to decide whether there was an overall significant difference between the mean reaction times of each age group. The complete and reduced sets of data[13] produced the same results when different age groups were compared in terms of mean RT (Table 4.17).

[13] Only strict scoring was used in RT analysis.

TABLE 4.17

Age Group	Age Groups	Mean Difference	Std. Error	P-value	95% Confidence Interval	
					Lower Bound	Upper Bound
Multiple Comparisons using Strict Scoring in the Complete Set						
76-85 years	20-29 years	550.65*	137.744	.003	143.68	957.61
	30-45 years	442.65*	137.744	.026	35.68	849.61
	46-55 years	601.85*	137.744	.001	194.88	1008.81
	56-65 years	537.35*	137.744	.003	130.38	944.31
	66-75 years	241.99	141.696	.533	-176.65	660.63
Multiple Comparisons using Strict Scoring in the Reduced Set						
76-85 years	20-29 years	590.37*	138.747	.001	180.45	1000.30
	30-45 years	440.17*	138.747	.029	30.25	850.10
	46-55 years	599.87*	138.747	.001	189.95	1009.80
	56-65 years	560.97*	138.747	.002	151.05	970.90
	66-75 years	199.16	142.728	.730	-222.52	620.85

*. The mean difference is significant at the 0.05 level.

Mean RT Post Hoc multiple comparisons using different sets of data

Both the complete and the reduced set post hoc tests found significant differences between the mean RTs of the eldest age group 6 (76 – 85 years) and age groups 1 (20 – 29 years), 2 (30 – 45 years), 3 (46 – 55 years) and 4 (56 – 65 years), but not significant differences between the eldest age group ((76 – 85 years) and the second eldest age group (66 – 75 years).

Hypothesis 1c

The respondents will produce fewer English responses as age increases.

Rationale

This hypothesis is based on knowledge about the use of the English and Maltese languages in Malta. It is proposed that older individuals are less likely to use English words when naming pictures, as older Maltese individuals tend to use less English daily when conversing with each other and when addressing children (Sciriha 1997). One possible reason for this may lie in the fact that there are better education opportunities (including a better acquisition of the English language) now than there were in the past, as formal education had not been obligatory by law before 1947 (Zammit Mangion 2000).

4.6.1.3 Correlations between age and English responses

A Kolmogorov-Smirnov test confirmed that the dependent variable, number of English responses, had a normal distribution as the p-value exceeds the 0.05 level of significance in both sets of data.

TABLE 4.18

	Number of English responses Complete Set	Number of English responses Reduced Set
Kolmogorov-Smirnov Z	1.039	.927
P-value (2-tailed)	.230	.357

One-Sample Kolmogorov-Smirnov Test (English responses)

Simple linear correlations between age and number of English responses showed a significant moderate negative correlation (see Table 4.19) in both the complete set and in the reduced set ($r = -.431$, $p < .01$; $r = -.591$, $p < .01$ respectively). This confirms the hypothesis (1c) that English is used less as age increases. A one-way ANOVA was also performed to show the use of English responses across the six age groups available in the present study. The ANOVA results are presented in Table 4.20 and Figure 4.4 below. This also replicates the findings of the simple linear correlations, showing that English responses decreased with increasing age. It would be interesting to look at these results again in, say, 30 years' time, or so, as the possibility of cohort effects may be significant.

TABLE 4.19

		Number of English responses Complete set	Number of English responses Reduced set
Age	Pearson Correlation	-.431**	-.591**
	P - value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between age and English responses

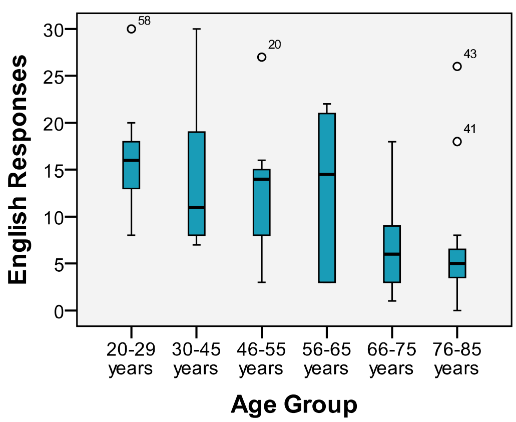
Post hoc comparison tests were also carried out in order to identify the true significant differences between the different age groups in terms of the mean number of English responses used. Although age group and the number of English responses were found to be negatively correlated to each other, post hoc comparisons did not identify significant differences between the various age groups when the complete set of data was used. On the other hand, significant differences were evident when post hoc comparisons were carried out on the reduced set. Significant differences in the number of English responses were found between the mean scores of the eldest age group 6 (76 – 85 years) and age groups 1 (20 – 29 years) and 2 (30 – 45 years). A significant difference between age group 5 (66 – 75 years) and age group 1 (20 – 29 years) was also evident (see Table 4.21).

TABLE 4.20

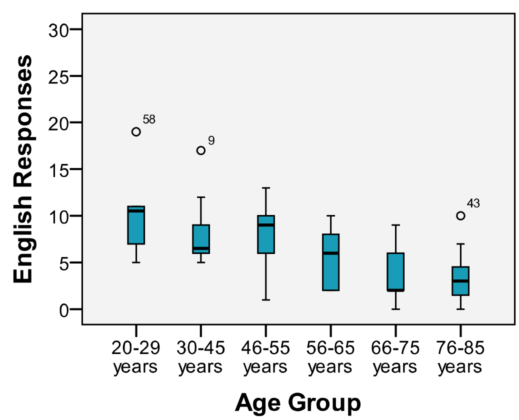
		Number of English responses Complete Set		Number of English responses Reduced Set	
Age Group	**N**	**Mean**	**SD**	**Mean**	**SD**
20-29 years	10	16.00	6.236	9.90	3.929
30-45 years	10	14.00	8.069	8.20	3.706
46-55 years	10	12.70	6.767	7.70	3.561
56-65 years	10	12.70	8.070	5.60	3.026
66-75 years	9	7.11	5.862	3.89	3.296
76-85 years	11	7.27	7.799	3.45	3.012
Total	**60**	**11.63**	**7.680**	**6.45**	**4.048**

Mean number of English responses for each age group in the complete and reduced sets

FIGURE 4.4



Complete Set



Reduced Set

Number of English responses produced with increasing age

TABLE 4.21

Age Group	Age Groups	Mean Difference	Std. Error	P-value	95% Confidence Interval	
					Lower Bound	Upper Bound
Multiple Comparisons using the Reduced Set						
76-85 years	20-29 years	-6.45*	1.500	.001	-10.88	-2.01
	30-45 years	-4.75*	1.500	.029	-9.18	-.31
	46-55 years	-4.25	1.500	.068	-8.68	.19
	56-65 years	-2.15	1.500	.709	-6.58	2.29
	66-75 years	-.43	1.543	1.000	-4.99	4.13
66-75 years	20-29 years	-6.01*	1.578	.005	-10.67	-1.35
	30-45 years	-4.31	1.578	.085	-8.97	.35
	46-55 years	-3.81	1.578	.169	-8.47	.85
	56-65 years	-1.71	1.578	.885	-6.37	2.95
	76-85 years	.43	1.543	1.000	-4.13	4.99

*. The mean difference is significant at the 0.05 level.

Post Hoc multiple comparisons for the mean number of English responses

4.6.2 Hypothesis 2

Hypothesis 2a

The respondents will produce more correct responses if more educated.

Hypothesis 2b

The respondents will have a shorter mean reaction time if more educated.

Rationale

Numerous researchers have found a correlation between education and naming performance (e.g. Worral et al. 1995; Tallberg 2005; Patricacou et al. 2007; Zec et al. 2007a). Thus, it is also proposed that the current study will identify a difference between education groups in terms of naming accuracy and naming latency.

4.6.2.1 Correlations between education and score

Simple linear correlations were performed in order to identify the effect of education on score. One-way ANOVAs with education category as the independent variable were also performed. A Kolmogorov-Smirnov test confirmed that the dependent variable, score, had a normal distribution. Simple linear correlations with years of education as the independent variable showed that there is a significant correlation between education and score (Table 4.22).

TABLE 4.22

		Strict Score Complete set	Strict Score Reduced set	Lenient Score Complete set	Lenient Score Reduced set
Years of Education	Pearson Correlation	.679**	.634**	.678**	.659**
	P - value	.000	.000	.000	.000
	Sample size	60	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between education and score

All the four sets of data showed that there is a strong positive correlation between years of education and Score. Thus, this demonstrates that more correct responses are produced by individuals who are more educated, as proposed in hypothesis 2a.

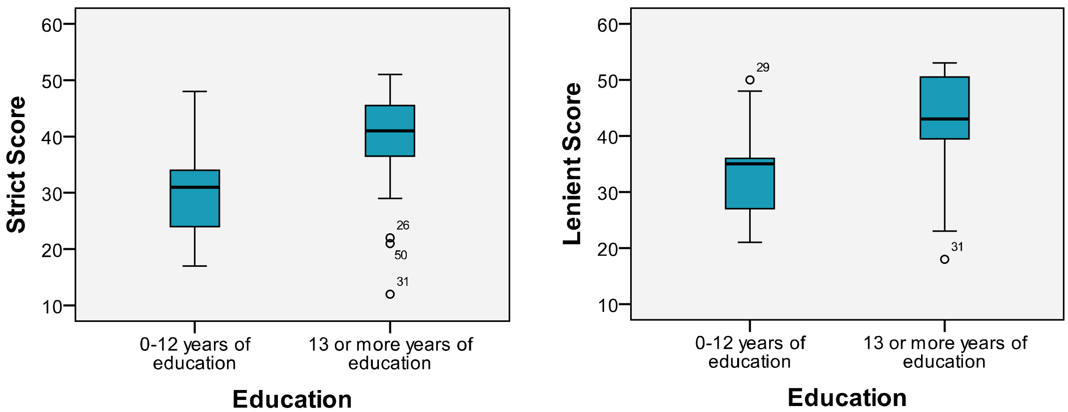
A one-way ANOVA was performed to make it possible to compare scores between the two education categories of the present study. A clear increase in the mean score is apparent in the higher educated group (Table 4.23 and Table 4.24). Figures 4.5 and 4.6 show a graphical representation of the mean score according to education category. A significant difference between the two education categories was further demonstrated by an independent samples t-test when comparing the two education categories in the complete set with strict scoring ($t(58) = -4.731$, $p = .000$), in the complete set with lenient scoring ($t(58) = -4.748$, $p = .000$), in the reduced set with strict scoring ($t(58) = -4.016$, $p = .000$), and in the reduced set with lenient scoring ($t(58) = -4.373$, $p = .000$).

TABLE 4.23

		Strict Score Complete Set		Lenient Score Complete Set	
Education category	**N**	**Mean**	**SD**	**Mean**	**SD**
0 – 12 years of education	29	29.45	7.371	32.79	7.208
13 or more years of education	31	39.42	8.831	43.00	9.241
Total	**60**	**34.60**	**9.523**	**38.07**	**9.723**

Mean scores for each education category using strict and lenient scoring in the complete set

FIGURE 4.5



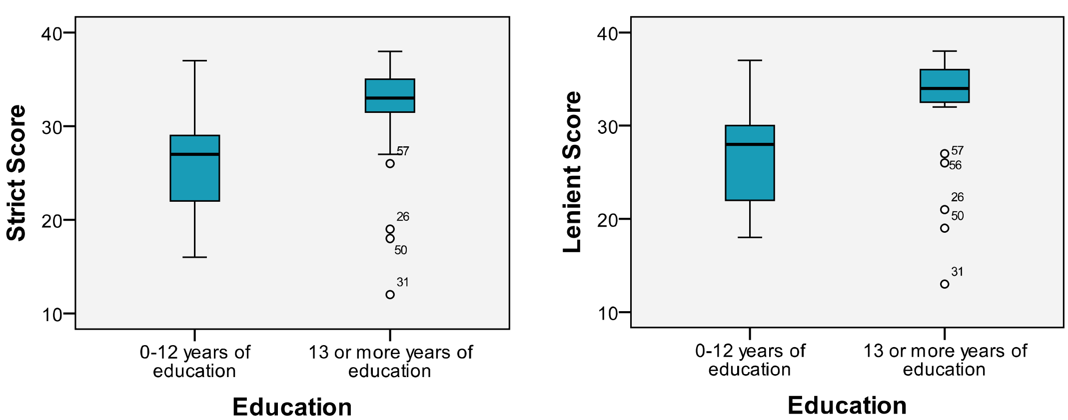
Mean scores with education category when using the complete set

TABLE 4.24

		Strict Score Reduced Set		Lenient Score Reduced Set	
Education category	**N**	**Mean**	**SD**	**Mean**	**SD**
0 – 12 years of education	29	25.97	5.088	26.52	4.983
13 or more years of education	31	31.68	5.867	32.65	5.805
Total	**60**	**28.92**	**6.171**	**29.68**	**6.201**

Mean scores for each education category using strict and lenient scoring in the reduced set

FIGURE 4.6



Mean scores with education category when using the reduced set

4.6.2.2 Correlations between education and reaction time

Simple linear correlations were performed in order to identify the effect of education on RT. One-way ANOVAs with education category as the categorical variable were also performed. A Kolmogorov-Smirnov test confirmed that the dependent variable, reaction time, had a normal distribution. Simple linear correlations with years of education as the independent variable showed that there is a significant negative correlation between education and reaction time. Two sets of data were analysed. These included only the strict scores in both complete and reduced sets (Table 4.25).

TABLE 4.25

		Mean RT Strict Score Complete set	**Mean RT Strict Score Reduced set**
Years of Education	Pearson Correlation	-.366**	-.457**
	P - value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between years of education and RT

Thus, both complete and reduced sets produced a moderate negative correlation when using strict scores ($r = -.366$, $p < .01$, and $r = -.457$, $p < .01$, respectively). This demonstrates that, when taken on its own, education has a negative effect on naming latency implying that reaction time becomes shorter with increasing years of education, as proposed in hypothesis 2b.

TABLE 4.26

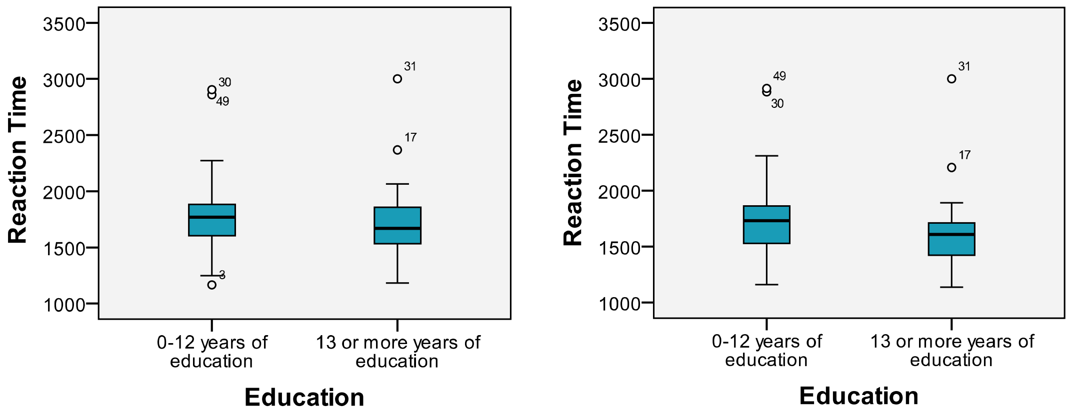
		Mean RT Complete Set using Strict scoring only		**Mean RT Reduced Set using Strict scoring only**	
Education Category	**N**	**Mean**	**SD**	**Mean**	**SD**
0 – 12 years of education	29	1818.83	402.403	1782.55	412.095
13 or more years of education	31	1727.87	343.054	1623.10	340.000
Total	**60**	**1771.83**	**372.54**	**1700.17**	**381.88**

Mean RT for each education category using strict scoring only

Means were compared to make it possible to compare reaction times between the two education categories of the present study. Table 4.26 and Figure 4.7 show the mean

RTs (in milliseconds) for each education category. Although a shorter mean reaction time is apparent in the higher educated group, there is not a significant difference in reaction time between the less and the more educated groups when using the complete set ($t(58) = .944$, $p = .349$) and when using the reduced set ($t(58) = 1.639$, $p = .107$).

FIGURE 4.7



Complete Set **Reduced Set**

Mean reaction time of the two education categories

Hypothesis 2c

The respondents will produce more English responses as education increases.

Rationale

The degree of English and Maltese that is used in schools in Malta varies depending on the type of school attended. While state schools use Maltese as their main language of instruction, the majority of church and private schools use English when addressing and teaching their pupils (Camilleri 1995). On the other hand, individuals who choose to further their studies beyond the obligatory years of education are required to use the English language as their main medium of study as most text books are in the English language. Therefore, individuals with more years of education are expected to produce more responses in English, than those with fewer years of Education, as a result of this greater exposure to the English language.

4.6.2.3 Correlations between education and English responses

Simple linear correlations were performed in order to identify the effect of education on the amount of English responses produced. As expected, a strong positive correlation between the number of years of education and the amount of English responses produced emerged. This was evident in both the complete set (r = .702, p < .01) and in the reduced set (r = .686, p < .01). This confirms the hypothesis that English is used more in naming the pictures used in this study as education increases.

TABLE 4.27

		Number of English responses complete set	**Number of English responses reduced set**
Years of Education	Pearson Correlation	.702**	.686**
	P-value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between education and English responses

Means were compared to show the use of English responses in the two levels of education in the present study. As shown in Table 4.28 and Figure 4.8 below individuals of higher education levels tend to use more English responses during a naming task than those with lower levels of education. This discrepancy is mostly evident in the complete set which resulted with a mean of 6.83 English responses for the lower education levels and a mean of 16.13 English responses produced by those who had 13 or more years of education. An independent samples t-test showed that there is a significant difference between the number of English responses produced by the two

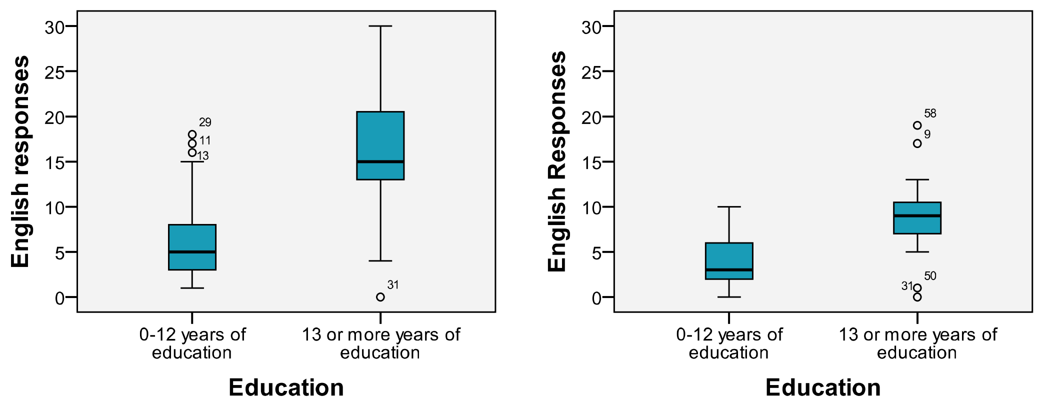
TABLE 4.28

		Mean number of English responses complete set		**Mean number of English responses reduced set**	
Education Category	**N**	**Mean**	**SD**	**Mean**	**SD**
0 – 12 years of education	29	6.83	4.796	4.14	2.863
13 or more years of education	31	16.13	7.164	8.61	3.818
Total	**60**	**11.63**	**7.680**	**6.45**	**4.048**

English responses produced by each level of education in the complete and reduced sets

education groups when using the complete set ($t(58) = -5.868$, $p = .000$), and when using the reduced set ($t(58) = -5.109$, $p = .000$). A greater variability in the amount of English responses produced by the higher education respondents is evident. This is demonstrated by the higher standard deviation of the group with 13 or more years of education. In fact, the number of English responses produced by this higher education group ranged from 0 to 30 English responses (out of 60).

FIGURE 4.8



Complete Set **Reduced Set**
Number of English responses produced by level of education

4.6.3 Hypothesis 3

Hypothesis 3a

Male respondents will produce more correct responses.

Hypothesis 3b

Male respondents will have a shorter mean reaction time.

Rationale

Different studies have found conflicting results concerning the relationship between gender and naming performance. Some researchers have proposed that the BNT is biased towards better performance for males, as a number of items included in the test are easier for males to name than females (Randolph et al. 1999). Thus, it is proposed that the males involved in the present study will also exhibit better naming performance.

Furthermore, the above hypotheses are also based on the consideration that today's older Maltese females had fewer opportunities for formal education and occupation outside the home, and therefore, were less exposed to a variety of linguistic and cultural experiences. Thus, it is expected that males would perform better than females on the BNT.

4.6.3.1 The relationship between gender and score

A one-way ANOVA with gender as the independent variable was performed. A Kolmogorov-Smirnov test confirmed that the dependent variable, score, had a normal distribution. A significant difference between the scores obtained by males and females was found in the complete set only. Table 4.29 presents these results.

TABLE 4.29

Complete Set					
	Type III Sum of Squares	df	Mean Square	F	P-value
Gender and Strict score	420.758	1	420.758	4.950	.030
Error	4929.642	58	84.994		
Gender and Lenient score	437.166	1	437.166	4.932	.030
Error	5140.567	58	88.630		
Reduced Set					
Gender and Strict score	115.406	1	115.406	3.141	.082
Error	2131.177	58	36.744		
Gender and Lenient score	116.705	1	116.705	3.145	.081
Error	2152.278	58	37.108		

Tests of Between-Subjects Effects (gender and score)

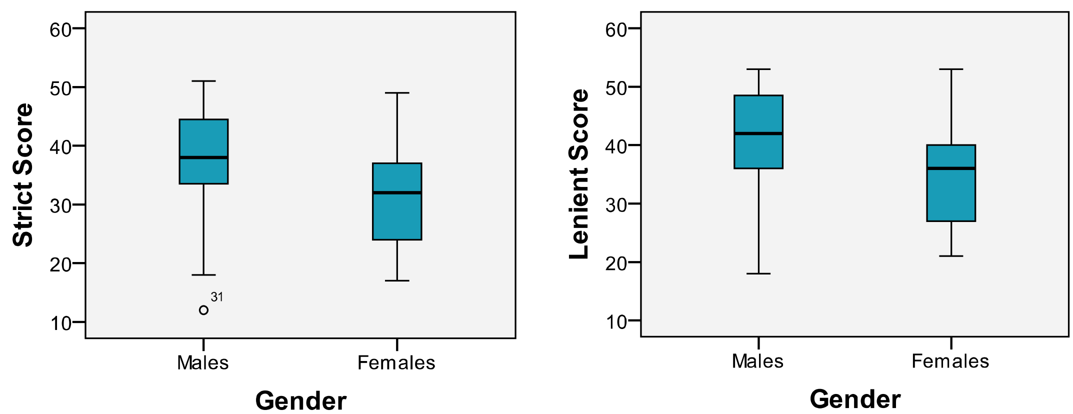
The complete set of data (using the full 60 items of the BNT) showed that there is a significant difference between scores obtained by males and females. This difference was evident when using both strict and lenient scoring as their p-values (0.030 for both sets) were both less than the 0.05 level of significance. An independent samples t-test confirmed this result ($t(58) = 2.225$, $p = .030$) and ($t(58) = 2.221$, $p = .030$) for strict and lenient scoring respectively. The coefficient of determination (R Squared) is 0.79, implying that 79% of the total variation of the dependent variable, score, is explained by the model. Therefore, as shown in Table 4.30 and Figure 4.9, males scored significantly higher than females when using only the complete set of the BNT, and when using both strict and lenient scoring.

TABLE 4.30

		Strict Score Complete Set		Lenient Score Complete Set	
Gender	**Sample size**	**Mean**	**SD**	**Mean**	**SD**
Males	31	37.16	9.543	40.68	9.428
Females	29	31.86	8.859	35.28	9.399
Total	**60**	**34.60**	**9.523**	**38.07**	**9.723**

Mean scores for males and females using strict and lenient scoring in the complete set

FIGURE 4.9



Mean scores with gender when using the complete set

The reduced set did not demonstrate the same difference in score between males and females. A significant difference was not found in both strict and lenient scoring when the reduced set of pictures was used. This was once again confirmed by a t-test $t(58) = 1.772$, $p = .082$ for strict scoring and $t(58) = 1.773$, $p = .081$ for lenient scoring. As shown in Table 4.29, the p values of the strict and lenient scoring (0.082 and 0.081) were greater than the 0.05 level of significance. A possible reason for not identifying a gender effect in the reduced set may be because once the *difficult* items were removed both males and females were able to produce a similar amount of correct responses.

4.6.3.2 The relationship between gender and reaction time

A one-way ANOVA with gender as the independent variable and reaction time as the dependent variable was performed. A Kolmogorov-Smirnov test confirmed that the dependent variable, reaction time, had a normal distribution. A significant difference in reaction time between males and females was not found in the complete or reduced sets, as the corresponding p-values (0.317 and 0.603 respectively) exceed the 0.05 level of significance (Table 4.31). This was also shown by an independent samples t-test for both the complete set ($t(58) = 1.009$, $p = .317$), and the reduced set ($t(58) = 0.522$, $p = .603$). The hypothesis stated above was rejected as males did not respond significantly quicker than females in the current study.

TABLE 4.31

Complete Set					
	Type III Sum of Squares	df	Mean Square	F	P-value
Gender and Reaction time	141324.362	1	141324.362	1.019	.317
Error	8047213.971	58	138745.068		
Reduced Set					
Gender and Reaction time	40276.086	1	40276.086	.273	.603
Error	8563710.247	58	147650.177		

Tests of Between-Subjects Effects (gender and reaction time)

Hypothesis 3c

Females will produce more English responses than males.

Rationale

Sciriha (1997) found that in Malta females tend to use more English than males in their everyday use of language. Therefore, female respondents in this study were expected to produce more responses in English, than males.

4.6.3.3 The relationship between gender and English responses

A one way ANOVA was performed in order to identify the effect of gender on the amount of English responses produced. A Kolmogorov-Smirnov test confirmed that the dependent variable, 'number of English responses', had a normal distribution. A significant difference between the number of English responses used by males and females was not found in either set (Table 4.32). The p-values of the complete and reduced set (0.166 and 0.280 respectively) exceed the 0.05 level of significance. Thus, the hypothesis stated above was rejected as the current study did not demonstrate a significant difference between males and females in the number of English responses produced.

TABLE 4.32

	Type III Sum of Squares	df	Mean Square	F	P-value
Complete Set					
Gender and English responses	114.207	1	114.207	1.968	.166
Error	3365.726	58	58.030		
Reduced Set					
Gender and English responses	19.402	1	19.402	1.188	.280
Error	947.448	58	16.335		

Tests of Between-Subjects Effects (gender and English responses)

4.6.4 Summary of findings of previous sections

The previous sections analyzed relationships and differences between individual demographic variables (age, education and gender) and dependent variables (score, reaction time and English responses). It is now essential to look into the interactions between the said demographic variables in order to determine the importance of each variable when all are analysed together. However, it is first necessary to summarise the findings of the sections described above.

When different sets of data (i.e. strict scoring complete set, lenient scoring complete set, strict scoring reduced set and lenient scoring reduced set) were used to analyse the effect of the predictor variables age and education on scores it was evident that different sets were correlated with each other. Thus, respondents who scored high on the complete set would also score high on the reduced set. Likewise, those who scored high when the strict scoring method was used also scored high when a lenient scoring method was used. Therefore, whichever set is chosen for analysis, similar results would emerge for the predictor variables age and education. On the other hand, the scores that resulted from the different sets were found to be significantly different and the highest mean score was achieved in the reduced set when lenient scoring was used.

From the analysis described in the previous sections it is evident that age and score are negatively correlated with each other (i.e. score decreases with increasing age), age and reaction time are positively correlated (i.e. RT increases with increasing age) and age is negatively correlated with number of English responses (i.e. the number of English responses decreases with increasing age). These correlations were evident in all sets of data.

Education was also found to be a significant predictor. Education was found to be highly correlated with score (i.e. score increases with higher education), negatively correlated with reaction time (ie. RT decreases with higher education) and also positively correlated with the number of English responses (i.e. English responses increased with higher education). These correlations were also evident in all sets of data.

Predictor variable gender was found to be correlated with score in the complete set of data only. Therefore, males scored higher than females when using both strict and lenient scoring in the complete set. However, this difference in scores obtained by males and females disappeared when the reduced set of items was used. Further, there was no evidence of a significant difference between males and females with regards to reaction time and the number of English responses.

4.7 REGRESSION ANALYSES

The previous sections used simple linear correlations and one-way ANOVAs to examine differences between a dependent variable and a predictor variable. However, the aim of this study is to investigate the collective effect of the predictor variables on the dependent variable. After all, particular predictors may lose their importance when these are examined in the presence of other predictors (Camilleri 2009). The sections that follow aim to identify the most important predictor variable for score, reaction time and English responses by means of regression analyses. Regression analysis may take the form of multiple regression or ANOVA regression analysis. ANOVA regression models are extensions of multiple regression models because they accommodate categorical variables and may need to be used when such variables are present. Only the reduced strict scoring set will be used during this part of the analysis. Since only the reduced set of data will be used when analyzing word variables (see Chapter 5) it was decided that only this set will be used during regression analysis of demographic variables too. The interaction between the most significant demographic variable(s) and the most significant word variable(s) will be investigated at the end of Chapter 5.

Therefore, it is necessary to use the same set of data when performing regression analysis of both demographic and word variables. Another reason for using the reduced set of data during analysis is that the performance of the Maltese Sample A was similar to that of studies on other populations when using the reduced set and not when using the complete set.

4.7.1 Interaction of demographic variables on score

This section looks at the interaction between predictor variables age and education and dependent variable strict score in the reduced set. Since gender was not found to be significantly related to strict score in the reduced set it was not included in this regression analysis. It is necessary to note that some of the sets of data involve very small numbers of participants. For example, there are only 5 respondents over 66 years with 13 or more years of education, and 15 with 0 – 12 years of education.

Table 4.33 clearly shows a summary of the correlations that exist between the dependent variable, strict score, and the predictor variables age and years of education. It is evident that the two predictor variables are correlated with each other. This creates multicollinearity problems in the regression analysis and interferes with identifying the true effect of each predictor on the dependent variable (Camilleri 2009). Thus, a stepwise regression analysis was carried out to identify the most significant predictor variable. Education emerged as the most important predictor for strict score. The coefficients for education (0.890) indicated that for every extra 1 year of education, the strict score is expected to increase by 0.89. Age remained a significant predictor variable as it was not excluded from the model (Table 4.34).

TABLE 4.33

		Age	Years of Education	Strict Score
Age	Pearson Correlation P-value	1.000	-.552** .000	-.602** .000
Years of Education	Pearson Correlation P-value	-.552** .000	1.000	.634** .000
Strict Score	Pearson Correlation P-value	-.602** .000	.634** .000	1.000

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between strict score and age and education

TABLE 4.34

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	17.659	1.908		9.256	.000
	Years of education	.890	.143	.634	6.241	.000
2	(Constant)	26.845	3.370		7.966	.000
	Years of education	.609	.159	.434	3.834	.000
	Age	-.107	.033	-.362	-3.205	.002

Coefficients (score - age and education)

The coefficient of determination (R squared) was 0.402 implying that this model explains 40.2% of the total variation of the dependent variable.

A two-way ANOVA regression analysis was also performed in order to analyze the differences between the different groups of the strongest predictor variable, education. The parameter estimates table shows that the mean strict score for respondents that have between 0 and 12 years of education is 5.712 words less than those with 13 or more years of education.

TABLE 4.35

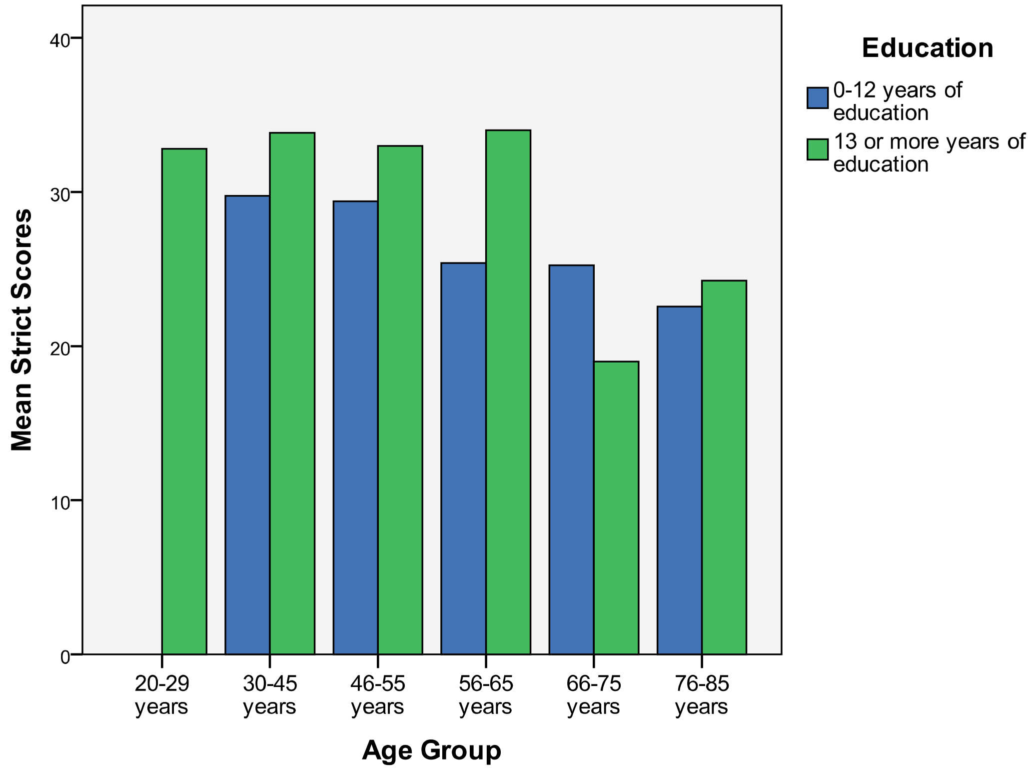
Parameter	B	Std. Error	t	P-value	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	31.677	.989	32.038	.000	29.698	33.657
[education=1]	-5.712	1.422	-4.016	.000	-8.559	-2.865
[education=2]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Parameter Estimates (strict score)

Figure 4.10 shows that the respondents with higher levels of education all obtained higher mean scores than those with lower education, except for the 66 – 75 years age group. The latter may be explained by the number of respondents in each group. As noted above, the older age groups include only a few respondents with higher education. On the other hand, the youngest age group included only people who were enrolled in University courses and, therefore, were all with 13 or more years of education. Figure 4.11 provides a graphical description of the number of respondents in each age group.

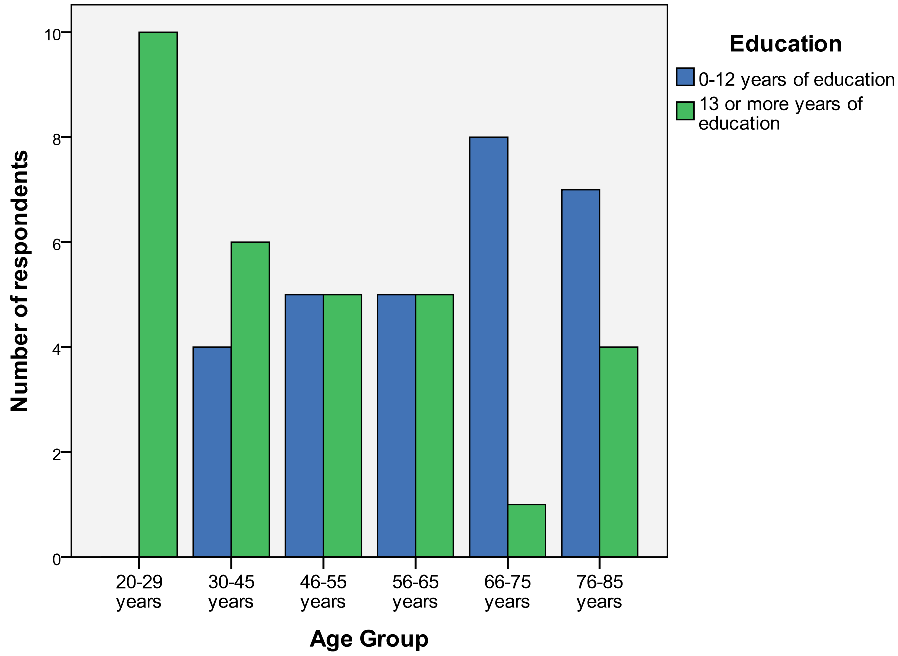
Figure 4.10



Mean strict score in each group – using the reduced set

There is only one respondent with higher education aged 66 – 75 years, in contrast to the eight respondents with 0 – 12 years of education. In the 76 – 85 years age group there are 4 respondents with higher education and 7 respondents with lower education. Therefore, cohort effects are inevitable as the older respondents involved in this study are a reflection of the older Maltese population. With time this situation is gradually changing as more opportunities for education have become available. It is possible that a similar study, in say 10-20 years' time, may produce different results, implying that regular updating of normative data and standardized tests is necessary.

Figure 4.11



Number of respondents in each group

4.7.2 Interaction of demographic variables on reaction time

This section aims to interact the predictor variables age and years of education with dependent variable reaction time by using multiple regression analysis. The predictor variable gender will not be included in the analysis as there was no evidence of a

TABLE 4.36

		Age	Years of education	Mean reaction time
Age	Pearson Correlation	1.000	-.552**	.475**
	P - value		.000	.000
	Sample size	60	60	60
Years of education	Pearson Correlation	-.552**	1.000	-.457**
	P - value	.000		.000
	Sample size	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between reaction time and age and education

significant difference between males and females in terms of RT (see section 4.6.3.2). There is, however, evidence that age and years of education are correlated with reaction time (see sections 4.6.1.2 and 4.6.2.2). Table 4.36 reproduces the correlation results.

Apart from the correlations mentioned, it is also obvious that the quantitative predictor variables, age and years of education, are correlated with each other (r = -.552, $p < .001$). This may lead to multicollinearity problems which have the potential of interfering with the analysis of the effects of each predictor. Therefore, a stepwise method in the regression analysis was used to identify the best model.

The stepwise regression analysis (Table 4.37) indicates that age is the most significant predictor of reaction time and it is followed by years of education. The coefficients for age (8.669) indicates that for every extra year the reaction time is expected to increase by 8.669 milliseconds.

A two-way ANOVA regression analysis was also performed in order to analyze the differences between the groups of the strongest predictor variable, age. The parameter estimates table shows that the mean reaction time of respondents aged between 20 and 65 is between 440 and 600 milliseconds less than that of a respondent aged at least 76 years. The mean reaction time increases markedly after 66 years of age.

TABLE 4.37

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	1243.312	119.387		10.414	.000
	Age	8.669	2.108	.475	4.113	.000
2	(Constant)	1699.366	248.516		6.838	.000
	Age	5.850	2.459	.321	2.379	.021
	Years of Education	-24.309	11.712	-.280	-2.076	.042

Coefficients (reaction time - age and education)

TABLE 4.38

Parameter	B	Std. Error	t	P-value	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	2095.273	95.745	21.884	.000	1903.316	2287.229
[Age group=20-29 years]	-590.373	138.747	-4.255	.000	-868.544	-312.201
[Age group=30-45 years]	-440.173	138.747	-3.172	.002	-718.344	-162.001
[Age group=46-55 years]	-599.873	138.747	-4.323	.000	-878.044	-321.701
[Age group=56-65 years]	-560.973	138.747	-4.043	.000	-839.144	-282.801
[Age group=66-75 years]	-199.162	142.728	-1.395	.169	-485.314	86.990
[Age group=76-85 years]	0[a]	.	.	.	.	.

Parameter Estimates (reaction time)

4.7.3 Interaction of demographic variables on English responses

This section aims to explore the interaction between the predictor variables, age and years of education, with dependent variable number of English responses, by using multiple regression analysis. The predictor variable, gender, will not be included in the analysis as there was no evidence of a significant difference between males and females in terms of the number of English responses (see section 4.6.3.3). Table 4.39 reproduces the correlations found in relation to the number of English responses.

As seen in section 4.7.1, the predictor variables, age and years of education, are correlated with each other. Therefore, once again a stepwise multiple regression analysis was performed to avoid possible multicollinearity problems. Stepwise regression analysis (see Table 4.40) showed that years of education is the most significant predictor of number of English responses. This is followed by predictor

TABLE 4.39

		Age	Years of Education	Number of English responses
Age	Pearson Correlation	1.000	-.552**	-.591**
	P – value (2-tailed)		.000	.000
	Sample size	60	60	60
Years of education	Pearson Correlation	-.552**	1.000	.686**
	P – value (2-tailed)	.000		.000
	Sample size	60	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between English responses and age and education

variable age. The coefficients for years of education (0.632) indicates that for every extra year of education the number of English responses is expected to increase by 0.632 words.

TABLE 4.40

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	-1.547	1.177		-1.314	.194
	Years of Education	.632	.088	.686	7.187	.000
2	(Constant)	3.521	2.116		1.664	.102
	Years of Education	.477	.100	.518	4.785	.000
	Age	-.059	.021	-.305	-2.815	.007

Coefficients (English responses – age and education)

A two-way ANOVA regression analysis was also performed in order to analyze the differences between the different groups of the strongest predictor variable, education. The parameter estimates table shows that the mean number of English responses for respondents that have between 0 and 12 years of education is 3.193 words less than those with 13 or more years of education.

TABLE 4.41

Parameter	B	Std. Error	t	P-value	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	5.486	1.129	4.862	.000	3.223	7.750
[Age group=20-29 years]	4.414	1.505	2.932	.005	1.394	7.433
[Age group=30-45 years]	3.991	1.395	2.860	.006	1.192	6.789
[Age group=46-55 years]	3.810	1.383	2.755	.008	1.036	6.584
[Age group=56-65 years]	1.710	1.383	1.237	.222	-1.064	4.484
[Age group=66-75 years]	1.241	1.437	.864	.392	-1.641	4.122
[Age group=76-85 years]	0[a]	.	.	.	.	.
[Education=0-12 years]	-3.193	.957	-3.336	.002	-5.113	-1.273
[Education=13 years or more]	0[a]	.	.	.	.	.

Parameter Estimates (English responses)

4.8 CONCLUSION

The present chapter has looked at many different aspects related to the performance of Maltese speakers on the BNT. It will conclude by describing the most important features that emerged.

First and foremost, the need for a modified version of the BNT became evident from the name agreement analysis of the preliminary study (using Sample A). Items with low name agreement were removed to produce a reduced set. This reduced set consists of 38 items which have at least 70% name agreement. It is believed that this set provides a better reflection of the naming abilities of Maltese speakers since it assesses naming performance on items which are more culturally appropriate.

Since the BNT was originally developed and normed on a diverse linguistic and cultural population, a number of items of the test had more than one possible response. The dominant response for each BNT item was obtained from the data collected from Sample A. This dominant response was taken to be the target response during the main study. All the other possible responses for each item – including English translation equivalents – were regarded as acceptable alternative responses since they could not be regarded as incorrect responses per se. In order to give credit to target responses as well as to alternative responses, a *lenient* scoring method is being recommended. This method allows clinicians to cater for the bilingual situation that is present in Malta, such that English (when the target word is in Maltese) or Maltese (when the target word is in English) responses are credited as correct in lenient scoring. Strict scoring (where target responses only are scored as correct) procedures may still be used if preferred.

The present chapter also compared the findings obtained from this study to those reported by other studies. Comparisons were made in terms of scores and name agreement across some studies. This chapter also compared the results of the complete set with those of the reduced set, the results of the preliminary study with those of the main study, and also compared the results of the two types of scoring methods (strict scoring and lenient scoring). These comparisons are discussed in Chapter 6.

The summary of the results derived from simple linear correlations and one-way ANOVAs provided preliminary information about the importance of the different predictor variables (see section 4.6.4). From the previous section, conclusions regarding the most important predictors can be made. It was evident that education was the strongest predictor of score and of the number of English responses produced. The mean strict score of an individual with more than 12 years of education is expected to be 5.712 points higher than that of an individual with less education. English responses are expected to increase by 0.632 words for every extra year of education and less educated individuals are expected to produce 3.193 English words less than their more educated counterparts.

Although age did not emerge as the strongest predictor it is still an important predictor of score as it was not excluded from the regression analysis model. Overall there was a decline in the mean score obtained as age increased. This occurred in both levels of education with a sharp decline in score from 66 years of age in those with 13 or more years of education and a slight decline in score from 56 years of age in the lower education group (see Figure 4.10). The small numbers of participants in each age and education group should be taken into consideration when attempting to interpret these results. In fact, there are only 5 respondents with 13 or more years of education over 66 years of age, in contrast to the 15 respondents with 0-12 years of education over 66 years of age. This is a reflection of the sociological reality that exists in Malta, as older individuals (especially females) with higher levels of education are more scarce (see section 3.5.4 in Chapter 3). A similar situation has also been reported in the Greek study on the BNT (Patricacou et al 2007).

On the other hand, age was found to be the strongest predictor of RT. After 66 years of age there is evidence of a marked increase in the mean RT, and for every extra year of age, RT is expected to increase by 8.669 milliseconds. The number of years of education also remained a predictor of RT as it was not excluded from the regression model. However, when ANOVA analysis was performed, the difference between the mean RTs of the two groups of education (0-12 years and 13 or more years) was not found to be significant.

The following chapter presents the results related to word variables and their influence on naming performance. The interaction between demographic variables (results of this chapter) and word variables will also be investigated towards the end of the next chapter. The findings are discussed further in Chapter 6.

Chapter 5

Results

Word variables and people without aphasia

5.0 INTRODUCTION

The previous chapter was concerned with analysing the demographic variables that influence naming accuracy, reaction time and the number of English responses. There was evidence that age, education and (to a smaller extent) gender may all affect naming performance. However, it is also necessary to analyse the influence of the material that is used to obtain information on naming ability. This chapter aims to investigate the potential effect of word variables on the performance of Maltese speakers on the BNT. Chapter 2 provided an overview of the literature related to word variables (section 2.4.2 in Chapter 2) and several proposals were made regarding the possible findings of this study in relation to word variables. These hypotheses will be stated and tested in a similar fashion to that used in the previous chapter on demographic variables. However, prior to commencing with result presentation, it is essential to provide a brief description of the data available.

5.1 THE AVAILABLE DATA

The word predictor variables that are available in the present study include both categorical and quantitative variables. The categorical variables include animacy and cognate status, while the quantitative variables include age of acquisition, familiarity, imageability, syllable and phoneme length and name agreement. Chapter 3 provides a description of how the data related to word variables was obtained (see sections 3.3 and 3.4).

Three different dependent variables emerge from the present data. These include the *number correct* (i.e. the number of correct responses for each item), the *mean reaction time* for each item (RT) and the number *of English responses* produced for each item. Each of these dependent variables will be correlated with the above mentioned predictor variables.

Although four sets of data are also available for word variables, namely complete and reduced sets, each including strict and lenient number correct, it was decided to use only the reduced set during analysis with word variables. Using the reduced set ensures that only the words/items that had at least 70% name agreement are included in the study. If the complete set of words were to be used in the analysis, the results would have included data from a substantial number of words with low name agreement and words with different word variables (see also section 4.7 in Chapter 4). Both strict and lenient number correct will be used when analyzing dependent variable 'number correct'. However, the analysis concerned with the mean reaction time for each word will only utilise the strict correct responses.

5.1.1 Reaction time measurement

Some studies have observed word initial fricative/affricate effects on naming latency. These effects are said to arise from a delay in the detection and registration of fricative and affricate sounds by the voice key apparatus that is used to record and measure Reaction times (Sakuma et al 1997, cited in Tyler et al 2005). This results in a longer mean reaction time on items that have a fricative or an affricate in word initial position. Thus, it was necessary to investigate whether a similar effect would be observed in the present study too.

A DMDX voice reaction timer was utilised to record reaction times during the present study. The reduced set was made up of 12 (31.6%) items that started with a fricative/affricate and 26 (68.4%) items without a fricative/affricate at the beginning of the word.

A one-way ANOVA was performed in order to identify the effect of word initial fricative/affricate on reaction time. In this study, reaction time was not affected by

TABLE 5.1

	Type III Sum of Squares	df	Mean Square	F	P-value
Word Initial Fricative and Reaction Time	106188.624	1	106188.624	.456	.504
Error	8385426.455	36	232928.513		

Tests of Between-Subjects Effects (word initial fricative and reaction time)

word initial fricative/affricate as the corresponding p value (p = .504) exceeds the 0.05 level of significance. Therefore, in the present study, there appears to be no difference in reaction time between words which start with a fricative or affricate and those which do not (Table 5.1).

5.2 HYPOTHESES TESTING

This section on word variables will state each hypothesis and present the related findings from the present study. Regression analysis was performed to identify the most important predictor/s for each dependent variable. Simple linear correlations and one-way ANOVAs are given first, however. The predictors that are being used in the analysis described below include name agreement, age of acquisition, familiarity, imageability, word length, cognate status and animacy.

5.2.1 Hypothesis 1

Hypothesis 1a

The items will have more correct responses as name agreement increases.

Hypothesis 1b

The items will have a less mean reaction time as name agreement increases.

Rationale

The reduced set of data includes only items which had high name agreement (i.e. at least 70% name agreement). Analysis of correlations between name agreement and number of correct responses were carried out. When a dominant name for a particular item is used by the majority of people, name agreement is said to be high and it is proposed that this would be easier to name as a result of less competition between alternative names. In contrast when a number of alternative names may be used to name a particular item, name agreement is low. The latter case may pose difficulties in naming as retrieval of the chosen name may be affected by the competing alternative names. Thus, it is proposed that words with high name agreement will be named more correctly and named faster than those with lower name agreement.

5.2.1.1 Correlations between name agreement and correct responses

Simple linear correlations were performed in order to identify the effect of name agreement on the number of correct responses. A Kolmogorov-Smirnov test confirmed that the dependent variable, number of correct responses, had a normal distribution, as the p-value exceeds the 0.05 level of significance in both strict number correct and lenient number correct (Table 5.2).

TABLE 5.2

	STRICT number correct	LENIENT number correct	Reaction Time
Kolmogorov-Smirnov Z	1.036	1.193	.725
P-value (2-tailed)	.233	.116	.669

One-Sample Kolmogorov-Smirnov Test (correct responses)

Simple linear correlations with name agreement as the independent variable showed that there is a significant positive correlation between name agreement and number of correct responses (Table 5.3). Both strict and lenient sets produced a strong positive correlation ($r = .529$, $p < .01$ and $r = .504$, $p < .01$ respectively). This demonstrates that, when taken on its own, name agreement has a positive effect on naming accuracy, with more correct responses being produced the higher the name agreement, as proposed in hypothesis 1a.

5.2.1.2 Correlations between name agreement and reaction time

Simple linear correlations were performed in order to identify the effect of name agreement on mean reaction time. A Kolmogorov-Smirnov test confirmed that the dependent variable, reaction time, had a normal distribution as the p-value exceeds the 0.05 level of significance (Table 5.2).

Simple linear correlations with name agreement as the independent variable showed that there is a significant negative correlation between name agreement and reaction time (Table 5.3). A moderate negative correlation ($r = -.415$, $p < .01$) was found between name agreement and reaction time. When taken on its own, name agreement has a negative effect on naming latency, implying that items with higher name agreement are named faster than those with lower name agreement, as proposed in hypothesis 1b.

Hypothesis 1c

The items will have fewer English responses as name agreement increases.

Rationale

In hypothesis 1c it is proposed that the possibility of producing an English response increases when name agreement is low. This is based on the notion that when people do not fully agree on the name of an item they are more likely to look for alternative names, which may also include the English translation equivalents of the Maltese target word. On the other hand, when Maltese speakers agree on the name of an item (especially in the reduced set, which included only the high name agreement items) it is proposed that there is less competition from other words, including the English equivalent words, and that the Maltese target word would be retrieved and produced, rather than the English equivalent.

5.2.1.3 Correlations between name agreement and English responses

Simple linear correlations were performed in order to identify the effect of name agreement on the number of English responses produced for different items. A significant negative correlation between name agreement and English responses was found (Table 5.3). A moderate negative correlation ($r = -.412$, $p < .05$) was found between name agreement and English responses. When taken on its own, name agreement has a negative effect on the number of English words produced, implying

TABLE 5.3

		Strict Number Correct	Lenient Number Correct	Reaction Time	Number of English responses
Name Agreement	Pearson Correlation	.529**	.504**	-.415**	-.412**
	P - value	.001	.001	.010	.010
	Number of items	38	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between name agreement and number correct, RT and English responses

that items with higher name agreement tend to have fewer English responses, as proposed in hypothesis 1c.

5.2.2 Hypothesis 2

Hypothesis 2a

The items will have more correct responses as age of acquisition decreases.

Hypothesis 2b

The items will have a less mean reaction time as age of acquisition decreases.

Rationale

Chapter 2 presented the findings of a number of studies which indicated that age of acquisition plays a very important role in naming performance. It has been said to be a very strong predictor of both reaction time and number correct for each item. Thus, it is proposed that similar results will emerge from this study too.

5.2.2.1 Correlations between age of acquisition and correct responses

Simple linear correlations were performed in order to identify the effect of age of acquisition on the number of correct responses. A significant negative correlation between age of acquisition and number of correct responses was found (Table 5.4). Both strict and lenient sets produced a strong negative correlation ($r = -.761$, $p < .01$ for both sets). This demonstrates that, when taken on its own, age of acquisition has a strong negative effect on naming, indicating that words that are learned early tend to be named more correctly than late acquired words. This satisfies hypothesis 2a.

5.2.2.2 Correlations between age of acquisition and reaction time

Simple linear correlations were performed in order to identify the effect of age of acquisition on mean reaction time. A strong positive correlation ($r = .746$, $p < .01$) was found between age of acquisition and reaction time (Table 5.4). When taken on its own, age of acquisition has a positive effect on naming, implying that items with young age of acquisition are named faster than those with older age of acquisition, as proposed in hypothesis 2b.

Hypothesis 2c

The items will have more English responses as age of acquisition increases.

Rationale

In hypothesis 2c it is proposed that there is a possibility of producing fewer English responses in items with early/young age of acquisition. Since the sample recruited in the present study was purposely selected to include only native speakers of Maltese it is proposed that words which were learned early in life were acquired primarily in Maltese. Thus, this may increase the possibility that early age of acquisition words are named in Maltese rather than in English, as they had been primarily acquired in Maltese rather than in English.

5.2.2.3 Correlations between age of acquisition and English responses

Simple linear correlations were performed in order to identify the effect of age of acquisition on the number of English responses produced for different items. A moderate positive correlation ($r = .326$, $p < .05$) was found between age of acquisition and the number of English responses (Table 5.4). Thus, when taken on its own, age of

TABLE 5.4

		Strict Number Correct	Lenient Number Correct	Reaction Time	Number of English responses
Age of Acquisition	Pearson Correlation	-.761**	-.761**	.746**	.326*
	P - value	.000	.000	.000	.046
	Number of items	38	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Correlations between AoA and number correct, RT and English responses

acquisition has a positive effect on the number of English responses produced, implying that items with young age of acquisition are more commonly named in Maltese than in English, as proposed in hypothesis 2c.

5.2.3 Hypothesis 3

Hypothesis 3a

The items will have more correct responses as familiarity increases.

Hypothesis 3b

The items will have a less mean reaction time as familiarity increases.

Rationale

Previous research has shown that familiarity is highly correlated with other word variables, such as age of acquisition and word frequency (see section 2.4.2.4 in Chapter 2). Since age of acquisition has been found to be highly correlated with number correct and reaction time in the present study, it is also proposed that similar findings will emerge with regards to familiarity too.

5.2.3.1 Correlations between familiarity and correct responses

Simple linear correlations were performed in order to identify the effect of familiarity on the number of correct responses. A significant positive correlation between familiarity and number of correct responses was found (Table 5.5). Both strict and lenient sets produced a strong positive correlation ($r = .640$, $p < .01$ and $r = .643$, $p < .01$ respectively). This demonstrates that, when taken on its own, familiarity has a positive effect on naming, indicating that words that are more familiar tend to be named more correctly, as proposed in hypothesis 3a.

5.2.3.2 Correlations between familiarity and reaction time

Simple linear correlations were performed in order to identify the effect of familiarity on mean reaction time. A strong negative correlation ($r = -.546$, $p < .01$) was found between familiarity and reaction time (Table 5.5). When taken on its own, familiarity has a negative effect on naming, implying that items which are highly familiar are named faster than those that are less familiar, as proposed in hypothesis 3b.

Hypothesis 3c

The items will have fewer English responses as familiarity increases.

Rationale

It is hypothesised that the more familiar one is with particular words and/or items, the easier it is to retrieve and name the word in one's native language, without requiring selection of the English translation equivalents. This is proposed while taking into consideration that the respondents involved in the study were all native speakers of Maltese, and that all learned Maltese as their first language. Thus, it is expected that words which are highly familiar will be named more in Maltese than in English.

5.2.3.3 Correlations between familiarity and English responses

Simple linear correlations were performed in order to identify the effect of familiarity on the number of English responses produced. No significant correlation was found (Table 5.5). Therefore, hypothesis 3c must be rejected.

TABLE 5.5

		Strict Number Correct	**Lenient Number Correct**	**Reaction Time**	**Number of English responses**
Familiarity	Pearson Correlation	.640**	.643**	-.546**	-.148
	P - value	.000	.000	.000	.374
	Number of items	38	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between familiarity and number correct, RT and English responses

5.2.4 Hypothesis 4

Hypothesis 4a

The items will have more correct responses as imageability increases.

Hypothesis 4b

The items will have a less mean reaction time as imageability increases.

Rationale

Previous research identified correlations between imageability and naming performance, in both the normal population and in people with aphasia (see section 2.4.2.5 in Chapter 2). Both the number of correct responses and reaction time are expected to be significantly correlated with imageability in the present study as well.

5.2.4.1 Correlations between imageability and number of correct responses

Simple linear correlations were performed in order to identify the effect of imageability on the number of correct responses. A significant positive correlation between imageability and number of correct responses was found (Table 5.6). Both strict and lenient sets produced a strong positive correlation ($r = .665$, $p < .01$ and $r = .653$, $p < .01$ respectively). This demonstrates that, when taken on its own, imageability has a positive effect on naming, indicating that words that are highly imageable tend to be named more correctly than less imageable words. This satisfies hypothesis 4a.

TABLE 5.6

		Strict Number Correct	Lenient Number Correct	Reaction Time	Number of English responses
Imageability	Pearson Correlation	.665**	.653**	-.625**	-.191
	P - value	.000	.000	.000	.250
	Number of items	38	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between imageability and number correct, RT and English responses

5.2.4.2 Correlations between imageability and reaction time

Simple linear correlations were performed in order to identify the effect of imageability on mean reaction time. A strong negative correlation ($r = -.625$, $p < .01$) was found between imageability and reaction time (Table 5.6). When taken on its own, imageability has a negative effect on naming, implying that items which are high in imageability are named faster than those with low imageability, as proposed in hypothesis 4b.

Hypothesis 4c

The items will have fewer English responses as imageability increases.

Rationale

It is hypothesised that the more imageable a word is, the easier it is to retrieve and name the word in one's native language, without requiring selection of the English translation equivalents. Since the respondents were all native speakers of Maltese, it is expected that words which are high in imageability will be named more in Maltese than in English.

5.2.4.3 Correlations between imageability and English responses

Simple linear correlations were performed in order to identify the effect of imageability on the number of English responses produced. No significant correlation was found (Table 5.6). Therefore, hypothesis 4c must be rejected.

5.2.5 Hypothesis 5

Hypothesis 5a

The items will have more correct responses as word length decreases.

Hypothesis 5b

The items will have a less mean reaction time as word length decreases.

Rationale

As noted in Chapter 2 there is disagreement regarding whether it is syllable length or phoneme length that has an effect on naming performance, if at all. Therefore, this section will attempt to correlate both syllable length and phoneme length with the number of correct responses and with reaction time, in order to decide whether word length is correlated with naming performance (accuracy and latency), and, if it is, to decide whether syllable or phoneme length influences naming in Maltese speakers.

5.2.5.1 Correlations between word length and correct responses

Simple linear correlations were performed in order to identify the effect of phoneme length and syllable length on the number of correct responses. A significant moderate negative correlation between word length and number of correct responses was found. Both syllable and phoneme length were correlated with number correct in both strict and lenient data (Table 5.7). This demonstrates that, when taken on its own, word length has a negative effect on naming, indicating that words that are shorter tend to be named more correctly than longer words, as suggested in hypothesis 5a.

5.2.5.2 Correlations between word length and reaction time

Simple linear correlations were performed in order to identify the effect of syllable length and phoneme length on mean reaction time. Only phoneme length was found to

TABLE 5.7

		Strict Number Correct	Lenient Number Correct	Reaction Time
Syllable Length	Pearson Correlation	-.401**	-.391**	.238
	P – value	.013	.015	.151
Phoneme Length	Pearson Correlation	-.426**	-.417**	.326*
	P – value	.008	.009	.046
	Number of items	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Simple linear correlations between word length and number correct and RT

be significantly correlated with reaction time ($r = .326, p < .05$). There was no significant correlation between syllable length and reaction time (Table 5.7). This implies that words with fewer phonemes may be named faster than words with more phonemes.

5.2.6 Hypothesis 6

Hypothesis 6a

The items that are cognate words will have more correct responses.

Hypothesis 6b

The items that are cognate words will have a less mean reaction time.

Rationale

Since cognate words are very similar translation equivalents it is proposed that cognate words would have less competition during lexical selection and retrieval than those words which are not cognates, and will, therefore, be named faster and more accurately than non cognate words.

5.2.6.1 The relationship between cognates and correct responses

The reduced set included 13 items (34.2% of the 38 items) with cognate status (e.g. *pyramid* – 'piramida'), and 25 items (65.8%) without cognate status (e.g. *funnel* – 'lembut').

A one-way ANOVA was performed in order to identify the effect of cognates on the number of correct responses. A significant difference between cognate and non cognate words with regards to the number of correct responses was found as the corresponding p-values (0.040 and 0.027 for strict and lenient number correct respectively) were less than the 0.05 level of significance (Table 5.8). This implies that cognate words are more likely to be named correctly, as proposed in hypothesis 6a.

5.2.6.2 The relationship between cognates and reaction time

A one-way ANOVA was performed in order to identify the effect of cognates on reaction time. There is evidence of a significant difference between cognate and non cognate words with regards to reaction time, as the corresponding p-value (0.007) is less than the 0.05 level of significance. Thus, this means that cognate words are named faster than non cognate words as proposed by hypothesis 6b.

Hypothesis 6c

The items that are cognate words will have more English responses.

Rationale

It is proposed that the English responses produced by the respondents would occur more in cognate words than in non cognate words as a result of the similarity between the English/Maltese names. Less similar words (non cognates) are expected to be named in Maltese, as the production of an English response would be too effortful to select and retrieve.

5.2.6.3 The relationship between cognates and English responses

A one-way ANOVA was performed in order to identify the effect of cognates on English responses. There is no evidence of a significant difference between cognate and non cognate words with regards to English responses as the corresponding p-value (0.278) exceeds the 0.05 level of significance (Table 5.8). Therefore, hypothesis 6c must be rejected.

TABLE 5.8

	Type III Sum of Squares	df	Mean Square	F	P-value
Strict Number Correct	2289.970	1	2289.970	4.523	.040
Error	18225.08	36	506.252		
Lenient Number Correct	2866.599	1	2866.599	5.335	.027
Error	19342.76	36	537.299		
Reaction Time	1576429.1	1	1576429.147	8.207	.007
Error	6915185.9	36	192088.498		
English Responses	584.178	1	584.178	1.215	.278
Error	17307.71	36	480.770		

Tests of Between-Subjects Effects (Cognates and dependent variables)

5.2.7 Hypothesis 7

Hypothesis 7a

The items that are animate words will have fewer correct responses.

Hypothesis 7b

The items that are inanimate words will have a less mean reaction time.

Rationale

The items in the present study were categorised into *living* and *non living* items in order to differentiate between animate and inanimate items respectively. It is proposed that during naming of animate items there is a greater activation of the numerous shared characteristics of the living/animate item. This creates conflict or competition to select and retrieve the appropriate name and, therefore, this leads to a disadvantage for both naming accuracy and latency (Janyan 2007).

5.2.7.1 Correlations between animacy and correct responses

The reduced set was made up of only 6 living items (15.8%) and 32 (84.2%) non-living items. A one-way ANOVA was performed in order to identify the effect of animacy on the number of correct responses. A significant difference was not evident between animate and inanimate words with regards to the number of correct responses, as the corresponding p-value (0.079) is greater than 0.05.

5.2.7.2 Correlations between animacy and reaction time

There is no evidence of a difference in reaction time between animate and inanimate words in the present study (Table 5.9).

TABLE 5.9

Source	Type III Sum of Squares	df	Mean Square	F	P-value
Strict Number Correct	1709.844	1	1709.844	3.273	.079
Error	18805.208	36	522.367		
Lenient Number Correct	2042.160	1	2042.160	3.645	.064
Error	20167.208	36	560.200		
Reaction Time	365188.860	1	365188.860	1.618	.212
Error	8126426.219	36	225734.062		

Tests of Between-Subjects Effects (Animacy and dependent variables)

It should be noted that the results obtained from the one-way ANOVA with animacy may not be reliable as a result of the disproportionate number of living (6) versus non living (32) items included in the reduced set. Thus, this implies that the data may not have been representative enough for animacy effects to appear.

5.3 SUMMARY OF FINDINGS OF PREVIOUS SECTIONS

The previous sections analysed the relationship between individual word variables (name agreement, age of acquisition, familiarity, imageability, word length, cognate status and animacy) and the dependent variables (number of correct responses for each item, reaction time and number of English responses for each item). It is now essential to look into the interactions between the predictor word variables in order to determine the importance of each variable when all are analysed together. A summary of the results found up to now will follow first, however.

The results related to word variables that have been obtained up to now from the present study provide evidence of numerous significant correlations between predictor and dependent variables. The number of correct responses has been found to be positively correlated with name agreement, familiarity, imageability and cognate status, but negatively correlated with age of acquisition and word length (including both syllable and phoneme length).

Reaction time has been shown to be negatively correlated with name agreement, familiarity and imageability, and positively correlated with age of acquisition and phoneme length. Cognate words are named faster than non cognates. On the other hand, animacy and syllable length were not found to be significant predictors of RT. There was no effect of word initial frictive or affricate on the measurement of reaction times (see Section 5.1).

The number of English responses was also found to be positively correlated with name agreement and age of acquisition. No relationship was found with familiairity, imageability and cognate status.

5.4 REGRESSION ANALYSES

The previous sections used simple linear correlations and one-way ANOVAs to examine relationships and differences between dependent variables and predictor variables. This section aims to investigate the collective effect of the predictor variables on each dependent variable in order to identify the most significant predictor(s) for each dependent variable. The significant predictor variables for the number of correct responses, the mean reaction time and the number of English responses for each item will be investigated by means of regression analyses (multiple regression analysis and ANOVA regression analysis).

As outlined in section 5.1 only the reduced set of the BNT is being used in the analysis on word variables, and only the strict correct items of the reduced set are being used to investigate reaction time. Thus, regression analysis on word variables will only utilize the reduced strict scoring set of data.

5.4.1 Interaction of word variables on number of correct responses

This section looks at the interaction between the predictor variables that were found to be significantly correlated with the number of correct responses for each item. These predictor variables include name agreement, age of acquisition, familiarity, imageability, syllable length, phoneme length and cognate status.

Table 5.10 shows a summary of the correlations that exist between the dependent variable, number of correct responses, and the predictor variables. It is also evident, however, that a number of predictor variables are correlated with each other. Syllable length is positively correlated with phoneme length ($r = .871$, $p < .01$); phoneme length is positively correlated with age of acquisition ($r = .351$, $p < .05$); imageability is positively correlated with familiarity ($r = .864$, $p < .01$); imageability is negatively correlated with age of acquisition ($r = -.813$, $p < .01$); imageability is positively correlated with name agreement ($r = .443$, $p < .01$); familiarity is negatively correlated with age of acquisition ($r = -.842$, $p < .01$); familiarity is positively correlated with name

TABLE 5.10

		Syllable Length	Phoneme Length	Imageability	Familiarity	AoA	Name Agreement	Number of Correct Responses
Syllable Length	Pearson Correlation	1.000	.871**	-.144	-.258	.252	-.066	-.401*
	P-value		.000	.389	.118	.128	.693	.013
Phoneme Length	Pearson Correlation	.871**	1.000	-.173	-.225	.351*	-.152	-.426**
	P-value	.000		.300	.174	.031	.362	.008
Imageability	Pearson Correlation	-.144	-.173	1.000	.864**	-.813**	.443**	.665**
	P-value	.389	.300		.000	.000	.005	.000
Familiarity	Pearson Correlation	-.258	-.225	.864**	1.000	-.842**	.474**	.640**
	P-value	.118	.174	.000		.000	.003	.000
Age of Acquisition	Pearson Correlation	.252	.351*	-.813**	-.842**	1.000	-.558**	-.761**
	P-value	.128	.031	.000	.000		.000	.000
Name Agreement	Pearson Correlation	-.066	-.152	.443**	.474**	-.558**	1.000	.529**
	P-value	.693	.362	.005	.003	.000		.001
Number of Correct Responses	Pearson Correlation	-.401*	-.426**	.665**	.640**	-.761**	.529**	1.000
	P-value	.013	.008	.000	.000	.000	.001	
	Sample size	38	38	38	38	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations between number of correct responses and predictor variables

agreement ($r = .474$, $p < .01$); and age of acquisition is negatively correlated with name agreement ($r = -.558$, $p < .01$). This creates multicollinearity difficulties in the regression analysis and interferes with identifying the true effect of each predictor on the dependent variable (Camilleri 2009). Therefore, it is preferable to use a stepwise regression analysis in order to identify the most significant predictor variable or variables.

From the stepwise regression analysis it is evident that age of acquisition is the most important predictor for number of correct responses for each item. This is followed by syllable length (Table 5.11). The predictor variables name agreement, familiarity, imageability and phoneme length, were excluded from the model as they became redundant when a stepwise regression analysis was carried out. The coefficients for age of acquisition (-14.77) indicates that for every extra 1 point on the scale of AoA ratings, the number of correct responses is expected to decrease by 14.77%. The coefficient of determination (R squared) is 0.579, implying that this model explains 57.9% of the total variation of the responses (dependent variable).

TABLE 5.11

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	123.56	7.507		16.460	.000
	AoA_Ratings	-14.77	2.102	-.761	-7.029	.000
2	(Constant)	135.84	9.277		14.643	.000
	AoA_Ratings	-13.68	2.077	-.704	-6.588	.000
	Syllable_Length	-6.66	3.188	-.223	-2.089	.044

Coefficients (number correct – AoA and syllable length)

The above multiple regression analysis included only the significant *quantitative* predictor variables and did not include cognate status. This categorical predictor was also found to be significant during hypothesis testing (section 5.2.6.1). However, it could not have been included in the multiple regression analysis, as the latter only allows regression between quantitative variables. Therefore, an ANCOVA analysis was used as this accommodates both quantitative and categorical variables. Table 5.12 illustrates the results of the ANCOVA analysis.

TABLE 5.12

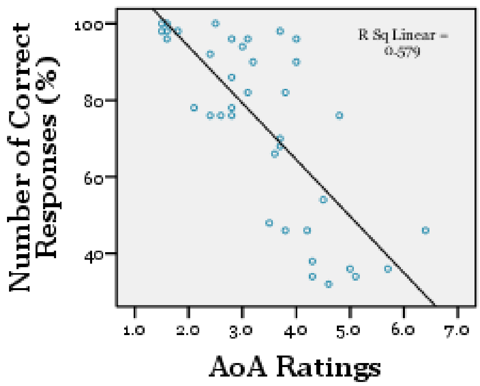
Source	Type III Sum of Squares	df	Mean Square	F	P-value
Syllable Length	597.893	1	597.893	2.685	.112
Phoneme Length	69.428	1	69.428	.312	.581
Imageability	497.011	1	497.011	2.232	.146
Familiarity	320.331	1	320.331	1.439	.240
AoA	1257.195	1	1257.195	5.646	.024
Name Agreement	524.445	1	524.445	2.355	.135
Cognate status	15.597	1	15.597	.070	.793
Error	6680.417	30	222.681		

Tests of Between-Subjects Effects (number correct and dependent variables)

Similar results to multiple regression analysis were obtained through ANCOVA analysis as age of acquisition remained the most significant predictor variable for number of correct responses. This is shown by the corresponding p-value (0.024) which is less than the 0.05 level of significance.

The overlay scatter plot in Figure 5.1 shows the correlation between AoA rating (from a scale of 1 to 7) and the percentage of correct responses.

FIGURE 5.1



Percentage of correct responses with AoA[1]

5.4.2 Interaction of word variables on reaction time

This section analyzes the interaction between the predictor variables that were found to be significantly correlated with mean reaction time for each item. Mean reaction time was found to be correlated with name agreement, age of acquisition, familiarity, imageability, phoneme length and cognate status.

TABLE 5.13

		Phoneme Length	Imageability	Familiarity	Age of Acquisition	Name Agreement
Reaction Time	Pearson Correlation	.326*	-.625**	-.546**	.746**	-.415**
	P-value	.046	.000	.000	.000	.010
	Sample size	38	38	38	38	38

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations between RT and predictor variables

Table 5.13 provides a summary of the correlations that exist between the dependent variable, reaction time, and the predictor variables. It is also necessary to keep in mind that a number of predictor variables were found to be highly correlated with each other (see Table 5.10). Once again a stepwise regression analysis will be used to investigate

[1] AoA ratings scale : 1.0 = 0-2 years; 2.0 = 2-4 years; 3.0 = 4-6 years; 4.0 = 6-8 years; 5.0 = 8-10 years; 6.0 = 10-12 years; 7.0 = 13 + years

the true influence of the predictor variables on the dependent variable, in order to avoid the multicollinearity problems described in the previous section.

A stepwise regression analysis showed that age of acquisition is once again the most important predictor for mean reaction time for each item. The predictor variables name agreement, familiarity, imageability and phoneme length, were excluded from the model when a stepwise regression analysis was carried out (Table 5.14). The coefficients for age of acquisition (294.71) indicates that for every extra 1 point on the scale of AoA ratings, the mean reaction time is expected to increase by 294.71 milliseconds. The coefficient of determination (R squared) is 0.556 implying that this model explains 55.6% of the total variation of the results (dependent variable).

The multiple regression analysis described above (Table 5.14) did not include the significant predictor, cognate status as this is a categorical variable. An ANCOVA analysis was, once again, used to accommodate both quantitative and categorical variables. Table 5.15 illustrates the results of the ANCOVA analysis.

TABLE 5.14

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	800.47	156.707		5.108	.000
	AoA_Ratings	294.71	43.870	.746	6.718	.000

Coefficients (reaction time – AoA)

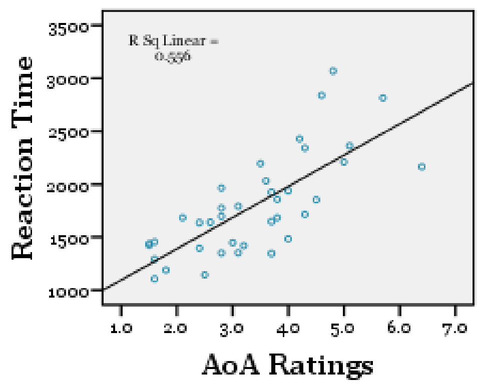
TABLE 5.15

Source	Type III Sum of Squares	df	Mean Square	F	P-value
Phoneme Length	68576.947	1	68576.947	.699	.410
Imageability	305595.944	1	305595.944	3.115	.087
Familiarity	565241.946	1	565241.946	5.761	.023
AoA	1015453.130	1	1015453.130	10.35	.003
Name Agreement	9545.962	1	9545.962	.097	.757
Cognate status	308146.021	1	308146.021	3.141	.086
Error	3041594.585	31	98115.954		

Tests of Between-Subjects Effects (reaction time and dependent variables)

When cognate status was included with the other predictor variables in an ANCOVA analysis, age of acquisition still remained the most significant predictor variable for reaction time. This is shown by the corresponding p-value (0.003) which is the smallest p-value and it is less than the 0.01 level of significance. The overlay scatter plot in Figure 5.2 presents the correlation between AoA rating (from a scale of 1 to 7) and the mean reaction time.

FIGURE 5.2



Reaction time with AoA

5.4.3 Interaction of word variables on number of English responses

Name agreement and age of acquisition were found to be significantly correlated to the number of English responses produced for each item. This section aims to identify the most significant predictor of English responses by means of multiple regression analysis.

TABLE 5.16

		Age of Acquisition	Name Agreement
Number of English Responses	Pearson Correlation	.326*	-.412*
	P-value	.046	.010
	Sample size	38	38

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations between English responses and predictor variables

Table 5.16 summarises the correlations that exist between the dependent variable, number of English responses, and the two significant predictor variables. As shown in Table 5.10 above, these two predictor variables are also highly correlated with each other. Thus, a stepwise regression analysis was carried out in order to avoid multicollinearity problems.

The stepwise regression analysis shows that name agreement is the most significant predictor of English responses as age of acquisition was excluded from the model.

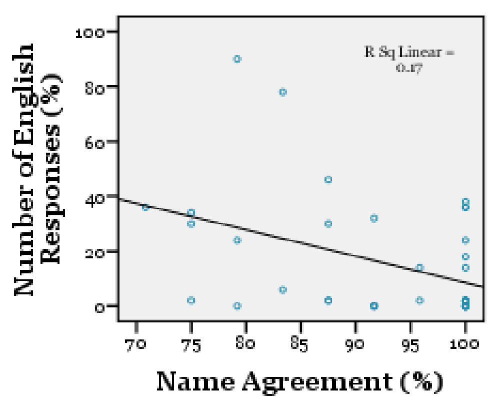
TABLE 5.17

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	104.76	32.902		3.184	.003
	Name Agreement	-.96	.355	-.412	-2.713	.010

Coefficients (English responses – name agreement)

The coefficient of determination (R squared) is 0.170 implying that this model explains only 17% of the total variation of the results (dependent variable). The coefficients for Name Agreement (-.96) indicates that for every extra 1% in name agreement, the number of English responses tends to decrease by 0.96%. The overlay scatter plot in Figure 5.3 presents the correlation between the percentage of name agreement and the percentage of English responses produced.

FIGURE 5.3



Number of English responses produced with name agreement in percentages

5.5 SUMMARY OF FINDINGS ON WORD VARIABLES

The first part of the chapter presented an investigation of the correlations that exist between 3 dependent variables and their corresponding predictor variables. The second part of the chapter then focused on regression analysis with the aim of identifying the most significant predictors. Age of acquisition has been shown to be extremely important in influencing both the number of correct responses for each item and also the mean reaction time for each item. Thus, it is apparent that the earlier a word is learned in life, the quicker it is named and the more accurate it will be. In fact, it has been shown that for every extra point on the AoA ratings scale (i.e. older age of acquisition), the number of correct responses is expected to decrease by 14.77% and the mean reaction time is expected to increase by 294.71 milliseconds.

Finally, although both AoA and name agreement were shown to be significantly correlated with the number of English responses produced for each item, name agreement emerged as the most important predictor. In fact, for every 1% more of name agreement, the English responses are expected to decrease by 0.96%.

5.6 INTERACTION OF SIGNIFICANT DEMOGRAPHIC AND WORD VARIABLES

This section aims to analyse the interaction between the variables that were found to be predominantly significant in the regression analysis carried out in Chapters 4 and 5. It is necessary to recall that in the chapter on demographic variables (Chapter 4), regression analysis using the complete set (and strict scoring) of data showed that education was the most significant predictor variable for score (as dependent variable). When regression analysis was carried out using the reduced set (and strict scoring) of data, the same result emerged, i.e. that years of education emerged as the most significant predictor variable for score. Age remained significant and was not excluded from the model, however (section 4.7.1 in Chapter 4). On the other hand, age emerged as the most significant predictor of reaction time, and years of education was not excluded from the regression analysis model (section 4.7.2 in Chapter 4).

Regression analysis on word variables showed that age of acquisition is the most significant influencing variable of both the number of correct responses for each item (section 5.4.1) and the mean reaction time for each item (section 5.4.2). Therefore, it is necessary to identify the possible interactions between the results of the two different types of variables. The sections below report the analysis of the relationship that exists between demographic and word variables in relation to score (sections 5.6.1 – 5.6.4).

Score on early AoA and score on late AoA words will be correlated with the demographic predictors, age and education (5.6.1). Multiple regression analysis will then be carried out to identify the most important predictor of naming accuracy on words which have early AoA and late AoA (section 5.6.2).

Syllable length also remained significant when multiple regression analysis was performed (see section 5.4.1). Therefore, it is also necessary to investigate whether there is an interaction between word variable, syllable length and demographic variables, age and education. The same method described above for early and late AoA words will be applied to short and long syllable length (sections 5.6.3 and 5.6.4).

Reaction time on early AoA words and reaction time on late AoA words will also be correlated with demographic predictors, age and education (section 5.6.5) and regression analysis will follow to identify the most important predictor of reaction time on early AoA words and reaction time on late AoA words (section 5.6.6).

5.6.1 Correlations between AoA and predictors education and age

The reduced set of the BNT includes 19 words of early age of acquisition and 19 words of late age of acquisition. Simple linear correlations were performed in order to interact word variable AoA with demographic variables education and age. The mean score achieved by the respondents on words with early AoA and the mean score achieved by the respondents on words with late AoA will be used to identify any AoA effects on naming performance of respondents of varying ages and different years of education.

TABLE 5.18

		Score on Early AoA items	**Score on Late AoA items**
Years of Education	Pearson Correlation	.297*	.707**
	P – value	.021	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Simple linear correlations between score on early vs late AoA items and years of education

Table 5.18 presents the correlations between years of education and score on early AoA words versus score on late AoA words. Although both sets of scores were found to be positively correlated with years of education, a strong positive correlation was only evident with the score on late AoA words. This demonstrates that AoA effects were present, and that on the late AoA words the respondents with higher levels of education scored significantly higher than their less educated counterparts ($r = .707$, $p < .01$). Only a weak correlation was found when the score on early AoA words was used ($r = .297$, $p < .05$). Similar correlations were carried out with age as predictor variable (Table 5.19).

TABLE 5.19

		Score on Early AoA items	**Score on Late AoA items**
Age	Pearson Correlation	-.444**	-.580**
	P – value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between score on early vs late AoA items and age

Both sets of scores were found to be negatively correlated with age. A moderately strong negative correlation emerged between age and the scores on both late ($r = -.580$, $p < .01$) and early ($r = -.444$, $p < .01$) AoA words. These rather similar results associated with respondent age demonstrate that AoA effects were not present in this case. This implies that both early and late AoA words are similarly correlated with age.

5.6.2 Regression analysis using score on words of early and late AoA

Regression analysis was then performed to determine the most significant predictor of score on early AoA items and the most significant predictor of score on late AoA items. Table 5.20 presents the Coefficients table of the multiple regression analysis with score on early AoA items as the dependent variable. It is evident that age is the most significant predictor of naming accuracy on a list of words of early AoA. Predictor variable years of education was excluded from the model. On the other hand, when late AoA words were used in a regression analysis that included age and years of education as predictor variables, years of education emerged as the most significant predictor. Age was not excluded from the model, however (Table 5.21). Thus, on a list of words of late AoA, years of education is expected to be the most significant predictor of naming accuracy.

TABLE 5.20

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	19.663	.688		28.589	.000
	Age	-.046	.012	-.444	-3.770	.000

Coefficients (early AoA words – age and education)

TABLE 5.21

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	1.691	1.373		1.232	.223
	Years of Education	.782	.103	.707	7.623	.000
2	(Constant)	7.114	2.495		2.852	.006
	Years of Education	.616	.118	.557	5.241	.000
	Age	-.063	.025	-.272	-2.556	.013

Coefficients (late AoA words – age and education)

5.6.3 Correlations between syllable length and predictors education and age on score

When analysing the relationship between word variables and number of correct responses for each item by means of regression analysis, AoA was found to be the most significant variable. The variable syllable length was not excluded from the regression model, however (section 5.4.1). Therefore, it is also necessary to correlate the independent variables age and education with the score on words with short versus long syllable length.

Simple linear correlations with years of education (as predictor variable) and score on words of varying length[2] (as dependent variable) were also carried out (Table 5.22). Comparable correlations were found between years of education and score on words of

[2] Short words consist of 1 - 2 syllables and long words consist of 3 - 4 syllables.

varying length. Both lengths of words showed a strong positive correlation between score and years of education.

TABLE 5.22

		Score on words with short syllable length	Score on words with long syllable length
Years of Education	Pearson Correlation	.570**	.651**
	P – value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between score on words with short & long syllable length and education

Simple linear correlations were also carried out between age and score on words of short and long syllable length (Table 5.23).

TABLE 5.23

		Score on words with short syllable length	Score on words with long syllable length
Age	Pearson Correlation	-.485**	-.639**
	P – value	.000	.000
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between score on words with short & long syllable length and age

Thus, although both sets of scores were found to be negatively correlated with age, a stronger negative correlation was evident with the score on longer words ($r = -.639$, $p < .01$) than with the score on shorter words ($r = -.485$, $p < .01$). This demonstrates that naming accuracy decreases with increasing age for both short and long syllable words. However, as we age, longer words tend to become slightly more difficult to name than shorter words.

5.6.4 Regression analysis using score on words of short and long syllable length

Regression analysis was performed to determine the most significant predictor of score on words of short syllable length and of score on words with long syllable length. Table 5.24 presents the Coefficients table of the multiple regression analysis with score on short syllable length as the dependent variable. It is evident that years of education is the most significant predictor of naming accuracy on a list of short syllable words. The predictor variable age was excluded from the model.

TABLE 5.24

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	60.337	4.643		12.996	.000
	Years of Education	1.832	.347	.570	5.281	.000

**. Correlation is significant at the 0.01 level (2-tailed).

Coefficients (short syllable length – age and education)

On the other hand, the Coefficients table of the multiple regression analysis with score on long syllable length as the dependent variable is presented below. Once again years of education emerged as the most significant variable. However, age was not excluded from the model and was retained as significant (Table 5.25).

In both sets of regression analysis years of education emerged as the most significant predictor of score on words with short and long syllable length. Age remained significant only in the analysis involving longer syllable words. Thus, this result confirms that ageing has a more detrimental effect on words of longer syllable length than on words with shorter syllable length.

TABLE 5.25

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	18.447	7.206		2.560	.013
	Years of Education	3.516	.539	.651	6.528	.000
2	(Constant)	57.716	12.401		4.654	.000
	Years of Education	2.314	.584	.428	3.960	.000
	Age	-.457	.123	-.403	-3.722	.000

**. Correlation is significant at the 0.01 level (2-tailed).

Coefficients (long syllable length – age and education)

5.6.5 Effect of word and demographic variables on reaction time

Age was found to be the most important demographic predictor of RT, while years of education remained significant. AoA emerged as the most influencing predictor of RT.

Thus, simple linear correlations were performed in order to interact word variable AoA with demographic variables age and education. The mean RT achieved by the respondents on words with early AoA and the mean RT achieved by the respondents on words with late AoA will be used to identify any AoA effects on naming latency of respondents of varying ages and different years of education. Table 5.26 presents the correlations between age and RT on early AoA words versus RT on late AoA words.

TABLE 5.26

		RT on Early AoA items	**RT on Late AoA items**
Age	Pearson Correlation	.562**	.291*
	P – value	.000	.024
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Simple linear correlations between RT on early vs late AoA items and age

Although both sets of reaction times were found to be positively correlated with age, a strong positive correlation was only evident with the RT on early AoA words. This demonstrates that AoA effects were present, and that on the early AoA words the respondents' mean RT tended to increase more significantly with age ($r = .562$, $p < .01$). A weaker correlation was found when the RT on late AoA words was used ($r = .291$, $p < .05$), implying that RT on late AoA words was affected less with ageing. Similar correlations were carried out with years of education as the predictor variable (Table 5.27).

TABLE 5.27

		RT on Early AoA items	**RT on Late AoA items**
Years of education	Pearson Correlation	-.483**	-.422**
	P – value	.000	.001
	Sample size	60	60

**. Correlation is significant at the 0.01 level (2-tailed).

Simple linear correlations between RT on early vs late AoA items and years of education

Both sets of reaction times were found to be negatively correlated with years of education. A moderate negative correlation emerged between education and the RTs on both late ($r = -.422$, $p < .01$) and early ($r = -.483$, $p < .01$) AoA words. These rather similar results associated with the respondents' level of education demonstrate that AoA effects were not present in this case. This implies that both early and late AoA words are similarly correlated with education when reaction times are compared.

5.6.6 Regression analysis using RT of words of early and late AoA

Regression analysis was then performed to determine the most significant predictor of reaction time on early AoA items, and the most significant predictor of reaction time on late AoA items. Table 5.28 presents the Coefficients table of the multiple regression analysis with RT on early AoA items as the dependent variable. It is evident that age is the most significant predictor of naming latency on a list of words of early AoA. Predictor variable years of education was excluded from the model.

TABLE 5.28

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	881.276	122.677		7.184	.000
	Age	11.194	2.166	.562	5.168	.000

Coefficients (early AoA words - age and education)

On the other hand, when late AoA words were used in a regression analysis that included age and years of education as predictor variables, years of education emerged as the most significant predictor of reaction time. Age was excluded from the model (see Table 5.29). This implies that, on a list of words of late AoA, years of education is expected to be the most significant predictor of naming latency.

TABLE 5.29

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	2726.353	197.038		13.837	.000
	Years of Education	-52.195	14.727	-.422	-3.544	.001

Coefficients (late AoA words - age and education)

5.7 CONCLUSION

Regression analysis that combined both demographic variables and word variables produced interesting results regarding the impact of age and education on naming accuracy and naming latency. Age was found to be the most significant predictor of both naming accuracy and naming latency when using words of early AoA. On the other hand, years of education emerged as the most important predictor of both naming accuracy and naming latency when using words of late AoA. When using words of both short and long syllable length, years of education was also the most significant predictor of naming accuracy. This confirms that the material that is used in clinical and research settings must be taken into consideration when attempting to interpret results. The effect of different demographic variables (such as age and education) on naming performance relies heavily on the stimuli being used. Naming words that are acquired early produced different results to naming words that are acquired later in life. The following chapter (Chapter 6) aims to discuss the findings of Chapter 4 and Chapter 5.

Chapter 6

Discussion 1

People without aphasia

6.0 INTRODUCTION

The first part of this research focused on picture naming performance in Maltese-speaking individuals. Because of its widespread use, the Boston naming test (BNT, Kaplan et al 1983) was utilised in order to study the effect of demographic and word variables on naming accuracy and naming latency. While simple linear correlations and one-way ANOVAs were used to analyse differences and relationships between naming performance and predictor variables, regression analysis then identified the most significant predictors. The influence of ageing and education on naming ability was given special attention in relation to naming accuracy, naming latency and the number of English responses produced.

When using a neuropsychological test on populations other than it was originally normed on, it is essential to interpret results according to normative data obtained from those particular populations. Accurate and reliable interpretation of results is only achieved if the neuropsychological test is standardised on that particular population. Thus, this study also aims to present normative data for Maltese speakers for a modified version of the BNT. Since age and education were identified as significant predictors of naming performance, the normative data (Table 6.1) is stratified for age and education.

Chapter 3 provided a detailed explanation of the methodology used to collect and analyse the available data, while Chapters 4 and 5 outlined the results obtained with regards to demographic and word variables respectively. The following sections will discuss the findings related to particular areas (scores, effect of demographic variables, correct responses for each item and effect of word variables) and the implications of the results with regards to naming performance and ageing and bilingual contexts.

6.1 SCORES

Scores were compared in a number of ways during the analysis. The scores of the complete set were compared to those obtained from the reduced set, the strict scores were compared to the lenient scores, and the scores of Sample A and those of the main study were also compared.

6.1.1 Complete and reduced set scores

The mean percentage score of all participants was significantly lower when using the complete set than when using the reduced set. This means that when the low name agreement items were removed from the test, naming performance improved. Using the reduced set of the test with Maltese speakers ensures that individuals are tested on those items (i) for which there is a general agreement regarding an appropriate name, (ii) which have a fewer number of possible responses and (iii) whose names are likely to be part of the person's vocabulary.

When naming low name agreement items individuals are required to choose the most appropriate name from a selection of available words which they might have for the same item. This increases the probability of competition or conflict associated with lexical retrieval when more than one possible response is available. Thus, if we remove items which have low name agreement it is possible to decrease this conflict during naming. Bilinguals already experience some conflict in naming when having to select and activate a particular word from the two available languages, and low name agreement items would add further competition. Other studies on different populations have also concluded that some items may need to be removed, or replaced, from the test for linguistic and cultural reasons (Patricacou et al 2007; Cruice et al 2000; Kim and Na 1999).

Shorter versions of the BNT have also been investigated by different researchers. They have been found to be valid assessment tools, which are especially useful when used with particular populations, such as ageing individuals and people with head injuries who may have attention problems (e.g. Elkadi et al 2006). Thus, the reduced set of the Maltese BNT may also be useful for such individuals.

This study has also shown that the results of the reduced BNT set were correlated to the complete set, and similar education and age effects were evident on both sets. Thus, using the reduced set of the BNT on Maltese individuals should provide the same information as the complete set, with the added benefit of catering for people with lower attention abilities and of being less time consuming to administer. Furthermore, the reduced set of the BNT (as suggested here) should provide a more reliable reflection of the true lexical retrieval abilities of Maltese speakers since it includes only those items which are linguistically and culturally appropriate.

6.1.2 Strict and lenient scoring methods

Alternative responses for each item were not target responses since they were not the dominant responses. However, all the acceptable responses were scored as correct in lenient scoring. Although strict and lenient scoring methods were found to be correlated with each other, the lenient measures unsurprisingly produced higher mean scores. The reduced set of the BNT consists of high name agreement items only. However, most of the items still had more than one possible response, since for all items both English and Maltese responses were accepted as possible responses. In fact, only 6 of the 38 items had only one possible response (e.g. 'ħelikopter' for *helicopter*, 'iglù' for *igloo*) as the English and Maltese terms were the same response. A further 5 items simply had more than one possible Maltese response (e.g. 'moxt' and 'pettne' for *comb*, 'mikinsa' and 'xkupa' for *broom*). Since the BNT was initially developed on a population that differs linguistically and culturally from the Maltese population, particular items had alternative Maltese/English responses which could not be scored as inaccurate. Thus, when administering the BNT on Maltese speaking individuals lenient scoring allows the scorer to obtain a better indicator of the true naming abilities. Other researchers have also proposed and adopted this method of scoring. Worral et al (1995) and Cruice et al (2000) suggest that synonym responses should be accepted as correct responses, especially when the test being used is not originally developed on that particular population. In the case of bilingual individuals, Kohnert et al (1998) also recommend the use of an alternative (either English or Maltese words) scoring system as it provides a more accurate reflection of bilingual individuals' naming abilities. Thus, using lenient scoring with the reduced set of the BNT appears to be the best measure for assessing naming in Maltese adults.

Table 6.1 presents normative data for the reduced set of the BNT stratified by age and education. The reduced set and the particular age and education groups are being suggested following the findings made up to now. The less educated group (0 – 12 years) consists of individuals who received minimal or only obligatory education, while the more educated group refers to individuals who have furthered their studies and attended post secondary diplomas and degrees at college or university level. Since naming performance has been shown to experience a decline after 66 years of age, normative data is stratified accordingly (younger than 66 years and 66+ years).

TABLE 6.1

NORMS FOR STRICT SCORING						
	AGE GROUP 1 20 – 65 YEARS			AGE GROUP 2 66 + YEARS		
	M	SD	CI	M	SD	CI[1]
0 – 12 YEARS OF EDUCATION	28.6	3.1	27.2 – 30.0	23.3	4.5	22.1 – 24.5
13 OR MORE YEARS OF EDUCATION	33.2	3.0	32.2 – 34.2	27.9	9.9	26.2 – 29.6

NORMS FOR LENIENT SCORING						
	AGE GROUP 1 20 – 65 YEARS			AGE GROUP 2 66 + YEARS		
	M	SD	CI	M	SD	CI
0 – 12 YEARS OF EDUCATION	29.1	3.1	27.7 – 30.4	23.9	4.4	22.8 – 25.1
13 OR MORE YEARS OF EDUCATION	34.2	3.1	33.2 – 35.2	28.5	9.3	26.8 – 30.2

Normative data for the reduced set of the BNT

[1] 68% confidence interval (1 SD)

6.1.3 Sample A and main study scores

The scores obtained from the main study were significantly lower than those obtained from the participants of Sample A. An effect of the different methods of data collection (timed versus untimed picture naming) on scores was initially suspected. However, it was shown that, at least in the age group of 30 – 45 years, timed picture naming did not negatively affect the scores (see section 4.3.2 in Chapter 4). Unfortunately, it is not possible to analyse the effect of time pressure on older individuals, as Sample A did not include participants over 65 years of age. The significant difference in scores is, therefore, possibly a result of the different samples used, which involved a different number of participants, as well as different age groups and levels of education, rather than the different methods of data collection. The effects of age and education will be discussed in the sections below.

6.2 DEMOGRAPHIC VARIABLES

The influence of demographic variables on naming performance was investigated, and the findings were reported in Chapter 4. The influence of education, ageing, gender and language on naming performance is discussed below.

6.2.1 Level of education

In the present study, education emerged as the most significant predictor of naming accuracy. However, ageing was not excluded from the regression model and it also remained significant. While education has also been identified as a very important predictor of naming ability by a number of studies (e.g. Neils et al 1995; Allegri et al 1997; Tombaugh and Hubley 1997; Mansur et al 2006) other research failed to find an effect of education (e.g. Tsang and Lee (2003), Nicholas et al (1985), Van Gorp et al (1986), Kaplan et al (1983) and Farmer (1990). It has been argued that the latter may be a result of a sample bias towards more highly educated individuals, with samples including participants with higher levels of education. The present study intentionally aimed at acquiring a non biased sample in terms of education (2 categories), age (6 age groups) and gender, such that each variable group consisted of a similar number of individuals. An effort was made to have an approximately equal number of respondents in each age group (9-11 respondents in each group), and to have a similar number of people with different levels of education (29 respondents with lower education and 31 respondents with higher education). However, the older age groups consisted of more individuals who had very few years of education. In fact, five of the older participants had 6 or fewer years of education[2]. It has been mentioned before, that this is a reflection of the sociological reality of older persons in Malta, meaning that the effect of age on naming performance may really be a result of a cohort effect.

6.2.1.1 Education as strongest predictor

Education was found to be the strongest predictor of naming accuracy in both complete and reduced sets of the BNT. Since the test was originally developed on a culturally and linguistically diverse population and included items which may not have been culturally appropriate for the Maltese it had been suspected that the higher educated individuals would be at an advantage over their less educated counterparts as a result of more exposure to a wider vocabulary. However, the education effect remained strong even when the low name agreement items were removed (in the reduced set). Thus, this confirms that education is the most important variable with regards to BNT score or naming accuracy, even when the 'more' difficult items are removed from the test.

6.2.1.2 Lower mean scores

Lower mean scores were found in the present study in comparison to the scores of studies on other populations. For example, the mean score obtained from the main study complete set when using lenient scoring was 39.2 (out of 60). This score is significantly lower than the mean scores obtained from other studies (such as Van Gorp et al (1986), Nicholas et al (1989) and Worrall et al (1995) who obtained mean scores of 54.0, 54.5 and 52.8 respectively). One possible reason for the lower mean scores achieved by the Maltese sample could be the fact that the sample included individuals whose education experience ranged from 0 to 19 years and, therefore, included a number of individuals with very low levels of education. Patricacou et al (2007) argued that the lower mean scores found in their study on a Greek sample, were mainly evident

[2] One respondent (aged 76) had received no formal education at all, one respondent (aged 82) had 3 years of education, two respondents (aged 68 and 83) had 5 years of education, and one respondent (aged 79) had 6 years of education.

in those participants who were less educated individuals, as higher educated people produced a score of over 50. However, low education alone cannot account for the lower mean scores obtained in the present study as both categories of education scored well below 50 even in lenient scoring. The individuals with lower education had a mean lenient score of 33 and those with higher levels of education had a mean lenient score of 43 on the complete set of the BNT. Thus, other variables also influenced the lower scores obtained. An interaction between ageing and education was evident as both ageing and level of education contributed to the lower scores achieved. The highest mean score was obtained by the younger, more educated group (M = 45.04; SD = 6.34). Table 6.2 presents the mean lenient scores (of the complete set) for each age and education group, showing that both the younger and older age groups with higher levels of education still achieved a mean lenient score that was lower than 50. The mean score obtained by Sample A is also presented in Table 6.2. The participants of Sample A had a younger mean age and all had 13 or more years of education. Reaction times were not measured in this sample. Yet, even the mean score of Sample A is lower than 50. The possible influence of bilingualism in explaining the lower mean scores obtained from the Maltese sample is discussed in Section 6.2.4.2.

TABLE 6.2

Main Study[3]				
	Age group 20 – 65 years		**Age group 66 – 85 years**	
Level of education	**Mean Score**	**SD**	**Mean**	**SD**
0 – 12 years of education	36.17	5.04	30.12	7.83
13 or more years of education	45.04	6.34	32.40	14.89

Sample A[4]		
	Mean Score	**SD**
	48.6	6.10

Mean scores using lenient scoring in the complete set

6.2.1.3 Education and Age of Acquisition

The latter part of Chapter 5 investigated the interaction between the most significant predictor demographic variables and word variables. It was evident that there is an

[3] Main Study sample: N = 60, age range = 20 - 85 years, education = 0 - 19 years.
[4] Sample A sample: N = 24, age range = 20 - 60 years, education = 11 - 20 years.

interesting interaction between education and age of acquisition (AoA). Tsang and Lee (2003) used only early AoA items in their study and found no effect of education. In the present study naming accuracy on late AoA items was strikingly more affected by education than naming accuracy on early AoA words. This implies that individuals with higher education levels scored higher on late AoA than the less educated individuals. This finding replicates previously reported findings (e.g. Carroll and White 1973; Ellis and Morrison 1998) that state that people who are more educated coped much better with late AoA words.

6.2.1.4 Education and reaction time

A moderate negative correlation was found between reaction time (RT) and the number of years of education. This correlation was evident in both the reduced and the complete set. However, the difference between the mean RTs of the two groups of education (0-12 years and 13 or more years) was not found to be significant when ANOVA analysis was performed. Thus, RT cannot be conclusively said to be influenced by education. Ageing has been reported to slow down all processes (Section 2.5 in Chapter 2), and this effect of age is apparent in naming latency too (Section 6.2.2.3).

6.2.1.5 Education and English responses

The relationship between education and the number of English responses produced was also investigated in Chapter 4. The bilingual context of the Maltese islands results in using both Maltese and English even during the BNT. A strong positive correlation with education was evident. The further the level of education achieved the greater the chance of producing English responses. As individuals with higher levels of education tend to be exposed to and use more English on a daily basis, it is possible that English may be more highly activated than it would be in someone with less education. Thus, this may explain the tendency of individuals with higher education to produce more English words when naming pictures, as retrieval of the more highly activated word (English version in this case) may occur. This further justifies scoring both English and Maltese responses as correct.

It is necessary to consider the implications of the findings related to education and its impact on naming accuracy for clinical purposes. Normative data must be stratified according to education categories (and age) as people of different levels of education will undoubtedly score differently. Individuals of lower education may produce a low score while still being within *normal* limits according to their level of education. The interaction of education with demographic variable age and word variable AoA must also be taken into account when interpreting BNT test results of Maltese speakers.

6.2.2 Ageing

6.2.2.1 Significant effect of ageing

Although ageing did not emerge as the most significant predictor of naming accuracy, an effect of age is clearly evident in the present study, with a particular decline over 66 years of age. The results of this study replicate the findings of previous research (e.g. Van Gorp et al 1986; Belke and Meyer 2007; Figueredo Balthazar et al 2008) on naming and ageing. The complete set of the BNT showed a moderate negative correlation between naming accuracy and age, while the reduced set showed a strong negative correlation. ANOVA analysis produced evidence of a drastic decline in naming accuracy after 66 years of age. Post hoc investigations showed that significant differences between the eldest age groups (66 - 75 years and 76 – 85 years) and the younger age groups were particularly evident when using the reduced set of the BNT (see Table 4.13, in Chapter 4) and especially when using lenient scoring. Age group 6 (76 – 85 years) was found to have significantly different mean scores from age groups 1,2, 3 and 4 (ranging from 20 – 65 years) when using the reduced set and lenient scoring. Also, the mean scores of age group 5 (66 – 75 years) were significantly different from the two younger age groups (20 – 35 years and 36 – 45 years). When using the complete set, however, the mean scores of the older age groups were only found to be significantly different from the two younger age groups. Therefore, using the reduced set provides better evidence of an ageing effect on naming accuracy and, particularly, a decline from 66 years of age. Furthermore, syllable length was shown to affect naming accuracy as, with ageing, longer words tend to become slightly more difficult to name than shorter words (see Chapter 5). The effect of ageing on retrieving longer syllable words may be a result of an increasing difficulty at the phonological assembly level or at the level of articulation programming and execution (see section 2.2.1, in Chapter 2).

6.2.2.2 Ageing and cohort effects

Another possible explanation for the decrease in naming accuracy that occurs with ageing may be the presence of cohort effects. Cohort effects may be present in this study since it is a cross-sectional study rather than a longitudinal one. The older individuals who took part in this study undoubtedly had different experiences from the younger individuals. As a result of the socioeconomic changes that occurred in Malta over the past decades, the ageing individuals of today had different experiences from the individuals who will be ageing in 15-20 years' time from now. Fewer opportunities for education and occupations outside the rural areas and outside the home were experienced by the ageing individuals of today. The latter was predominantly experienced by females rather than males. Thus, a longitudinal study over a period of years could lead to clearer conclusions regarding the impact of ageing on naming performance. Cruice et al (2000) carried out a 4-year longitudinal study and found only a weak effect of ageing. However, the authors stress that a longer study should provide more accurate information about naming and ageing.

6.2.2.3 Ageing and reaction time

Age and reaction time were found to be positively correlated such that RT increased with increasing age. Post hoc comparisons using both complete and reduced sets showed that the mean RT of the eldest age group (76 – 85 years) was significantly different from age groups 1, 2, 3 and 4 (ranging from 20 – 65 years), but not from age group 5 (66 – 75 years). Thus, it is evident that, after 66 years, RT is significantly slower than that of younger individuals. A similar decline in naming performance was reported with regards to the effect of ageing on naming accuracy (see section 6.2.2.1). Thus, after 66 years of age both accuracy and latency showed a decline. Ageing emerged as the strongest predictor of RT. These results are compatible with those reported by other researchers who report a slowing down of cognitive processes with ageing (Tsang and Lee 2003; Madden et al 2005). Thus, the finding of slower RTs with increasing age may not simply be related to slower naming performance. Rather, this may be a reflection of the general slowing of psychomotor tasks associated with ageing.

The significant decline in naming accuracy, and the significantly slower RTs, after 66 years may be attributed to the physical changes that have been reported in brain volume and neural activity in the prefrontal and frontal regions of the left hemisphere (Tsang and Lee 2003; Wierenga et al 2008) rather than simply resulting from cohort effects. Yet, a longitudinal study (of more than 4 years) and a larger sample might produce different explanations for this general decline in naming performance over 66 years of age. Future studies must also collect personal data regarding common conditions associated with ageing from potential participants, since even hypertension and diabetes mellitus have been associated with cognitive and lexical retrieval difficulties.

6.2.2.4 Ageing and English responses

The relationship between ageing and the number of English responses produced during administration of the BNT was also investigated during this study. The findings showed that the number of English responses produced decreased with increasing age. In fact, post hoc comparisons showed that the number of English responses produced by age group 6 (76 – 85 years) was significantly different from that of age groups 1 and 2 (20 – 35 years and 36 – 45 years). Age group 5 (66 – 75 years) was also found to be significantly different from age group 1 (20 – 35 years) with regards to the number of English responses used. It has been previously reported (Sciriha 1997) that younger people, especially those with higher levels of education, in Malta tend to use more English than older people. Thus, it is possible that the socioeconomic changes that have occurred may be influencing factors in the use of English versus Maltese. The introduction of cable television over the past years may be one contributing factor to the increase of exposure and use of English. Previously, most people watched Italian TV apart from the local national TV station and exposure to English speaking TV stations was minimal. Furthermore, a substantial number of new private schools have been opened and these use English as their main medium of instruction. In 30 years' time a different picture may be seen as the young people of today, who use mainly English to communicate, or else use both English and Maltese, may use more English than individuals who are ageing today. Thus, there is a possibility of cohort effects in this situation too.

6.2.2.5 Ageing and variation in performance

This study also showed evidence of a greater variation in naming accuracy after 66 years of age. Ageing and variability in performance have been attributed to the variation in cognitive abilities of ageing individuals (Lindeboom and Weinstein 2004). Other reasons for this variation may lie in the cultural and education experiences of each individual (Zec et al 2007a) as well as other psychogenic factors such as leading cognitively and physically active lifestyles.

6.2.3 Gender

Gender was found to be related to score only in the complete set of data and it was not correlated to RT and the number of English responses in any of the sets used. No interaction between age and gender was identified in relation to score either. Thus, once the low name agreement items were removed from the set of pictures, females performed just as well as males in picture naming, implying that there is no gender bias in the reduced BNT.

6.2.4 Bilingualism

All the respondents of the main study were mainly Maltese speakers and all were exposed to English to some degree. The extent of exposure and use of English largely depended on one's level of education and type of occupation. However, English is encountered continuously especially through the media (television programmes, adverts, announcements, written text in one's environment and so on). Thus, even those participants who claimed not to use any English on a daily basis were exposed to some English too.

During the administration of the BNT, the participants and respondents of the study were requested to say the first name that came to mind when a picture was shown (see section 3.5.5.1). They were intentionally not told to use English or Maltese only during the administration of the BNT and no emphasis on choice of language was made. This was necessary in order to allow the respondents to perform in a manner that was natural for each person involved and to avoid the added pressure of having to block activation of the lexical representations of one of their languages.

6.2.4.1 Bilingualism and English responses

As expected, the participants and respondents who took part in the study used a number of English responses. 19.4% of the total responses were produced in English in the complete set while 16.9% of the responses were produced in English in the reduced set. Considering the bilingual context of the Maltese islands and the long history of British rule it was expected that the responses would include English words. Further, since the BNT was initially developed on an English–speaking population some items are not culturally associated with the Maltese. Thus, words such as *beaver*, *pretzel* and *hammock* would have been learned in English only (through television, reading, etc) and a Maltese translation equivalent may not even be known. Other item names are simply used more in English than in Maltese (e.g. *toothbrush*, *acorn* and *wheelchair*) and have English words as target responses. When the number of English responses was investigated in Chapter 5, it became evident that name agreement was the most significant predictor of the number of English responses produced. It became clear that the lower the name agreement, the greater the tendency to use an English name rather than a Maltese one. A possible explanation for this might be that in the case of uncertainty about the Maltese name of an object, the respondents tended to retrieve the English equivalent, as they were possibly more certain about the accuracy of the English name. This highlights the importance of giving credit to English responses even when these are not the target names for items.

6.2.4.2 Bilingualism and lower mean scores

Bilingualism has been shown to have a negative effect on spoken word production with slower naming, less accuracy and a higher tendency for tip of the tongue states (Section 2.6.4 in Chapter 2). Reports of mean scores of bilingual studies on the BNT are fewer than those on monolingual populations. Kohnert et al (1998) reported the BNT mean scores obtained from their study on Spanish-English bilinguals. A mean score of 32.0 was obtained when the participants were told to name the BNT in Spanish only, and a mean score of 46.7 when told to name items in English only. It is apparent that the lower score was obtained when the first, non dominant language was used, while higher scores were obtained when English was used. The latter study differed from the present one in several ways. First and foremost, it involved a different experience of bilingual language acquisition. A switched dominance is often experienced by Spanish-English bilinguals, since Spanish is the first language acquired and English (which is acquired from school age) generally becomes the dominant language. On the other hand, for the respondents of this study, Maltese is both the first *and* the dominant language used (National Statistics Office 2006), even if a person is highly proficient in English too (see Section 2.6.2 in Chapter 2). There are also differences in the samples used (including sample size, age group and level of education), since Kohnert and colleagues used a sample of 100 participants, with a mean age of 20.82 years and 12 or more years of education. It is interesting to note that the younger and more educated Maltese respondents (age range of 20 – 65 yrs with 13 or more years of education) obtained a mean score of 45.04 (Table 6.1), which is similar to the score (46.66) obtained by the young, educated sample of Kohnert et al's study when the dominant language was used. Therefore, when matched for age and education these two bilingual groups performed similarly.

Roberts et al (2002) also reported a lower mean score of bilinguals in comparison to a monolingual control group. Their study involved a monolingual (English) group, a Spanish-English bilingual group and a French-English bilingual group. Once again the sample involved younger and more educated participants than the present study[5]. All bilingual participants were said to be proficient in both languages and were all asked to name the BNT items in English. Lenient mean scores were reported for the monolingual group (53.9), the Spanish-English group (43.9) and the French-English group (41.4). Once again, despite the different sample and methodological characteristics across studies, the younger, more educated Maltese category of participants scored (45.04) similarly to the bilingual groups studied by Roberts et al (2002). Thus, in addition to the influence of education and age, bilingualism may also be regarded as a significant issue in explaining the lower scores obtained. This further highlights the need for interpreting BNT scores according to the specific normative data for each population.

[5] The samples of Roberts et al (2002) : English unilinguals (mean age 37.1 yrs; mean years of education 16.29 yrs); Spanish bilinguals (mean age 39.6 yrs; mean years of education 17.91 yrs); French bilinguals (mean age 34.9 yrs; mean years of education 15.0 yrs);

6.2.4.3 Bilingualism and reaction time

It has been reported that bilingual individuals are slower (Gollan and Acenas 2004; Bialystok 2008a) in naming than monolinguals, even when they are tested in their dominant language (Gollan et al 2007). The CRL – IPNP[6] database, which is available online[7], was used to compare the RT data of the Maltese bilingual sample to a monolingual sample. This database includes data from seven different college-aged populations from different countries (American English, German, Mexican, Italian, Bulgarian, Hungarian and Mandarin Chinese). From an email correspondence to Anna Szekely (May 2010), it was confirmed that the studies were based on monolingual participants. However, the researcher could not exclude the possibility that some participants (e.g. the American English sample from San Diego might have been exposed to Spanish as well as English, and the Hungarian students may have been exposed to English too). On the other hand, it was assumed that the Italian participants (of the CRL – IPNP) were monolingual. Thus, the RT data of the Italian students was compared to the RT data of a segment of the Maltese (bilingual) sample. The Maltese participants selected for comparison were matched for age and education to the Italian students. Considering the close geographical proximity and the cultural similarities between the Italian and Maltese populations it was assumed that the main difference between the two samples was the element of bilingualism in the latter population. Differences in word variables (such as word length and complexity) between the two groups were not accounted for.

Unfortunately the CRL database does not include all the items of the BNT. Comparison between the monolingual Italian group and the bilingual Maltese group was based on the 36 items which were found in both the BNT and the CRL database. Of these 36 items found in the database, 30 items were included in the reduced BNT test. Thus, most of the items compared had over 70% name agreement in the Maltese study. While keeping in mind the obvious linguistic differences in the names of each item in the different languages, it was evident that the mean RT of the Italian group (*Mean RT* = 1055 msecs, *SD* = 264) was faster than the mean RT of the Maltese group (*Mean RT* =

[6] Centre for Research in Language – International Picture-Naming Project.

[7] http://crl.ucsd.edu/ ~ aszekely/ipnp/1database.html

1488 msecs, *SD* = 384). A *t*-test showed that there is a significant difference ($t(35) = 7.77$, $p < .01$) between the matched groups of Italian and Maltese samples.

Thus, assuming that the main difference between the two samples (Italian and Maltese) was bilingualism, it may be concluded that the bilingual individuals in the present study were significantly slower than monolinguals, replicating the findings of previous research on different populations. In their study on Spanish – English bilinguals Hernandez and Kohnert (1999) found that the participants in their sample were most accurate and most rapid on the BNT when they used naming in a blocked condition. In a blocked condition, participants are instructed to use only one language before commencing test administration. In future studies it might be interesting to investigate BNT naming performance in blocked conditions with Maltese speakers, in order to determine whether being told to consciously block or deactivate one language actually reduces the lexical conflict between the two available languages resulting in faster RTs. On the other hand, it must be noted that when a naming test is used clinically, say with someone who has aphasia, blocked conditions would not be a possible option and the individual would simply be instructed to name the pictures presented.

6.3 NUMBER OF CORRECT RESPONSES FOR BNT ITEMS

The mean number of correct responses of Sample A was compared to the results reported by some studies on other populations (see Table 4.3 in Chapter 4). Despite the sample differences across the studies reported some interesting results emerged. It is evident that particular items in the present study have very low name agreement in comparison to other studies. This is evidently a result of the cultural differences across populations. For example, *pretzel* (item 19) and *beaver* (item 29) both had low name agreement (16.7% and 33.3% lenient number of correct responses respectively) in the Maltese participants of Sample A. The respondents of the main study produced an even lower number of correct responses (6.7% and 20% respectively). The Canadian and Swedish studies reported 92% and 99% correct responses for *pretzel*, and 98% and 90% correct responses for *beaver* respectively. On the other hand, the Greek study showed a tendency towards exhibiting low name agreement also on these same items. In fact, the

Greek version replaced *pretzel* with a more culturally-appropriate item and only 28% responded accurately for *beaver*. It is, therefore, possible that the similar findings that emerged in the Greek and Maltese studies were a result of the geographical proximity of the countries and possibly because of some cultural similarities between them.

6.3.1 Comparing studies - complete set strict scoring

When the complete set results of the present study were compared to other studies, the Swedish, Canadian and New Zealand studies all revealed a significantly higher number of correct responses than the Maltese study (Table 4.3, in Chapter 4). However, the Greek study did not differ significantly from the Maltese study. Thus, once again, it is possible that the two Mediterranean populations (Greek and Maltese) produced a similar number of correct responses for each item as a result of the cultural similarities between the two populations, as familiarity with the items of the BNT would probably be similar. As expected, particular items would be less familiar to populations that differ culturally and linguistically from the original population that was used to develop the test.

6.3.2 Comparing studies – complete set lenient scoring

When all the alternative responses were accepted as correct responses (i.e. lenient scoring) the Maltese study produced a significantly higher number of correct responses than the Greek study. In the present study synonyms and English responses were regarded as possible responses in lenient scoring. In the complete set, only 9 items had only one possible response and 51 items had more than one possible response. Despite using this very lenient scoring measure the Swedish, Canadian and New Zealand studies still produced a significantly higher number of correct responses than the Maltese study. On the other hand, in the Greek study only 4 items (*wheelchair*, *latch*, *tongs* and *globe*) had more than one possible response. Thus, this explains the advantage of the Maltese number of correct responses over those of the Greek study, as it was easier for the Maltese to score correct as a result of the numerous possible responses. Since the test was originally in English, the Maltese are at an advantage since both Maltese and English responses are accepted in lenient scoring. Therefore, scoring methods also need to be taken into account when comparing studies and must be specified when scoring and interpreting tests.

6.3.3 Comparing studies – reduced set strict and lenient scoring

The reduced set of the BNT consisted only of those items which had at least 70% name agreement in the initial study with Sample A. A total of 22 items which had less than 70% agreement were removed from the reduced set. In contrast to the complete set, the Maltese Sample A showed no significant difference from the Swedish, Canadian and New Zealand samples, and only the Greek study produced a significantly lower number of correct responses than the Maltese study (see Table 4.5, in Chapter 4). This was evident when using both strict and lenient scoring. Therefore, once the low name agreement items were removed from the BNT, the Maltese sample A could easily reveal similar naming accuracy to other populations (including Swedish, Canadian and New Zealand). On the other hand, both the complete and the reduced sets showed significant differences in performance (number of correct responses per item) between the main study Maltese sample and the Swedish, Canadian and New Zealand studies. Only the Greek study was not significantly different from the Maltese main study. As seen above these results may be a reflection of the similarities between the Greek and the Maltese populations.

Only the Maltese sample A (which included individuals with higher education than the Maltese main study) performed similarly to the Swedish, Canadian and New Zealand samples, and this was evident only when using the reduced set. When removing the low name agreement items from the test, the Maltese *main study* sample performed significantly poorer than the other studies. This has also been shown in section 6.2.1.2, which provided evidence of lower mean respondent scores in the Maltese sample than the scores obtained by other studies. There are considerable differences between the samples used in different studies which may account for the significantly different results. All the other studies were based on monolingual samples. The main study sample differed from the other studies in terms of culture, language, number of participants, education and age group. For example, the Canadian study (Tombaugh and Hubley 1997) involved a similar age range (25 – 88 years) to the main study (20 – 85 years). However, the former involved 219 participants who ranged from 9 to 22 years of education, while the Maltese study involved 60 participants with 0 – 19 years of education. Furthermore, the Canadian participants may have been much more familiar with the vocabulary used in the BNT than the Maltese respondents, as the test was originally developed in North America. The Swedish study (Tallberg 2005) involved a similar sample to the Canadian study, as it included 111 participants, aged between 21 and 80 years and with 6 to 22 years of education. On the other hand, the New Zealand study (Barker-Collo 2001) involved only 58 University students. Thus, the Swedish, Canadian and New Zealand studies all included participants of higher levels of education than the individuals of the Maltese main study. The Greek study (Patricacou et al 2007) included 100 participants, aged between 20 and 71+ years, and with 0 to 13+ years of education. Both the Greek and the Maltese main study included participants with very low levels of education and this factor may have been another reason for the similar findings across these two studies, and may account for the low mean scores obtained.

6.3.4 Word Variables

6.3.4.1 Word variables and number of correct responses

The number of correct responses for each item was found to be correlated with name agreement, familiarity, imageability, age of acquisition, syllable and phoneme length, cognate status and animacy. Stepwise regression analysis showed that, although a number of correlations were evident between different word variables, age of acquisition (AoA) emerged as the most significant predictor variable of the number of correct responses, while syllable length was not excluded from the model.

Unfortunately the available data on Maltese word frequency is only based on written text and items, such as *bed* (the first BNT item), were listed as low frequency items. It was concluded that this frequency data was not reliable enough to be used in analysis on word variables. Thus, it is not known whether word AoA would remain as the most important predictor if word frequency had to be included in a regression analysis too.

6.3.4.2 Word variables and reaction time

Regression analysis also showed that AoA was the most significant word variable predictor for naming latency. Thus, from the results that emerged it may be assumed that AoA is the most influential word variable for both naming accuracy and naming latency for Maltese individuals performing a picture naming task. These results replicate the findings of other studies on AoA (e.g. Iyer et al 2001; D'Amico et al 2001; Cortese and Khanna 2007). This information has important implications for assessment purposes as the words chosen must reflect these age of acquisition effects. Picture naming tests should, therefore, include both early AoA and late AoA words in order to provide an opportunity to retrieve both *easy* and *difficult* names. Incidentally, the reduced set of the BNT includes an equal amount of (19) early AoA and (19) late AoA items.

As noted above (section 6.2.1.3) there is also evidence of an interaction between AoA and level of education. No interaction was found between AoA and age. Thus, the most significant demographic variable (education) was found to be highly correlated with the most significant word variable (AoA). Words which are learned later in life tend to be more difficult to name for individuals who have a low level of education. On the other hand, late AoA words are easier to name for individuals who have a higher level of education. It is possible that people who acquire higher levels of education continue reading and learning new words, thus finding late AoA easier to retrieve than individuals with lower levels of education. Furthermore, since cognitively active individuals have been shown to be protected against cognitive decline, less educated individuals may be at a disadvantage in naming accuracy as a result of disuse of cognitive processes.

6.4 CONCLUSION

The first part of this research identified the main predictor variables for naming performance. There appears to be substantial evidence that education and age play an influential role in naming ability. Thus, norms for the Maltese population for an adapted BNT must be stratified for age and education. Further, some items of the BNT were found to be very low in name agreement and were, therefore, removed from the test in order to remain with a set of pictures that could be recognised and named by the Maltese population. Age and education effects were still evident on this reduced set and gender was not found to be correlated to performance on the reduced set. It was also shown that alternative responses (including English/Maltese translations of the target response) should be given credit as this (lenient) scoring method offers a better reflection of bilinguals' naming abilities. Therefore, a reduced set of the BNT with lenient scoring may be the most appropriate adaptation of the BNT for Maltese speakers. This is especially necessary since the test was originally normed on a different population.

While education was found to be the best predictor of naming accuracy, age was found to be the most influential predictor of naming latency. An ageing effect on naming accuracy was especially stronger when using the reduced set of the BNT. There is a particularly evident decline in naming performance after 66 years of age affecting both naming accuracy and naming latency. This replicates the results of other studies which found that ageing is accompanied by a slowing of cognitive processes resulting in longer reaction times. Ageing individuals with lower levels of education are, therefore, at a further disadvantage for naming performance. A longitudinal study and personal medical information (including presence of hypertension and diabetes) could provide valuable information regarding the true effect of ageing in order to eliminate the possibility of having cohort effects in cross-sectional studies.

The effect of bilingualism on naming performance also appears to be an important factor. Bilingualism may be contributing to the lower scores and slower reaction times observed in the present study as a result of the increase in conflict between competing names during lexical selection, activation and retrieval. A comparison of Maltese and Italian RT data using the same list of words showed significantly slower RTs for Maltese bilingual individuals in contrast to the (matched) Italian monolingual sample.

Age of acquisition effects are also expected to occur when the reduced set of the BNT is administered on Maltese speaking individuals, even in the absence of neurological disorders. There was evidence of an interaction between AoA and education, but no interaction between AoA and ageing. Thus, individuals with lower levels of education tend to be affected most by late AoA words. The potential effect of word variables on naming performance should also be considered when interpreting results of picture naming tests.

Different scores and number of correct responses per item were evident across studies (including differences between Sample A and the main study, and differences between the Maltese study and studies on other populations). It may be concluded that, although culture and language may play a part in producing different results, different sample characteristics (especially differences in levels of education of the participants) may play a more influential role. This was especially shown by the similarities that emerged between the Maltese and the Greek studies.

The following chapters will report the research carried out on people with aphasia. This research aimed at determining whether the proposed BNT is really a suitable test to assess Maltese speakers with naming impairments. The methodology used in this part of the research is reported in Chapter 7. Results and discussion of findings and conclusions are presented in the following chapters. The final chapter provides practical information regarding the use of a modified version of the BNT with Maltese speaking individuals. Recommendations regarding administration, scoring and interpretation of results obtained are provided. Normative data stratified by age and education will be presented for a modified BNT. Suggestions for qualitative analysis of test results will also be provided, according to the findings made during the second part of the research.

Chapter 7

Methodology 2

People with aphasia

7.0 INTRODUCTION

This chapter introduces the second part of the current research and focuses on the data obtained from people with aphasia. The use of the Boston Naming Test (Kaplan et al 1983) with this population is widespread, although studies which utilise the BNT to investigate naming in aphasia are less common in the literature (e.g. Laine and Martin 1996, Lambon Ralph et al 2002) in comparison to the numerous studies of the BNT on 'normal' populations. The BNT has been used as an additional test to other measures. For example, Lambon Ralph et al (2002) used the BNT as a screening tool in their research, and administered a set of other tasks in addition to it.

The following is a description of the aims of the study, the participants, the methods and the materials used with the people with aphasia (hereafter PWA).

7.1 THE AIMS OF THE STUDY

After presenting the results of the study on the people without neurological impairments, it is necessary to investigate the naming performance of PWA on the BNT. The first part of the research has already identified specific variables (such as years of education and age of acquisition) which affect naming performance in people without aphasia (hereafter PwithoutA). Possible modifications (such as reducing the number of items and an alternative scoring method) to the original BNT have already been proposed to cater for the linguistic and cultural characteristics of the Maltese population. It is necessary to determine whether the findings and proposals suggested may be applied clinically in the assessment of individuals with suspected word finding impairments. The main aim of the second part of the research is, therefore, to determine whether a modified version of the BNT (i.e. the reduced set of the BNT with lenient scoring) is an appropriate tool for assessing Maltese people with word finding difficulties. A set of four questions will be used to reach this aim:

1. Does the proposed test differentiate between Maltese PWA and PwithoutA?
2. Is this test a valid test for assessing Maltese PWA?
3. Is this test a reliable test for assessing Maltese PWA?
4. Is this test adequate for differential diagnosis in Maltese PWA?

Final recommendations for using the BNT on Maltese speakers will be made, with suggestions regarding possible modifications to test administration, test scoring and error profile interpretation.

7.2 MATERIALS AND PROCEDURE

Apart from the BNT, a set of materials were administered on each participant in the same order of presentation. An explanation of the aims of the study as well as the method of administration was provided to each participant prior to the actual administration of the tests. Also, prior to commencing the set of tasks, the participants (and sometimes, a relative) were interviewed in order to obtain relevant personal information about each person. Personal details that were noted down included age, gender, level of education, occupation, language experience, literacy, medical history, date of neurological impairment, and CT and/or MRI reports for nature and localization of neurological damage. An informed consent form was signed by each participant. The tests administered are described below.

7.2.1 Language Screening Test

The Language of the Elderly Screening Test (LEST, Delia et al, unpublished material) is the first standardized language screening test for Maltese adults. It has been developed by a group of Maltese speech and language pathologists (SLPs) and, as yet, it is still unpublished. This screening test consists of two sections: a receptive language section and an expressive language section. The receptive language section targets simple commands, object identification, sequential commands, complex commands, differences in meaning and narrative. The expressive language section includes object naming, definitions, verbal reasoning, abstract word finding, fluency naming (number of animals named in 1 minute) and sequencing. Clear instructions for administration, scoring and interpretation of the test results are available. The total possible score is 36, and each section (receptive and expressive section) has a highest possible score of 18 points. The cut off scores for the receptive language, expressive language and total score of the LEST are 13, 14 and 27 respectively.

Sixteen of the eighteen PWA who participated in this study failed this screening test (see Table 8.13 in Chapter 8 for LEST results). Two PWA (A01 and A13) did well on the screening test despite showing evidence of word finding difficulties in spontaneous speech and performing below the confidence intervals suggested on the reduced set of the BNT (see Table 6.1 in Chapter 6 for normative data and Table 8.13 in Chapter 8 for the scores obtained by the PWA). The receptive language and expressive language score of the LEST were both recorded. The fluency naming result (i.e. the number of animals named in one minute), which is part of the expressive language section, was also recorded and used as an independent variable during the analysis.

7.2.2 Additional tests of comprehension and repetition

Additional tests of comprehension and repetition were administered on the PWA as picture naming alone does not provide sufficient information to identify the level of breakdown in spoken word production (see section 2.2 in Chapter 2). The tests were designed by the researcher, except for 10 of the items in the comprehension test which were created together with a group of Maltese SLPs from the Aphasia Specialized Division at the Speech Language Department in Malta. The members of the Aphasia Specialized Division kindly provided consent for the researcher to use the mentioned items during the current research.

The test of auditory comprehension of single words contains 30 items (plus one practice item) and involves a word-picture matching task. The pictures are all colour photographs. Participants were requested to listen carefully to the word uttered by the researcher and to point to the corresponding picture from a set of four pictures. For each item the participants were presented with a semantic distractor, a phonological distractor, an unrelated distractor and the target word.

A 30-item (plus one practice item) real word repetition test and a 10-item (plus one practice item) non word repetition test were used. The repetition items were graded in syllable length and complexity, starting from simple CV monosyllables (e.g. 'il') to more complex 4-syllable words (e.g. 'irnexxielhom'). The PWA were requested to listen carefully to what the researcher said and to repeat the words and non words.

The responses were recorded on prepared forms. See Appendix E and F for the test of auditory comprehension of single words and test of repetition respectively. Each correct response (of both the comprehension and repetition tests) was given a score of 2 points. Self corrections, requests for repetition and delays were given a score of 1 point, and incorrect responses were given 0 points. Thus, a maximum score of 60 points was possible on the comprehension test and the real word repetition test. A score of 20 was the maximum number of points possible in non word repetition.

All the responses of each participant were inserted into a Microsoft Excel File and the scores were inserted into the *demographic variables* PASW (Predictive Analytics Software Statistics 18 for Windows, 2009) file described below. The scores of each participant are presented in Table 8.11 in Chapter 8 while section 7.3.1, below, explains how data was obtained for the test of auditory comprehension of spoken words and for the tests of nonword and real word repetition from a group of healthy controls.

7.2.3 Boston Naming Test (BNT, Kaplan et al 1983)

A hard copy of the Boston Naming Test (Kaplan et al 1983) was used during the second part of the study. The publishers of the test, Pro-Ed. Inc., kindly provided the researcher with a copy of the test for research purposes. Reaction times were not recorded.

The researcher tested all the PWA individually. The participants were told that a set of 60 pictures would be shown to them one by one, and that they were expected to name each picture with a single word whenever possible. They were encouraged to say the first word that came to mind when they saw the picture. No time limit was given. However, hesitations longer than 12 seconds were noted down (since the people without aphasia were allowed up to 11 seconds to respond – see section 3.5.5 in Chapter 3). If a participant showed that he or she was having some difficulty with producing a name, a phonemic cue was provided. Correct responses that followed a phonemic cue were not scored as correct. A note was made every time a phonemic cue was given to indicate whether the cue facilitated production or not. In the cases when the participants still could not provide a name, a note was made according to the individual's reaction. When participants stated that they *forgot* a name or that they *didn't know* a name, the researcher noted this down as *no response.* Whenever a response reflected a visual misinterpretation of the item a note, *visual association*, was made.

The researcher transcribed the response for each item and wrote down any comments regarding the participants' ability to name particular items, whenever necessary. The scoring procedure followed the same method utilised during the first part of the research. Strict and lenient scores were added for each participant. English responses and cognate responses were scored in the same manner as the research on people without aphasia. This means that English responses were regarded as lenient correct when they were not a target name and cognate responses were regarded as correct in both strict and lenient scoring (see section 3.3.7 in Chapter 3).

All the responses of each participant were inserted into a Microsoft Excel File, and errors were coded in a similar manner to the first part of the research. The PWA produced a more varied error profile than the PwithoutA and new error categories had to be added to the original list found in Table 3.8 (Chapter 3). The types of errors produced by the PWA are described in Chapter 8 (see section 8.2.3). The Percentages for each of the categories listed were worked out automatically by Microsoft Excel.

The data was then transferred to two different PASW files for further analysis and statistical procedures. The *demographic variables* PASW file consists of data and scores for each participant with aphasia and includes both categorical (e.g. age group, pattern of language acquisition and localization of neurological impairment) and quantitative predictor variables (e.g. age, score of the LEST, fluency naming, auditory comprehension score and repetition score). The following dependent variables were derived from the data collected: strict score, lenient score, and total number of English responses produced. Both complete and reduced versions of the demographic PASW file are available.

The second PASW file is the *word variables* file and consists of a number of quantitative predictor variables (imageability rating, age of acquisition rating, familiarity rating, phoneme and syllable length), and categorical predictor variables (word initial fricative/affricate, animacy and cognate status). The dependent variables derived from the data consisted of the strict number of correct responses for each item, the lenient number of correct responses for each item and the number of English words produced for each item. The strict and lenient number of correct responses of the PwithoutA were also included in this file for the possibility of comparing results. The results of this part of the study are reported in Chapter 8.

7.3 SAMPLES

Data was collected from two samples during the second part of the research. Since no norms were available for the comprehension and the repetition tests, the researcher administered these tests on 20 healthy Maltese individuals over 60 years of age. These extra participants served as controls (Sample E) and made it possible to acquire some normative data on these new tests (see section 7.3.1). Sample F consisted of 18 PWA.

7.3.1 Additional tests control sample (Sample E)

Sample E was utilised to acquire data on performance on the comprehension and the repetition tests from healthy individuals. This sample consisted of 20 healthy individuals aged between 61 and 87 years. The mean age of the participants was 73.45 years (*SD* = 6.60) and the mean years of education was 7.80 years (*SD* = 3.58). 10% were males and 90% were females. A total of 18 females made part of sample E since more women than men tend to attend day centres for the elderly.

The participants were mainly recruited from elderly day centres. Letters were addressed to the heads of department, including a description of the aims and procedures involved in the study. Permission for the researcher to visit the mentioned settings and to meet the volunteers was granted. The local Ethics Committee of the University of Malta approved the study and an informed consent form was signed by each participant before commencing the tests.

All the participants of sample E satisfied the inclusion criteria described in section 3.5.2 in Chapter 3. All controls were *healthy* Maltese-speaking individuals. To ensure that the participants could be included in the study, each person was interviewed about some basic personal information (e.g. age, years of education and medical history). The researcher filled in the information on the same form used during the main study (see Appendix D). The Clock Drawing Test was also carried out on each individual and all acquired an acceptable score of 4 or 5 points (see section 3.5.3 in Chapter 3). All the respondents reported normal or corrected-to-normal visual acuity. Any participant with a history of stroke, head injury, cerebral tumour, learning disability, epilepsy, history of concussion or loss of consciousness, psychiatric disorder or alcohol/drug abuse was excluded from the study. Table 7.1 presents demographic information about Sample E, as well as the comprehension and repetition scores obtained from the control sample (Sample E).

From the results of the comprehension test, it is evident that all the controls obtained scores of 58 (out of 60) or higher. Therefore, a “cut-off” score of 57 may be used to identify comprehension difficulties on this test. All participants of sample E were able to repeat the list of real word without difficulties (i.e. all obtained 60 points). On the other hand, the lowest score obtained on the non word repetition test was 14 points, leading to a “cut-off” score of 13 to identify repetition difficulties on the non word repetition test.

7.3.2 People with aphasia (Sample F)

People with aphasia (PWA) were assessed on the tests described in section 7.2. The participants were recruited from different settings, including an acute hospital, a rehabilitation hospital as well as community clinics. Speech Language Pathologists (SLPs) working in these settings were informed about the research and about recruiting potential participants from their case loads. Letters were addressed to the respective heads of departments including a description of the aims and procedures involved in the study. Permission for the researcher to visit the mentioned settings and to meet potential participants was granted. The local Ethics Committee of the University of Malta approved the study on people with aphasia.

Consent from participants was initially acquired when the SLPs informed them about participation in the study. Once this consent was obtained, the researcher contacted the participants and/or their relatives to set up an appointment. The assessments were carried out at a place that was convenient to each participant. The researcher met some people at their own home, and others at their community clinic or hospital. A consent form was signed before commencing the assessments. The personal details and medical history of each participant were recorded by the researcher on the form shown in Appendix G.

TABLE 7.1

	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
Age	79	77	67	64	70	77	66	66	73	73	70	74	79	79	82	74	72	79	61	87
Years of Education	12	6	8	11	8	13	14	10	6	10	7	9	8	7	5	8	4	1	9	0
Clock Drawing Test[1]	4	4	4	5	5	5	5	5	4	4	5	5	4	5	4	5	5	4	5	5
Scores on Auditory Comprehension and Repetition tasks																				
Comprehension of spoken words[2]	58	60	60	60	60	60	60	60	60	60	60	60	58	60	58	60	60	58	60	60
Repetition of real words[3]	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60
Repetition of non words[4]	20	18	16	20	18	20	20	20	20	18	20	20	20	14	14	16	18	14	20	20

Summary of results of Sample E (controls for additional tasks)

[1] CDT scoring involves a simple 0 – 5 point system with a "cut-off" score of 3 (Shulman 2000). 5 points refers to a 'perfect' clock & 4 points refers to 'minor visuo-spatial errors'.
[2] The test of auditory comprehension of spoken words is a 30-item word-picture matching task. Total possible score is 60 points.
[3] The test of repetition of real words involves repetition of 30 real words of increasing syllable length. Total possible score is 60 points.
[4] The test of repetition of non words involves repetition of 10 non words. Total possible score is 20 points.

All the participants of Sample F had experienced a cerebrovascular accident (CVA, 15 ischemic and 3 haemorrhagic). Localization of neurological impairment and time post onset varied across the PWA. Table 7.2 shows some demographic information about each person with aphasia.

During the recruitment stage, SLPs were given a small number of criteria to help them identify potential participants. The following criteria were given:

- Participants must be Maltese-speaking. They may be bilingual, but the dominant language must be Maltese;
- Participants must be over 20 years of age;
- They must have aphasia secondary to stroke or head injury;
- They must be able to participate in a picture naming task.

Due to the difficulties encountered in recruiting a number of PWA, the time post onset was not listed as a criterion. All participants who were accepted as part of the study were medically stable at the time of assessment by the researcher.

Sample F consisted of 18 participants, aged between 36 and 90 years. Their mean age was 70.17 years (*SD* = 12.79) and their mean years of education was 8.67 years (*SD* = 3.98), with a range of 0 to 17 years of education. There are two main education categories. 77.8% (N = 14) of the group had 0 to 12 years of education, while 22.2% (N = 4) had 13 or more years of education.

Table 7.3 tabulates the number of participants of each gender, age group and education category. Age and education were not contrasted orthogonally as the older age groups tended to consist of individuals with lower levels of education. No females with higher education levels participated in this study. This sample is a reflection of the sociological reality that exists in the Maltese islands, since older people (especially females) had fewer opportunities for education than the younger people of today.

Most of the participants had a history of manual types of occupations (55.6%). A further 16.7% of the PWA had worked in clerical jobs, and the same amount had managerial roles. 11.1% had a supervisory position. None of the PWA held jobs at the time of assessment.

TABLE 7.2

PWA	Age	Years Of Education	Gender	CVA	Localization	Time post onset
A01	62	17	male	ischaemic	Periventircular	2 mths
A02	36	10	male	ischaemic	L MCA	9 days
A03	73	7	male	ischaemic	L ischaemic frontobasal	6 mths
A04	73	9	male	ischaemic	L MCA	2½ yrs
A05	66	13	male	ischaemic	L basal ganglia; & several periventricular & deep white matter lesions	11 days
A06	69	13	male	haemorrhagic	L parietal region; & small lacunar ischemic lesions in basal ganglia	4 years 5 mths
A07	62	9	male	haemorrhagic & ischemic	L subarachnoid haemorrhage & parietal ischemic infarct	2 years 4 mths
A08	58	14	male	ischaemic	L posterior parietal cortex	7 mths
A09	65	0	female	ischaemic	L frontal lobe & temporal lobe	5 mths
A10	88	7	female	haematoma	L peri-orbital and supra-orbital	2 mths
A11	76	8	female	2 ischaemic	L corona radiata & R paraventricular region	2 yrs 3 mths
A12	72	9	female	2 ischaemic	(May 2010) L parietal lobe (extensive) & R parietal lobe (smaller) & (Jan 2011) new foci L side of pons & midbrain, & R thalamus	8 mths
A13	89	6	male	2 ischaemic	L MCA infarcts (2006 and January 2011)	4 yrs
A14	72	8	female	ischaemic	left Parieto-occipital region	3 mths
A15	61	10	male	ischaemic	L MCA	8 mths
A16	72	7	male	ischaemic	unavailable	5 mths
A17	90	3	female	ischaemic	L thalamus	6 mths
A18	79	6	female	ischaemic	left internal and external capsules (sub-acute ischaemia)	7 days

Demographic information about PWA

TABLE 7.3

		0 – 12 years of Education	13 or more years of Education	Total
Males	**Age Group**			
	30 – 45 years	1	0	1
	56 – 65 years	2	2	4
	66 – 75 years	3	2	5
	86 + years	1	0	1
	Total	**7**	**4**	**11**
Females	**Age Group**			
	56 – 65 years	1		1
	66 – 75 years	2		2
	76 – 85 years	2		2
	86 + years	2		2
	Total	**7**		**7**

The number of participants in each gender, age group and education category

All the respondents were native speakers of Maltese and all were exposed to some degree of English. 14 (77.8%) of the participants reported that they used Maltese only (M) on a daily basis, 2 (11.1%) used Maltese with some English words (Me) daily and 2 (11.1%) used both Maltese and English (ME) daily. Individuals who used English with some Maltese words (Em) or English only (E) were excluded from the study (see section 2.6.1.1, in Chapter 2).

7.4 RESULTS AND ANALYSIS

Scores were worked out for each participant. The scoring procedure adopted in this part of the research was the same one described in the first part of the study. Thus,

- target responses only were awarded a strict score of 1;
- both target and other acceptable/alternative responses (including English responses, when English was not the target word) were given a lenient score of 1;

- cognate responses were regarded as target responses and given a strict score and lenient score of 1;
- immediate self corrections were scored as correct;
- correct responses following a phonemic cue were *not* scored as correct (phonemic cueing was only used to identify whether the PWA could access the correct name following a phonemic cue);
- correct responses following conduite d'approche were *not* scored as correct in the PWA data. Since this type of error did not occur in the people *without* aphasia, it cannot be regarded as correct;
- hesitations which were longer than 12 seconds were scored as incorrect since people *without* aphasia were only allowed up to 11 seconds to respond.

Error profiles were worked out to determine the nature of errors produced by each participant. New error categories were added to the list of errors produced by the people without aphasia.

The data was then inserted into Microsoft Excel and PASW statistics files, together with the data from the other tests (LEST, comprehension and repetition) carried out. Chapter 8 presents the results obtained from the data collected from PWA. General results about the naming performance of PWA are reported first, together with a comparison of the different sets of the BNT and different scoring methods. Four sections follow the general results, each section having the aim of answering one of the important questions listed at the beginning of this chapter (see section 7.1). The results of PWA and PwithoutA are compared to determine whether the proposed test (reduced set of the BNT with lenient scoring) is able to differentiate between individuals with impaired and unimpaired naming (Question 1). Validity and reliability of the proposed test are reported in Questions 2 and 3 respectively. Finally, differential diagnosis is investigated in the final question. Different sources of evidence are looked at for potential information regarding differential diagnosis of naming impairment. The nature of the errors produced in the BNT, the word variable effects on naming accuracy of the BNT, and the performance on the additional tasks are all used to determine whether the proposed test allows differential diagnosis to be made.

Chapter 8

Results

People with aphasia

8.0 INTRODUCTION

Chapter 8 is concerned with the analysis of naming performance in people with aphasia (hereafter PWA). Cross-sectional data was derived from 18 Maltese-speaking participants who were diagnosed with aphasia, secondary to an ischaemic or a haemorrhagic cerebrovascular accident. Naming performance was primarily assessed by means of the complete set of the Boston Naming Test (BNT: Kaplan et al 1983). Additional tasks of comprehension and repetition of single words were also carried out. The tests administered on the PWA are described in Chapter 7, together with a description of the methods and procedures used.

The BNT was also used to analyse naming performance in Maltese speaking people *without* neurological impairments. The results obtained were reported in Chapter 4 and Chapter 5. It was concluded that the BNT could be used to assess naming performance in Maltese adults and that the reduced set (of 38 items) of the BNT, with lenient scoring was the best means of assessing this population (see Section 6.1.2, in Chapter 6). It is necessary to determine whether this claim remains true when it is applied to PWA. A number of questions must be answered in order to ensure that the recommendations for a modified version of the BNT (i.e. using the reduced set with lenient scoring) may be applied to Maltese people who have naming difficulties. It is imperative to determine whether this modified test can really differentiate between people with impaired naming and people without naming difficulties. It is also necessary to establish the validity and reliability of the recommended test. Finally, the usefulness of this test in differential diagnosis of Maltese PWA must also be investigated.

The first part of this chapter will report the results of the BNT administered on the PWA. The subsequent sections will utilise the data collected from these individuals to answer questions about the usefulness of the modified BNT for Maltese speakers with aphasia (see above). The results that emerge from the analysis will then be discussed further in chapter 9, together with reflections regarding implications of the results obtained for differential diagnosis of level of breakdown within a framework that relies on the cognitive neuropsychological approach to lexical access.

8.1 BNT RESULTS FROM PEOPLE WITH APHASIA

A summary of demographic information about the participants of this part of the research is presented in chapter 7 (also see Table 8.1). Four sets of BNT data are available, including scores and error profiles obtained from (i) the complete set (of 60 BNT items) using strict scores only, (ii) the complete set using lenient scores, (iii) the reduced set (of 38 BNT items) using strict scores only, and (iv) the reduced set using lenient scores. The score results are reported in the following section. Error profiles are presented in section 8.2.3 and 8.5.4 of this chapter.

8.1.1 Scores obtained from people with aphasia

Table 8.1 presents the scores obtained from 18 PWA (A01 – A18), together with some demographic information. The scores of the four sets of the BNT are shown as well as the number of English responses produced in each set. It is evident that the different PWA had varying levels of naming performance ranging from a strict score of 3 to 27 (out of 38) on the reduced set of the BNT.

8.1.2 Comparing complete and reduced sets

The scores obtained by the PWA on the complete and the reduced set of the BNT were compared. A Kolmogorov-Smirnov test confirmed that the dependent variable, score, had a normal distribution as the p-value exceeds the 0.05 level of significance in strict score and lenient score of both complete and reduced sets (see Table 8.2).

A strong positive correlation emerged between the strict score results of the complete set and the strict score results of the reduced set ($r = .985$, $p < .01$), showing that participants who scored high on the complete set also scored high on the reduced set and vice versa. Lenient scoring also showed a strong positive correlation when complete and reduced sets were compared ($r = .986$, $p < .01$). Similar findings were

TABLE 8.1

SCORES OBTAINED BY PEOPLE WITH APHASIA																		
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**	**A13**	**A14**	**A15**	**A16**	**A17**	**A18**
Age (in years)	62	36	73	73	66	69	62	58	65	88	76	72	89	72	61	72	90	79
Gender	M	M	M	M	M	M	M	M	F	F	F	F	M	F	M	M	F	F
Education (in years)	17	10	7	9	13	13	9	14	0	7	8	9	6	8	10	7	3	6
Complete Set[1]																		
Strict Score	31	21	10	5	25	24	22	11	17	11	15	13	24	12	3	23	7	18
Lenient Score	33	21	10	5	27	26	26	13	20	11	21	22	27	14	3	24	7	19
Number of English responses	11	2	1	0	9	5	4	1	1	0	10	12	7	3	0	11	0	2
Percentage of English responses[2]	33.3	9.5	10	0	33.3	19.2	15.4	7.7	5.0	0	47.6	54.5	25.9	21.4	0	45.8	0	10.5
Reduced Set[3]																		
Strict Score	27	18	9	5	20	23	20	11	17	11	13	10	20	12	3	19	7	17
Lenient Score	28	18	9	5	22	24	21	13	17	11	18	19	20	14	3	19	7	17
Number of English responses	6	1	0	0	5	5	3	1	1	0	8	12	4	3	0	5	0	1
Percentage of English responses	21.4	5.6	0	0	22.7	20.8	15.0	7.7	5.9	0	44.4	63.2	20.0	21.4	0	26.3	0	5.9

[1] In complete sets the score is out of 60 responses.
[2] Percentage of English responses out of the number of lenient correct responses for each individual.
[3] In reduced sets the score is out of 38 responses.

TABLE 8.2

MAIN STUDY	STRICT score Complete Set	LENIENT score Complete Set	STRICT score Reduced Set	LENIENT score Reduced Set
Kolmogorov-Smirnov Z	.505	.613	.618	.763
P-value (2-tailed)	.960	.846	.839	.605

One-Sample Kolmogorov-Smirnov Test (score)

reported when correlations were performed using the data obtained from the people without neurological impairment (see section 4.3.1, in Chapter 4).

Table 8.1 shows that some participants obtained similar scores in the reduced and the complete sets, implying that the correct responses were all items from the reduced set. The following participants obtained similar scores in the complete and reduced sets: A03, A04, A06, A08, A10, A14, A15, A17 and A18. Thus, in these cases, the reduced set provided the same information as the complete set, in terms of scores. Paired samples t-tests were performed in order to determine whether there is a significant difference between the scores of the complete and the reduced sets. Both strict scoring (t (17) = 4.043, $p < .01$) and lenient scoring (t (17) = -4.043, $p < .01$) methods indicated that the complete and the reduced sets of the BNT produced a significantly different score. There was a significant difference of 1.7% in score when strict scoring was used, and a significant difference of 2.5% when lenient scoring was used to compare complete and reduced sets of the BNT. Thus, there is a small, but significant, difference between the percentage score of the complete and the reduced sets, as was evident in the results of the people without aphasia (hereafter PwithoutA) too. However, the difference is much smaller in the results of the PWA than those of the PwithoutA (18.5% and 14.8% difference in strict and lenient scores respectively).

8.1.3 Comparing scoring methods

Correlation statistics on data collected from PwithoutA showed a strong positive correlation between strict and lenient scoring methods (see section 4.33 in Chapter 4). The data obtained from PWA showed similar results. Strong positive correlations emerged between the two scoring methods when using the complete set (r = .962, p <

.01) and when using the reduced set (r = .939, $p < .01$). This implies that PWA who scored high when the strict scoring method was used also scored high when lenient scoring was used and vice versa. Lenient scoring methods produced higher scores, as expected. Only two participants (A11 and A12) showed a somewhat larger discrepancy between their strict and their lenient scores. This discrepancy is a result of the amount of English responses produced (8 and 12 English responses respectively), since English responses are given credit in lenient scoring only.

A paired-samples t-test showed a significant difference of 2.1% between the scores achieved when using strict and lenient scoring (t (17) = -3.652, $p < .05$) in the complete set, and a significant difference of 1.3% when using the reduced set (t (17) = -2.335, $p <$ .05). In PwithoutA, a small, but significant difference between scoring methods was also found in the complete set (5.7%) and in the reduced set (2%). In PWA this difference is even smaller.

8.1.4 Number of English responses used

The number (and percentage) of English responses produced by the PWA is shown in Table 8.1. When counting the number of English responses, all items produced in English were counted, including those items which had an English target word (e.g. mushroom). Those English target words which did not have any other alternative names in Maltese (e.g. protractor, helicopter) were not added to the number of English

TABLE 8.3

Complete Set		**Reduced Set**	
Mean number of English responses	**SD**	**Mean number of English responses**	**SD**
People With Aphasia			
4.39	4.41	3.06	3.33
18.8%		15.6%	
People Without Aphasia			
11.63	7.68	6.45	4.05
26.9%		20.2%	

Mean number of English responses produced by PWA and PwithoutA

responses, however. This procedure was applied both with PWA and PwithoutA. The percentage is derived from the number of English responses produced out of the number of lenient correct responses for each individual. Table 8.3 shows the mean number of English responses produced by PWA and PwithoutA respectively, in the complete and reduced sets.

It is apparent that the PWA produced a smaller mean number of English responses possibly because there was less opportunity to produce English responses as a result of the smaller number of correct responses produced in comparison to the PwithoutA.

8.2 DOES THIS TEST DIFFERENTIATE BETWEEN PEOPLE WITH APHASIA AND PEOPLE WITHOUT APHASIA?

The introduction to this chapter referred to a number of important questions that must be answered in order to ensure that a modified BNT is an appropriate tool for assessing Maltese people with impaired word finding abilities. The first question will be the focus of this section and will aim to investigate whether this test is really able to identify people with lexical retrieval difficulties and to differentiate them from people with unimpaired word finding. The data obtained from PWA will be compared to that obtained from the PwithoutA in terms of scores, range of scores, error profiles.

8.2.1 Comparing scores

The scores obtained by PWA will be compared to the scores obtained by PwithoutA. All four sets of the BNT (i.e. the complete set with strict and lenient scoring and the reduced set with strict and lenient scoring) will be compared.

Table 8.4 presents the mean scores and standard deviations for the two samples (the PWA and the PwithoutA). The mean strict score obtained by the PWA when using the reduced set (14.6) is much lower than that obtained from the PwithoutA in the main study (28.9). Similar differences between mean scores are evident when using the other sets of data (see Table 8.4 for lenient scoring in the reduced set, and strict and lenient scoring in the complete set). While there are a number of differences between the samples presented (including sample size, levels of education, age group and methodology), the most obvious reason for this discrepancy in scores is the presence of a neurological impairment in the PWA.

Independent samples t-tests were carried out to ensure that there is a significant difference between the scores of the PWA and the PwithoutA. A significant difference was evident in the complete set with strict scoring (t (76) = 7.46, p < .05), in the complete set with lenient scoring (t (76) = 7.76, p < .05), in the reduced set with strict scoring (t (76) = 8.54, p < .05), and in the reduced set with lenient scoring (t (76) = 8.15, p < .05). Thus, there is a significant difference between the mean scores obtained by PWA and by PwithoutA, indicating that all sets (including the reduced set with lenient scoring) may be used to differentiate between groups of people *with* aphasia and groups of people *without* aphasia (the main study).

TABLE 8.4

People *with* Aphasia Complete Set (score out of 60 items)				**People *with* Aphasia Reduced Set** (score out of 38 items)			
Mean Strict Score	**SD**	**Mean Lenient Score**	**SD**	**Mean Strict Score**	**SD**	**Mean Lenient Score**	**SD**
16.2	7.8	**18.3**	8.6	**14.6**	6.5	**15.8**	6.7
People *without* Aphasia Complete Set (score out of 60 items)				**People *without* Aphasia Reduced Set** (score out of 38 items)			
Mean Strict Score	**SD**	**Mean Lenient Score**	**SD**	**Mean Strict Score**	**SD**	**Mean Lenient Score**	**SD**
34.6	9.5	**39.2**	9.7	**28.9**	6.2	**30.3**	6.2

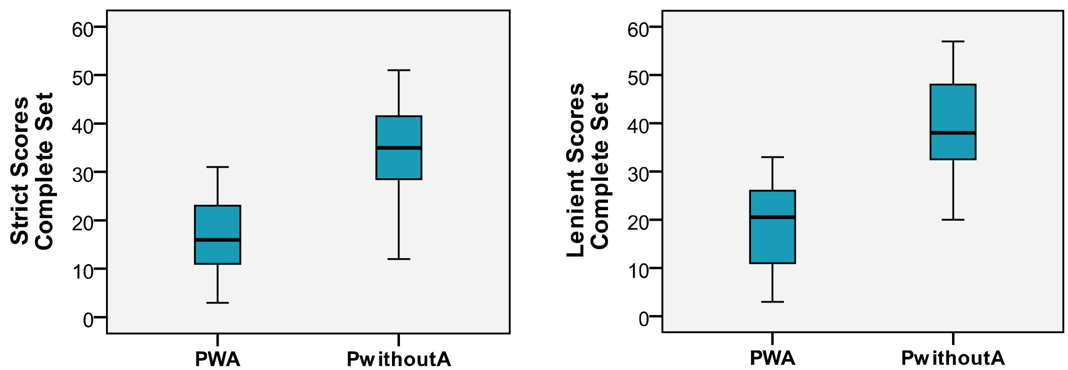
Global Mean Scores

8.2.2 Comparing range of scores

Different participants showed varying levels of naming performance as they obtained a range of scores (see section 8.1.1). The range of scores obtained by the PWA will now be compared to the normative data obtained from the first part of the research. The range of scoring for both populations is shown in the figures below and in Table 8.5.
It is evident that, although the mean scores are significantly different (see above), there is some overlap in the range of scores of the two samples. Some PwithoutA scored well

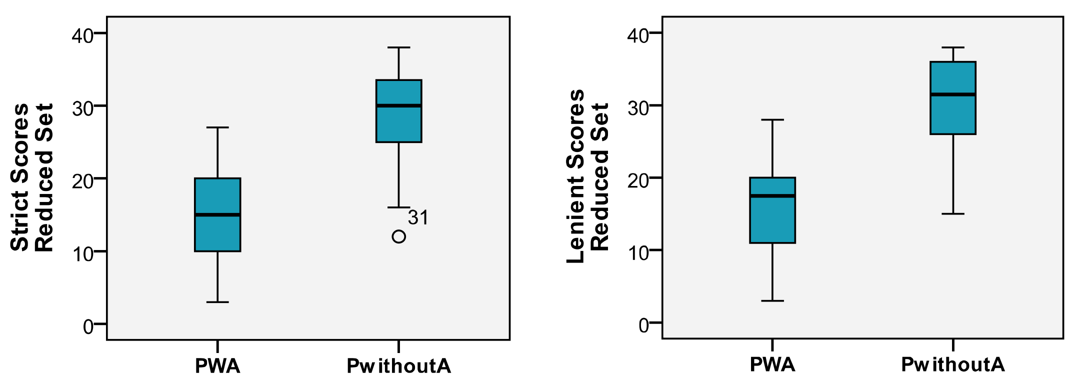
below the mean score of the ‘normal' population, and individual scores of the PWA were all below the mean score obtained by the PwithoutA.

FIGURE 8.1



Range of scores obtained by PWA and PwithoutA
Complete Set

FIGURE 8.2



Range of scores obtained by PWA and PwithoutA
Reduced Set

Independent samples t-tests were carried out to compare patients’ naming performance to that of each category of the PwithoutA (see normative data presented in Table 6.1 in

Chapter 6). The categories are: (1) < 66 years of age and ≥ 13 years of education, (2) < 66 years and 0-12 years of education, (3) ≥ 66 years and ≥ 13 years of education, and (4) ≥ 66 years and 0-12 years of education. The scores of all the different groups were found to be statistically different from the mean scores of the PWA (as a whole group) on all sets of the BNT. Each group of the PwithoutA was then compared to the respective group from the people with aphasia. The mean scores were found to be statistically different from each other. There was only evidence of one group (the older people with higher education) which did not show a significant difference in mean scores *($t(5) = .385$, $p = 0.72$)* on the reduced set with strict scoring). However, this result is not reliable as there were only 2 persons *with* aphasia (A05 and A06) and 6 people *without* aphasia in this category for comparison purposes.

TABLE 8.5

Complete sets				
	PWA strict score	**P without A strict score**	**PWA lenient score**	**P without A lenient score**
Mean score	16.2	34.6	18.3	39.2
Minimum score	3	12	3	20
Maximum score	31	51	33	57
Range	28	39	30	37

Reduced sets				
	PWA strict score	**P without A strict score**	**PWA lenient score**	**P without A lenient score**
Mean score	14.6	28.9	15.8	30.3
Minimum score	3	12	3	15
Maximum score	27	38	28	38
Range	24	26	25	23

Range of scores obtained by people with aphasia and unimpaired people

Although t-tests showed a significant difference between mean scores of each category (except for one category), the results need to be interpreted with caution since some PwithoutA obtained very low scores.

A number of individuals who participated in the main study on PwithoutA achieved scores which were lower than the 68% confidence interval for their age and education level. The lowest (reduced set) lenient score obtained by a respondent without aphasia was 15. Figures 8.3 and 8.4 demonstrate the frequency distribution of the scores of the PWA and the PwithoutA respectively. Figure 8.4 shows that 9 individuals scored below the 68% confidence intervals of one standard deviation from the mean on lenient scoring. These 9 PwithoutA obtained the lowest lenient scores on the reduced set. They are not outliers in the data as there appears to be a steady progression in the scores. Table 8.6 presents data on each of these individuals.

TABLE 8.6

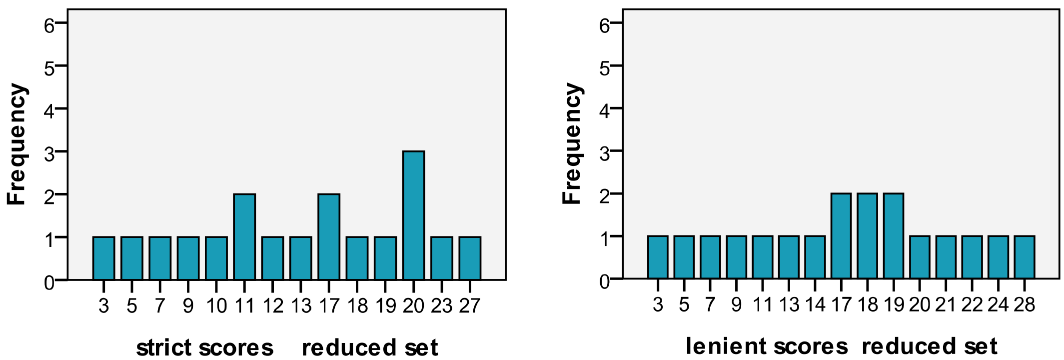
PwithoutA	Lenient Score	Age	Gender	Years of Educ	68% Confidence interval	95% Confidence interval	Main types of errors
R31	15	84	Male	14	26.8 - 30.2	15.9 – 36.1	8 ph cues; 4 V Assoc
R42	18	75	Male	7	22.8– 25.1	21.6 - 27.7	2 ph cues; 6 V Assoc
R49	18	75	Male	9	22.8– 25.1	21.6 - 27.7	7 ph cues; 5 V Assoc
R32	19	79	Female	6	22.8– 25.1	21.6 - 27.7	3 ph cues; 6 V Assoc
R 36	20	82	Female	3	22.8– 25.1	21.6 - 27.7	2 ph cues, 7 V Assoc
R50	20	85	Female	13	26.8 - 30.2	15.9 – 36.1	4 ph cues; 6 V Assoc
R34	21	77	Female	8	22.8– 25.1	21.6 - 27.7	1 ph cue, 6 V Assoc
R44	22	64	Female	9	27.7 - 30.4	26.4 – 29.9	2 ph cues, 5 V Assoc
R45	22	60	Female	9	27.7 - 30.4	26.4 – 29.9	2 ph cues, 4 V Assoc

PwithoutA who scored below confidence intervals

All of these respondents were over 60 years of age and 6 out of the 9 persons were females. These results (which are derived from the reduced set) may show a tendency for older females to be at a disadvantage on the reduced set of the BNT, possibly reflecting the socio-cultural experiences of these PwithoutA. However, the three individuals who obtained the lowest scores were males. This number of people (9 individuals out of 60 respondents) suggests that those obtaining scores below the mean score are not simply outliers in the data. This indicates that score alone is not a reliable enough measure for identifying people with naming difficulties.

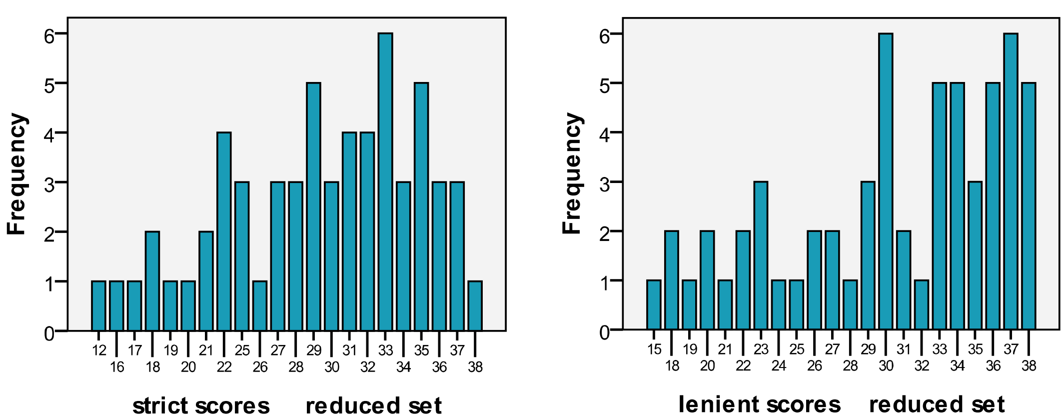
One of the most common types of errors produced by these 9 PwithoutA was visual associations. Visual associations may be a result of lack of familiarity with an item, resulting in the production of the name of objects which may be visually similar to the target item. Although the people in the main study were not screened for visual impairments, all the PwithoutA were screened on and passed the Clock Drawing Test to minimize the possibility of including individuals with visual impairments in the control study.

FIGURE 8.3



Frequency of scores obtained by PWA in reduced set

FIGURE 8.4



Frequency of scores obtained by PwithoutA in reduced set

These individuals were frequently able to produce a correct response following a phonemic cue (at the end of the test). This occurred when they experienced Tip of the tongue difficulties which are commonly encountered in older individuals (see section 2.5.2 in Chapter 2).

In a further attempt to differentiate the PWA from the PwithoutA, new confidence intervals were worked out to allow for 2 standard deviations from the mean rather than 1 standard deviation. However, all of the nine controls (except R50) still performed below the new confidence intervals (95% confidence intervals).

It has become increasingly evident that scores do not provide substantial evidence to differentiate between impaired and unimpaired naming, especially in individuals over 60 years of age. It is also essential to look at error profiles in order to obtain further information about naming performance. The next section will look at the error categories produced by the PWA and by PwithoutA.

8.2.3 Comparing error profiles

The PWA who participated in this research produced more *types* of errors than those produced by the PwithoutA. Table 8.7 lists the types of errors produced by both samples, together with the mean number of each type of error. It is evident that (a) some types of errors (e.g. phonologically related non words, neologisms, conduite d'approche) were not produced at all by PwithoutA (atypical errors), and (b) the common errors appeared more frequently in the PWA (see the mean number of each type of error). The types of errors which were produced by PWA only are listed in bold in Table 8.7. Only self corrections were more frequent in the data obtained from PwithoutA (mean of 0.65 self corrections versus mean of 0.11 in PWA). An independent samples t-test showed that there is a significant difference between the mean number of errors produced by PWA and those produced by the PwithoutA ($t(40) = 3.187$, $p < .05$). It is important to note that none of the PwithoutA produced any of the atypical errors and all of the PWA showed evidence of some atypical errors in their error profile.

The Error profiles of each individual with aphasia were also compared to the worst norms in order to determine whether PWA differ from the worst error profile obtained by the PwithoutA in terms of error categories. Table 8.8 shows the error profiles for each individual with aphasia as well as the worst norm (WN) for each type of error.

TABLE 8.7

Types of Errors	PWA mean	PwithoutA mean
Coordinate	1.28	0.75
Superordinate	1.11	0.65
Subordinate	0.06	0.03
Semantic association (S Assoc)	1.72	0.52
Circumlocution	2.28	1.75
Multiword (MW)	1.61	0
Phonologically related non word (PhRN)	0.67	0
Phonologically related real word (PhRR)	0.33	0.18
Morphological alteration	0.11	0
Perseveration	0.28	0
Unrelated	0.5	0
Forgot	0.67	0
Don't know	2.11	0.75
Visual Association/not recognized	2.44	1.68
Self correction	0.11	0.65
Successful conduite d'approche	1.00	0
Unsuccessful conduite d'approche	0.17	0
Neologism	0.06	0

Mean number of errors produced by PWA and PwithoutA

The following differences were noted:

- Particular error types (in bold) simply did not appear in the PwithoutA.
- The most common type of error produced by PwithoutA was circumlocution, and they also responded well to phonemic cues when retrieval was difficult.
- Visual Associations were also common errors made by PwithoutA.
- Semantic and phonological errors were sometimes produced by the PwithoutA, but they were more frequently produced by individuals with aphasia.

TABLE 8.8

Type of Errors	**A01**	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**	**A13**	**A14**	**A15**	**A16**	**A17**	**A18**	**WN[4]**
Coordinate	1	3	3		1		1	1	1		4	1	2	1	1	2	2	2	2
Superordinate	2	1			1		1		1	4	3	1		1		1		3	2
Subordinate					1														1
Semantic association	2	2	2		1	3	1			2	2	3	1	2	1	1	5	3	2
Circumlocution	2	8			2	1	3	1	3		3	5	3	3	3	2		2	6
Multiword		1			1		1		1						24				
Phonologically related non word		2	1				1	2	2			2						2	
Phonologically related real word						2	1	1			1							1	1
Morphological alteration		1													1				
Perseveration					1	1				1							1		
Unrelated	1									2		1		2			2		
Forgot	1		4	3	1					1		1							
Don't know			5	7	1	1	1	1	2	5	3		1	3			7	1	2
Visual Association/not recognized		1	2		1				1	6	5	5	1	5	2	2	8	5	4
Self correction						1							2				1		2
Correct following a phonemic cue			9	20	5	5		9	3	4			1	8	2	1	4		8
Incorrect following a phonemic cue			3	3				9	4	1					2		1		
Successful conduite d'approche							4		3				6			4	1	1	
Unsuccessful conduite d'approche							1	1					1						
Neologism											1								

Error profile of each person with aphasia vs worst error profile of PwithoutA

[4] PwithoutA – Worst Norms

The most striking difference between the errors produced by the PWA and PwithoutA in this research is that particular errors were only produced by the PWA (9 different types of errors listed in bold in Table 8.7). While PWA produced both circumlocutions and multiwords (i.e. empty connected speech that fails to describe the picture), PwithoutA only produced circumlocutions, showing that only impaired individuals had difficulties with describing a target when they could not name a picture. Further evidence of impaired naming is the production of neologisms, phonologically related non words and unrelated words, all of which were not produced by PwithoutA. Finally, conduite d'approche which resulted in correct or incorrect naming of a target was evident only in PWA. PwithoutA did not experience this repeated trial and error strategy to arrive at a target name.

8.2.4 Comparing number of English responses

Both PwithoutA and PWA produced a number of English responses (see Section 8.1.4, this Chapter). This section aims to investigate whether the same items tend to produce English words in these two samples. The same procedure described in section 8.1.4 was used to calculate the number of English responses produced for each item. Table 8.9 shows the percentage of English responses used for each item. Some items do not have a different Maltese-English translation as the same target name is used in both languages (e.g. helicopter, dart, etc.).

Simple linear correlations showed that there is a strong significant correlation between the use of English words by PWA and by PwithoutA ($r = .766$, $p < .01$). This implies that the same items of the reduced set tend to be named in English by both samples. Earlier it became apparent that the PwithoutA tend to use more English words than the PWA – possibly as a result of lack of opportunity to produce more correct words as a result of the naming impairment (see section 8.1.4). However, when item responses were compared (rather than mean number of English responses produced by each individual), t-tests failed to demonstrate that the difference between the two samples' use of English is significant. Another t-test was performed on the items having early age of acquisition only. However, even these items failed to show a significant difference between the number of English words used by PWA and by PwithoutA. It is worth noting that the first few words of the reduced BNT (see Table below) show a tendency for more use of English responses by the PWA. This tendency is no longer evident in the other items. Yet, the difference between the two samples in the use of English throughout the 38 items failed to reach significance. Some items only have one possible response (e.g. cactus, protractor) and the percentage has been left blank.

TABLE 8.9

Reduced set of the BNT	English responses used by PwithoutA (%)	English responses used by PWA (%)
Bed	0	7
Tree	0	17
Pencil	0	23
House	2	13
Whistle	3	11
Scissors	0	18
Comb	2	0
Flower	0	8
Saw	0	7
Toothbrush	91	100
Helicopter		
Broom	7	38
Octopus	19	0
Mushroom	96	100
Hanger	2	0
Camel	15	20
Mask	3	0
Bench	2	25
Snail	2	20
Volcano	2	0
Dart		
Globe	0	0
Rhinoceros	67	100
Igloo		
Dominoes	63	50
Cactus		
Escalator		
Harp	16	14
Knocker	6	0
Stethoscope	83	0
Pyramid	53	33
Unicorn	94	100
Funnel	2	0
Compass	32	0
Tripod	74	100
Palette	75	0
Protractor		
Abacus	100	100

English responses produced by PwithoutA and PWA (in percentages)

8.2.5 Comparing order of difficulty of BNT items

This section will look into the items of the BNT, rather than the performance of each individual who participated in the study. Only the reduced set will be used while comparing these results.

The data obtained from the Maltese PwithoutA and PWA was used to identify the frequency of correct responses for the items of the reduced set of the BNT. Items with a larger number (or percentage) of correct responses were regarded as less difficult and more frequently named correctly, while items with a smaller number of correct responses were regarded as more difficult and less frequently named correctly. Table 8.10 presents the sequence of the reduced set items for the PwithoutA and the PWA, starting with more frequent to less frequent naming accuracy. The first column presents the original order of the (reduced) BNT items. The order of difficulty varies across the 3 groups.

The sequence obtained from the PwithoutA may be a result of cultural and linguistic differences across populations, considering that the original BNT was normed on a very diverse population. On the other hand, the sequence of items which emerged from the PWA data may reflect the variety of levels of difficulties encountered by the PWA. These 18 people are not a homogenous group in terms of severity and level of breakdown in spoken word production. Therefore, it may not be appropriate to generalize the order of difficulty of BNT items to all people with aphasia. Twenty-two items are being omitted from the original BNT to accommodate the needs of the Maltese speakers. The original sequence of the remaining BNT items will be retained in order to match the standardization process.

TABLE 8.10

Original BNT	PwithoutA	PWA
bed	bed	tree
tree	tree	scissors
pencil	pencil	house
house	house	bed
whistle	scissors	saw
scissors	flower	pencil
comb	bench	flower
flower	comb	bench
saw	saw	comb
toothbrush	hanger	hanger
helicopter	snail	funnel
broom	camel	camel
octopus	funnel	snail
mushroom	octopus	globe
hanger	mushroom	whistle
camel	globe	helicopter
mask	helicopter	broom
bench	toothbrush	octopus
snail	harp	volcano
volcano	volcano	mushroom
dart	compass	harp
globe	broom	mask
rhinoceros	whistle	dart
igloo	mask	compass
dominoes	cactus	toothbrush
cactus	pyramid	igloo
escalator	dart	cactus
harp	rhinoceros	tripod
knocker	escalator	knocker
stethoscope	dominoes	pyramid
pyramid	igloo	dominoes
unicorn	tripod	escalator
funnel	knocker	rhinoceros
compass	protractor	unicorn
tripod	abacus	palette
palette	palette	abacus
protractor	stethoscope	stethoscope
abacus	unicorn	protractor

Sequence of reduced set BNT items ordered according to frequency of correct responses

8.3 IS THIS TEST A VALID TEST FOR ASSESSING MALTESE PEOPLE WITH APHASIA?

Concurrent validity will be investigated to ensure that the proposed reduced test of the BNT correlates well with other measures carried out on the same individuals at the same time. It is necessary to determine whether the scores obtained from the test are positively correlated with other standardized tests of language ability.

8.3.1 LEST

The LEST (Language of the Elderly Screening Test, Delia et al unpublished) screening measure will be used to validate the reduced set of the BNT. This is a standardized language screening tool, which was primarily developed for Maltese adults at risk of language difficulties. Four segments from the LEST will be used as independent variables: the fluency naming section (i.e. the number of different animals named in one minute), the receptive language score, the expressive language score and the total score. These will be correlated with dependent variables strict score and lenient score (on the reduced set of the BNT). The LEST is described in Chapter 7.

Sixteen of the eighteen PWA who participated in this study failed this screening test (see Table 8.11 for LEST results). Two PWA (A01 and A13) did well on the screening test despite showing evidence of word finding difficulties in spontaneous speech and performing below the confidence intervals suggested on the reduced set of the BNT. These two participants passed the LEST as a whole, and also passed the expressive and the receptive parts of the test. However, they both failed the fluency naming section, as they named 9 and 10 animals respectively. The fluency naming section involved saying as many animal names as possible in one minute. The authors of the LEST base the scoring criteria of the fluency section on a different study (Azuma et al 1997) and they suggest giving a score of 3 points if a person names between 12 and 21 animals, and a score of 0 if the person names less than 12 animals. The authors

TABLE 8.11

Reduced set of the BNT																		
	A01	**A02**	**A03**	**A04**	**A05**	**A06**	**A07**	**A08**	**A09**	**A10**	**A11**	**A12**	**A13**	**A14**	**A15**	**A16**	**A17**	**A18**
Strict Score	27	18	9	5	20	23	20	11	17	11	13	10	20	12	3	19	7	17
Lenient Score	28	18	9	5	22	24	21	13	17	11	18	19	20	14	3	19	7	17
Number of English words used	6	1	0	0	5	5	3	1	1	0	8	12	4	3	0	5	0	1
LEST																		
Expressive language score	14.0	6.0	1.5	1.0	6.5	5.5	8.0	5.0	4.5	3.0	12.0	7.0	14.0	3.5	0	11.0	4.0	6.0
Receptive language score	16.0	17.0	14.0	14.0	12.0	12.0	14.0	12.0	11.0	6.0	13.0	13.0	17.0	4.5	4.0	15.0	8.5	18.0
Total score	30.0	23.0	15.5	15.0	18.5	17.5	22.0	17.0	15.5	9.0	25.0	20.0	31.0	8.0	4.0	26.0	12.5	24.0
Fluency naming	9	6	2	2	4	7	9	6	7	2	6	2	10	3	2	11	3	7
Auditory Comprehension and Repetition tasks																		
Comprehension of spoken words	60	59	53	58	56	57	58	59	60	42	56	58	59	37	49	59	38	60
Repetition of real words	60	60	56	60	60	59	60	60	52	58	58	54	48	60	8	60	48	60
Repetition of non words	20	20	14	18	10	12	4	16	14	12	18	6	2	20	0	20	4	18

Summary of results

of the LEST did not use Maltese normative data to score fluency naming in their test. All the other participants also failed the fluency naming section.

Simple linear correlations were carried out between the subtests of the LEST and the scores obtained on the BNT. This was necessary in order to determine whether the proposed reduced set of the BNT correlates well with other measures of language that are carried out on the same people at the same time, thus, showing concurrent validity.

Simple linear correlations showed that all segments of the test were positively correlated with both types of scoring methods (see Table 8.12). The correlations were high and positive, showing evidence of strong criterion validity and indicating that the scores of the reduced set of the BNT are related to the scores obtained on the LEST segments. This further implies that individuals obtaining higher scores on the proposed test are likely to obtain high scores on the LEST as well, and vice versa. As predicted, the correlation between the receptive language score and the (lenient) score of the reduced set of the BNT is somewhat lower ($r = .531$, $p < .05$) than the correlations with fluency naming ($r = .696$, $p < .01$), expressive language score ($r = .777$, $p < .01$) and the total score ($r = .736$, $p < .01$) of the LEST, since all these scores are related to spoken word production rather than comprehension. In the future, predictive validity may be measured as an additional measurement of criterion validity, to ensure that when the same individuals are tested again at a later date their results will match the present results.

TABLE 8.12

		Fluency naming	**Receptive language score**	**Expressive language score**	**Total score of the LEST**
Strict score	Pearson Correlation	.795**	.550*	.721**	.715**
	P - value	.000	.018	.001	.001
Lenient score	Pearson Correlation	.696**	.531*	.777**	.736**
	P - value	.001	.023	.000	.000
	Sample size	18	18	18	18

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Simple linear correlations between scores and segments of LEST

8.3.2 Interaction of additional tests on score

Apart from the LEST, two other tests (developed by the researcher and described in section 8.5.2 below) were administered on the PWA. These tests were devised to screen repetition of real words and non words, and auditory comprehension of single words. This section investigates the interaction between these additional tasks and dependent variable score in the reduced set. Table 8.13 shows a summary of the correlations that exist between the dependent variables, and the predictor variables single word comprehension and repetition.

Moderately strong and positive correlations became evident between strict and lenient scores and real word repetition (r = .481, $p < .05$ and r = .511, $p < .05$ respectively) and auditory comprehension (r = .498, $p < .05$ and r = .520, $p < .05$ respectively). This is further evidence of criterion validity as the reduced set of the BNT correlates well with other measures of language ability carried out on the PWA at the same time. No significant correlation was found between scores and non word repetition. However, there is a strong positive correlation between real word repetition and non word repetition (r = .635, $p < .01$).

TABLE 8.13

		Strict score	Lenient score	Real word repetition	Non word repetition	Comprehension
Strict score	Pearson Correlation	1.00	.939**	.481*	.247	.498*
	P-value		.000	.043	.323	.035
Lenient score	Pearson Correlation	.939**	1.00	.511*	.215	.520*
	P-value	.000		.030	.391	.027
Real word repetition	Pearson Correlation	.481*	.511*	1.00	.635**	.235
	P-value	.043	.030		.005	.348
Non word repetition	Pearson Correlation	.247	.215	.635**	1.00	.190
	P-value	.323	.391	.005		.450
Comprehension	Pearson Correlation	.498*	.520*	.235	.190	1.00
	P-value	.035	.027	.348	.450	
N		18	18	18	18	18

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Summary of correlations between score and additional tests

Regression analysis was performed to identify the most significant predictor variable of score. All the additional tasks (except for non word repetition) were inserted into the model with dependent variables strict and lenient scores. Thus, the independent variables included fluency naming, expressive score of LEST, receptive score of LEST, total score of LEST, comprehension of spoken words and repetition of real words. Since some predictors are correlated with each other, a stepwise regression analysis was carried out to avoid the multicollinearity problems described in Chapter 4. Fluency naming emerged as the most important predictor for both strict and lenient score, while all the other variables were excluded from the model. This implies that fluency naming is the best predictor of score on the reduced set of the BNT.

TABLE 8.14

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	5.293	2.015		2.626	.018
	Fluency naming	1.701	.325	.795	5.234	.000

Coefficients using reduced set strict score

The coefficients for fluency naming (1.70) indicated that for every extra one word named in fluency naming, the strict score is expected to increase by 1.70 (Table 8.14). The coefficient of determination (R squared) was 0.631 implying that this model explains 63.1% of the total variation of the dependent variable strict score.

TABLE 8.15

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	7.472	2.453		3.046	.008
	Fluency naming	1.536	.396	.696	3.882	.001

Coefficients using reduced set lenient score

When lenient score was used as the dependent variable the coefficients for fluency naming (1.54) indicated that for every extra one word named in fluency naming, the lenient score is expected to increase by 1.54 (Table 8.15). The coefficient of determination (R squared) was 0.485 implying that this model explains 48.5% of the total variation of the dependent variable lenient score.

8.4 IS THIS TEST A RELIABLE TEST FOR ASSESSING MALTESE PEOPLE WITH APHASIA?

Inter-rater reliability was estimated by looking at scores and error profiles of five of the eighteen PWA. A Speech Language Pathologist, who was not involved in the research itself, acted as inter-rater. The inter-rater was given the data obtained from the complete set of the BNT, of 5 PWA of varying severity and she was asked to score the test for each one. She was given instructions to carry out both strict and lenient scoring. The inter-rater was also requested to categorize each error and to produce an error profile for each individual. Written and verbal explanations about scoring methods and error profiles were given to the inter-rater to ensure consistency across raters. The variation between the scores and the error profile of the researcher and the inter-rater was investigated.

8.4.1 Comparing scores across raters

Pearson's correlations were carried out to determine whether the scores given by the researcher and the inter-rater were correlated with each other. The table below shows the scores given by the researcher and the inter-rater.

TABLE 8.16

	Strict Score Researcher	Strict Score Inter-rater	Lenient Score Researcher	Lenient Score Inter-rater
A 03	9	9	9	9
A 05	20	19	22	22
A 09	17	17	17	17
A 11	13	13	18	18
A 13	20	19	20	20

Scores given by researcher and inter-rater

There is evidence of a strong positive correlation between the scores given by the researcher and the scores given by the inter-rater, for both strict scoring (r = .997, $p < .01$) and lenient scoring (r = 1.000, $p < .01$). This shows that there is consistency and general agreement between the scores given by the two raters.

8.4.2 Comparing error profiles across raters

The error profiles drawn by the researcher and the inter-rater were compared. Table 8.17 presents this data for each of the 5 PWA. The main differences that were evident between the error profiles lie in the classification of semantic errors. The inter-rater tended to generalize most semantic errors as semantic associations and failed to specify the particular type of semantic error in some cases:

- For participant A 03, the inter-rater classified response *giraffe* for *camel* as a semantic association, while the researcher classified it as a coordinate error.
- For participant A 05, the inter-rater classified response *Etna* for *volcano* and *stairs* for *escalator* as semantic associations, while the researcher categorized them as subordinate and superordinate errors respectively.
- There was complete agreement between the error profiles of participant A 09.
- For participant A 11, the inter-rater classified the semantic errors as semantic associations, while the researcher considered response *brush* for *broom* to be a coordinate, *stairs* for *escalator* to be a superordinate, and *cave* for *igloo* to be a visual association.
- Finally, there was only one mismatch in the error profile of participant A 13. While the response *girna*[5] for *igloo* was categorized as a visual association by the researcher, the inter-rater classified it as a coordinate.

Although there were some differences in the error profiles drawn by the two raters, it is worth noting that there was agreement that the errors were semantic in nature, and that the differences were only in the categorization of the type of semantic errors. These results emphasize the need to ensure that anyone scoring the test and classifying errors must know the exact procedure of scoring and must understand what each error category means. Unless there is consistency among raters to score and interpret results

[5] A small circular stone hut typically found in the Maltese countryside.

the test will not be regarded as reliable. Thus, testers and raters must read and abide by the scoring instructions and must be properly trained to interpret results.

TABLE 8.17

Type of Errors	**A 03**		**A 05**		**A 09**		**A 11**		**A 13**	
	R[6]	**I**[7]	**R**	**I**	**R**	**I**	**R**	**I**	**R**	**I**
Coordinate	3	2	1	1	1	1	2	1	2	3
Superordinate			1		1	1	3	2		
Subordinate			1							
Semantic association	2	3	1	3			2	5	1	1
Circumlocution			2	2	3	3	3	3	3	3
Multiword			1	1	1	1				
PhRN	1	1			2	2				
PhRR							1	1		
Morphological alteration										
Perseveration										
Unrelated										
Forgot	4	4	1	1						
Don't know	5	5	1	1	2	2	3	3	1	1
Visual Association/not recognized	2	2	1	1	1	1	5	4	1	
Self correction									2	2
Successful conduite d'approche					2	2			6	6
Unsuccessful conduite d'approche					1	1			1	1
Neologism							1	1		

The error profiles drawn by the researcher and the inter-rater

[6] Researcher's error profile
[7] Inter-rater's error profile

8.5 IS THIS TEST USEFUL FOR DIFFERENTIAL DIAGNOSIS IN MALTESE PWA?

This is the fourth and final question that will be targeted in the results chapter. The focus of this section is differential diagnosis and whether the proposed test enhances the identification of the level of breakdown in word retrieval.

Assessment must be comprehensive enough to enable the clinician to identify impaired and intact processes. Picture naming alone may not provide sufficient information to identify the level of breakdown in word retrieval. Other sources of evidence (including auditory/written comprehension and repetition tasks, analysis of word variable effects and types of errors) need to be utilized in order to reach more accurate conclusions regarding an individual's difficulties, and to plan treatment accordingly. The cognitive neuropsychological approach holds that intervention can be directly designed to target the person with aphasia's particular difficulties and abilities. However, it is necessary to keep in mind that there may not be a direct relationship between a specific kind of therapy and a particular level of breakdown, but rather that the identification of the impaired and intact processes limits the "range of possible treatments" (Whitworth et al 2005: 109).

The current research collected data from 18 PWA. Five main tasks were carried out with each individual: the LEST, the BNT, a 30-item single word auditory comprehension test, a 30-item real word repetition task, and a 10-item non word repetition task. Chapter 7 provides a detailed description of each task. Some results relating to picture naming have already been reported in the preceding sections of this chapter. Table 8.11 summarizes the results obtained from the different tasks for each person with aphasia.

8.5.1 Identifying the level of breakdown

The ultimate aim of this final question is to determine whether the reduced set of the BNT provides substantial information to identify the level of breakdown in spoken word production and, therefore, to make a differential diagnosis. The other sources of evidence will be used to identify the possible areas of breakdown, and the results obtained will, then, be compared to the information derived from the reduced set of the BNT.

- **Performance on additional tasks** (section 8.5.2)

  The first source of evidence to be used will be the performance on the additional tasks administered on the PWA, i.e. on the tests of auditory comprehension of spoken words and repetition. Simple linear correlations and regression analysis between the additional tasks and scores of the reduced set of the BNT have already been reported (see section 8.3), and the relationship between the additional tasks and the BNT was stated.

  The PWA were categorized into groups according to their performance on the repetition and comprehension tasks, in order to predict the level of breakdown for each individual. Four different groups are described, each one associated with a particular level of breakdown. Information derived from the other two sources of evidence (variable effects and the types of errors produced by each participant) may also be used to predict the level of breakdown for each participant and will be compared to the information derived from the additional tests (i.e. the group).

- **Effect of word variables** (section 8.5.3)

  The second source of evidence will be the effect of word variables. The current research on the PwithoutA has shown that age of acquisition and syllable length were the most important predictors of naming accuracy. It is necessary to determine whether the same holds true when using the data obtained from PWA. Logistic regression is used to determine whether each participant with aphasia experiences variable effects, and whether variable effects may be used to predict level of breakdown for each participant.

- **Types of errors** (section 8.5.4)

Finally, the types of errors will be used to predict the level of breakdown for each participant. The most common types of errors made by each participant will be identified and compared to the groups of PWA, in order to determine whether different groups tend to make particular types of errors. Chi Square tests are carried out to identify the relationship between them. If particular types of errors are found to be clearly associated with particular groups of PWA, it may be possible to predict the level of breakdown, and, therefore, to make a differential diagnosis from the error profile obtained from the reduced set of the BNT.

8.5.2 Information from additional tests

The tests of auditory comprehension of single words, and of repetition of words and non words served as additional sources of information for identifying intact and impaired levels of processing. Moderately strong positive correlations were found between the additional tasks and scores (section 8.3.2, this chapter), implying that PWA who performed well on the comprehension test and on the repetition test were more likely to obtain a higher score on the naming test.

8.5.2.1 Criteria for passing additional tests

A set of criteria to determine who fails or passes the additional tests was required. A cut off score for the comprehension and the repetition tasks was derived from the range of scores obtained by the controls on the same tests (see section 7.3.1 in Chapter 7). Some difficulties became apparent when applying this procedure to the repetition tests. Some (a) PWA performed within *normal* limits on the test of real word repetition, but failed the non word repetition test, while other (b) PWA tended to fail on both repetition tests. It was necessary to decide whether failure on non word repetition with adequate real word repetition should be regarded as impaired repetition or not. In order to solve this difficulty, two different repetition schemes were tested to determine which method was the most predictive of level of breakdown.

The error profiles of each participant were compared to the two different repetition schemes in order to determine which method was most predictive of level of breakdown. Information on participant A07 is presented here as an example (see Table

8.18). This person with aphasia had poor non word repetition, adequate real word repetition and adequate comprehension. When using scheme 1, A07 was said to have poor repetition and was categorized as having a possible phonological assembly deficit. When using scheme 2, he was said to have adequate repetition and was categorized as having a possible phonological output lexicon deficit. The error profile was then compared to the schemes. A07 produced both semantic and phonological errors. These type of errors are more consistent with impaired phonological output lexicon than phonological assembly deficits, implying that scheme 2 is preferred. On the other hand, it is also necessary to point out that A07 may be impaired in *both* phonological output lexicon and phonological assembly levels, since he produces both semantic and phonological errors. Yet, participant A05 may also be used to support scheme 2. In this case, A05 failed the comprehension test and the non word repetition test, but was perfectly able to repeat all the real words. His picture naming errors were semantic in nature and there was no evidence of phonological errors. His performance on the different tests is consistent with a lexical semantic level deficit without impairment at the phonological assembly level, confirming that scheme 2 provides more reliable information as it is possibly better at predicting level of breakdown.

Thus, although it may not always be possible to exclude possible additional deficits (as in A07 above) we may assume that adequate real word repetition reasonably excludes deficits at the phonological assembly level.

TABLE 8.18

Participant	Scheme 1	Scheme 2	Type of Errors
A07	Repetition is said to be impaired if non word repetition is impaired, despite good real word repetition	Repetition is said to be adequate if only non word repetition is impaired and real word repetition is good	
Impaired non word repetition; Adequate real word repetition; Adequate comprehension of single words	Possible phonological assembly deficit	Possible phonological output deficit	Semantic errors PHRN[8]& PHRR.

Repetition tests

[8] PHRN = Phonologically related non word; PHRR = Phonologically related real word.

Thus, only the real word repetition test results will be taken into account when determining who had impaired and intact repetition skills. The PWA who failed the non word repetition test, but did well on real word repetition were considered to pass the repetition test.

When repetition of non words is impaired and repetition of words is spared (or less impaired) as described in (a) above, damage to a post-lexical level may be the cause. A similar case is described by Franklin et al (2002), who describe client MB as having impairment at the phonological assembly level. On the other hand, selective impairment of non word repetition may also arise from damage to the nonlexical route of phonological input-to-output conversion, as in client MK (Franklin 1989) who was unable to perform a non word repetition. The direct nonlexical route from auditory phonological analysis to phonological assembly is impaired in this case. The non lexical route is not only associated with repetition of non words, however. Hanley et al (2004) showed that both the lexical-semantic route and the nonlexical route play a part in repetition of familiar words, while picture naming relies only on the lexical-semantic route. This is consistent with the *summation hypothesis* (Hillis and Caramazza 1991), which allows activation from both routes during repetition of words, allowing repetition to be better than picture naming of the same words. The present research did not test repetition of the same words of the picture naming test and, therefore, it is not possible to test the dual route here.

8.5.2.2 Groups of PWA

Using the additional tasks of comprehension and repetition it was possible to categorize the PWA into 4 different groups and to identify the possible levels of breakdown in word retrieval (see section 2.2 in Chapter 2). The groups' performance on the additional tests are described below. We shall assume that:

- participants who fail the auditory comprehension of single words test are possibly impaired at the lexical semantic level of spoken word production (Group 1);
- participants who fail the repetition test have possible difficulties with phonological assembly (Group 3);

- participants who have adequate comprehension and repetition skills are possibly impaired at the level of the phonological output lexicon (Group 2);
- participants who fail on both the comprehension and the repetition tests have a more severe impairment in spoken word production (Group 4).

Group 1	comprehension only is impaired (i.e. possible lexical semantic deficits)
Group 2	comprehension and repetition are adequate (i.e. possible phonological output lexicon deficits)
Group 3	Repetition only is impaired (i.e. possible phonological assembly deficits)
Group 4	both comprehension and repetition are impaired (i.e. a more severe group)

Table 8.19 summarizes the available information about each person with aphasia and includes the group number, scores, fluency naming and the performance on the additional tasks. The number and type of errors produced by each participant are also tabulated. The table also includes the number of English responses produced when English was not a target name. The main types of errors and the PWA groups will be compared in the final part of this chapter.

When referring to Table 8.19, it should be noted that:

- The error profile does not include errors made by controls. Therefore, the number of errors differs from the errors listed in Table 8.8.
- PWA were categorized into groups according to their performance on the additional tasks of comprehension and repetition (as described above).
- Fluency naming refers to the number of animal names produced in one minute.
- The PWA produced other errors which are not listed in the table – including some morphological alterations, visual associations and unrelated responses. Others responded correctly to phonemic cues when they could not produce a name (e.g. A04).

TABLE 8.19

GROUPS OF PWA

Group [9]	PWA	Strict Score	Lenient Score	English words	Fluency naming	Comprehension of single words	Word repetition	Types of Errors produced in naming		Main types of errors
1	A05	20	22	2	4	56	60	1 coordinate 1 subordinate 1 S Assoc 2 circumlocution	2 don't know	Sem
1	A06	23	24	4	7	57	59	3 S Assoc	2 PHRR 1 don't know	Sem and Phon
1	A10	11	11	1	2	42	58	4 superordinate	6 don't know	Sem and No Resp
1	A11	13	18	6	6	56	58	4 coordinate 1 superordinate 1 S Assoc 2 circumlocution	1 PHRR 3 don't know	Sem and No Resp
1	A14	12	14	2	3	37	60	2 coordinate 2 S Assoc 1 circumlocution	3 don't know	Sem and No Resp
2	A01	27	28	2	9	60	60	2 S Assoc 1 circumlocution	1don't know	Sem
2	A02	18	18	1	6	59	60	1 coordinate 3 S Assoc 7 circumlocution	2PHRN	Sem and Phon

9

Group 1 comprehension only is impaired (i.e. possible lexical semantic deficits)
Group 2 comprehension and repetition are adequate (i.e. possible phonological output lexicon deficits)
Group 3 repetition only is impaired (i.e. possible phonological assembly deficits)
Group 4 both comprehension and repetition are impaired (i.e. a more severe group)

Group	PWA	Strict Score	Lenient Score	English words	Fluency naming	Compre hension of single words	Word repetition	Types of Errors produced in naming		Main types of errors
2	A04	5	5	0	2	58	60		10 don’t know	No Response
2	A08	11	13	1	6	59	60	1 coordinate 1 circumlocution	1 PhRR 2 PhRN	Sem and Phon
2	A07	20	21	0	9	58	60	2 S Assoc 3 circumlocution	2 PhRN 1 PhRR 5 cond d’appr	Sem and Phon
2	A16	19	19	0	11	59	60	1 superordinate 1 S Assoc 1 circumlocution	4 cond d’appr 11 hesitations	Sem and Phon
2	A18	17	17	0	7	60	60	2 coordinate 1 S Assoc 2 circumlocution	1 PHRR 2 PHRN 1 cond d’appr	Sem and Phon
3	A09	17	17	0	7	60	52	1 circumlocution	2 PHRN 3 cond d’appr	Phon
3	A12	10	19	12	2	58	54	1 coordinate 1 superordinate 3 S Assoc 3 circumlocution	2 PHRN	Sem and Phon
3	A13	20	20	0	10	59	48	2 coordinate	1 don’t know 9 cond d’appr	Phon
4	A03	9	9	0	2	53	56	3 coordinate 2 S Assoc	1 PHRN 9 don’t know	Sem and No Resp
4	A15	3	3	0	2	49	8	1 coordinate 1 S Assoc 3 circumlocution	24 MWs	Sem and No Resp
4	A17	7	7	0	3	38	48	4 coordinate 1 superordinate 2 S Assoc	7 don’t know	Sem and No Resp

Groups of PWA

8.5.3 Effects of word variables on naming accuracy

Regression analysis on data collected from people without neurological difficulties, showed that syllable length and, especially, age of acquisition had the most significant effects on naming accuracy (see section 5.4 in Chapter 5). Simple linear correlations and regression analysis were performed in order to determine whether the same effects of predictor variables, AoA and syllable length, emerged when using the data collected from PWA.

8.5.3.1 Correlations and regression analysis

A Kolmogorov-Smirnov test confirmed that dependent variable number of correct responses had a normal distribution, as the p-value exceeds the 0.05 level of significance (Table 8.20). Simple linear correlations with AoA as the independent variable showed that there is a significant negative correlation between AoA and number of correct responses.

TABLE 8.20

	STRICT number Correct	**LENIENT number correct**
Kolmogorov-Smirnov Z	.572	.622
P-value (2-tailed)	.900	.834

One-Sample Kolmogorov-Smirnov Test (number correct)

Both strict and lenient sets produced a strong negative correlation ($r = -.708$, $p < .01$ and $r = -.729$, $p < .01$ respectively). A significant negative correlation was also found between syllable length, as independent variable, and number of correct responses. Both strict and lenient sets produced a moderately strong negative correlation ($r = -.505$, $p < .01$ and $r = -.498$, $p < .01$ respectively) (Table 8.21).

This demonstrates that similar effects of word variables occurred in both samples used in this research. Regression analysis was also performed to identify the most significant predictor of naming accuracy in the PWA. When all the word variables (also including phoneme length, imageability and familiarity) were inserted into a regression analysis model, AoA, once again, appeared as the most significant independent variable for number of correct responses. Syllable length was not excluded from the model, while all the other word variables were excluded (Table 8.22).

TABLE 8.21

		Strict Number Correct	Lenient Number Correct
Age of Acquisition	Pearson Correlation	-.708**	-.729**
	P - value	.000	.000
Syllable length	Pearson Correlation	-.505**	-.498**
	P - value	.001	.001
	Number of items	38	38

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations between number correct and Age of Acquisition and Syllable Length

The coefficients for age of acquisition (-14.38) indicates that for every extra 1 point on the scale of AoA ratings (as described in section 3.4 of Chapter 3), the number of correct responses is expected to decrease by 14.38%. The coefficient of determination (R squared) is 0.615, implying that this model explains 61.5% of the total variation of the responses (dependent variable).

TABLE 8.22

Model		Unstandardized Coefficients		Standardized Coefficients	t	P-value
		B	Std. Error	Beta		
1	(Constant)	86.695	8.538		10.154	.000
	AoA_Ratings	-14.377	2.390	-.708	-6.015	.000
2	(Constant)	106.748	9.827		10.863	.000
	AoA_Ratings	-12.594	2.200	-.620	-5.726	.000
	Syllable_Length	-10.880	3.377	-.349	-3.222	.003

Coefficients (number of correct responses and predictor variables)

8.5.3.2 Logistic regression

Logistic regression analysis was performed to determine whether the individual PWA experienced AoA and syllable length effects. Table 8.23 presents p-values, odds ratios and overall percentages for predicting AoA and syllable length effects for each person with aphasia. The table also includes information regarding the PWA's groups – as determined from the performance on the additional tasks. It is evident that AoA effects were experienced by all the PWA but A11, A12 and A15, while syllable length effects were only evident in A07 and A18. If variable effects are to be considered as predictive

TABLE 8.23

Participant	Group	Word variable	P-value	Odds ratios	Overall percentages
A01	**2**	**AoA**	**.014***	3.11	81.6%
		Syllable length	.078		
A02	**2**	**AoA**	**.037***	2.14	68.4 %
		Syllable length	.125		
A03	**4**	**AoA**	**.049***	2.18	68.4 %
		Syllable length	.716		
A04	**2**	**AoA**	**.012***	8.59	89.5%
		Syllable length	.513		
A05	**1**	**AoA**	**.049***	1.89	63.2%
		Syllable length	.274		
A06	**1**	**AoA**	**.020***	2.35	71.1%
		Syllable length	.401		
A07	**2**	**AoA**	**.033***	2.41	84.2%
		Syllable length	**.011***	5.45	
A08	**2**	**AoA**	**.004***	4.86	73.7%
		Syllable length	.714		
A09	**3**	**AoA**	**.004***	5.01	84.2%
		Syllable length	.229		
A10	**1**	**AoA**	**.039***	2.61	78.9%
		Syllable length	.069		
A11	**1**	AoA	.156		55.3%
		Syllable length	.923		
A12	**3**	AoA	.730		73.7%
		Syllable length	.730		
A13	**3**	**AoA**	**.037***	2.193	78.9%
		Syllable length	.256		
A14	**1**	**AoA**	**.010***	3.24	71.1%
		Syllable length	.629		
A15	**4**	AoA	.101		94.7%
		Syllable length	.187		
A16	**2**	**AoA**	**.011***	5.22	84.2%
		Syllable length	.201		
A17	**4**	**AoA**	**.027***	3.58	84.2%
		Syllable length	.368		
A18	**2**	**AoA**	**.008***	4.13	78.9%
		Syllable length	**.049***	3.60	

*. Correlation is significant at the 0.05 level (2-tailed).

Logistic regression

of level of breakdown, they should be associated with the results of the additional tasks. Therefore, it is proposed that anyone showing evidence of length effects would have deficits at the phonological assembly level of word production (Group 3). On the other hand, it is proposed that anyone showing age of acquisition effects in naming would have deficits at the lexical semantics level (Group 1) and possibly at the phonological output lexicon level (Group2).

The two PWA who experienced length effects (A07 and A18) are not the same participants who were categorized as Group 3 (phonological assembly group). Rather, from the results of the additional tests they have been categorized as being in Group 2 (POL deficits). On the other hand, the 15 PWA who were found to have age of acquisition effects were from different groups. In Group 1, four out of the five PWA experienced AoA effects, in Group 2, seven out of the eight PWA showed AoA effects, in Group 3, both PWA, and in Group 4, two out of the three PWA experienced AoA effects. Therefore, it appears that in the current research variable effects may not be used to predict the level of breakdown, as AoA effects were apparent in all the different groups and syllable length effects were not apparent in Group 3 as expected.

8.5.4 Types of errors

The relationship between types of errors and the level of breakdown in single word production will be investigated in relation to the four groups identified from the additional tests. It should be noted that errors which occurred in the data obtained from the PwithoutA were excluded from the error profiles of the PWA. For example, 'horse' for 'unicorn' was a common error made by both the controls and the PWA. Therefore, this error cannot be regarded as an 'aphasic' error.

8.5.4.1 Relationship between groups and type of errors

The different levels of breakdown in spoken word production and the typical errors that are associated with each level have been described in Chapter 2 (see section 2.2.1). The participants have already been grouped according to their performance on the tests of spoken word comprehension and repetition, and the possible levels of breakdown have been identified for each group. The error profiles of each participant (see Table 8.19) may also be used to identify possible levels of breakdown for each person. From each individual's error profile it was possible to identify the main type of error/s produced by each participant.

The relationship between the four groups and the main types of errors (the final column in Table 8.19) was investigated by means of a Chi-Square test. It is proposed that if particular errors are consistently associated with particular groups of PWA (reflecting their levels of breakdown), then error profiles alone may be useful in identifying levels of breakdown in spoken word production.

Two categorical variables, group and main type of errors are available. Categorical variable main type of errors has five levels: *semantic errors*, *semantic and phonological errors*, *semantic errors and no responses*, *phonological errors* (including conduite d'approche) and *no responses* (i.e. when participants say they *don't know, forgot or can't name a picture*), while categorical variable group has four levels (lexical semantic group, phonological output lexicon group, phonological assembly group and severe group). The results displayed in the tables below show that there is a significant association (Chi-Square value = 21.6, $p < .05$) between the four different PWA groups and the type of errors they produced.

TABLE 8.24

	Lexical Semantic Group	**Phonological output lexicon group**	**Phonological assembly group**	**Severe group**	**Total**
Semantic errors	1	1	0	0	2
Semantic and Phonological errors	1	5	1	0	7
Semantic errors and no responses	3	0	0	3	6
Phonological errors	0	0	2	0	2
No responses	0	1	0	0	1
Total	5	7	3	3	18

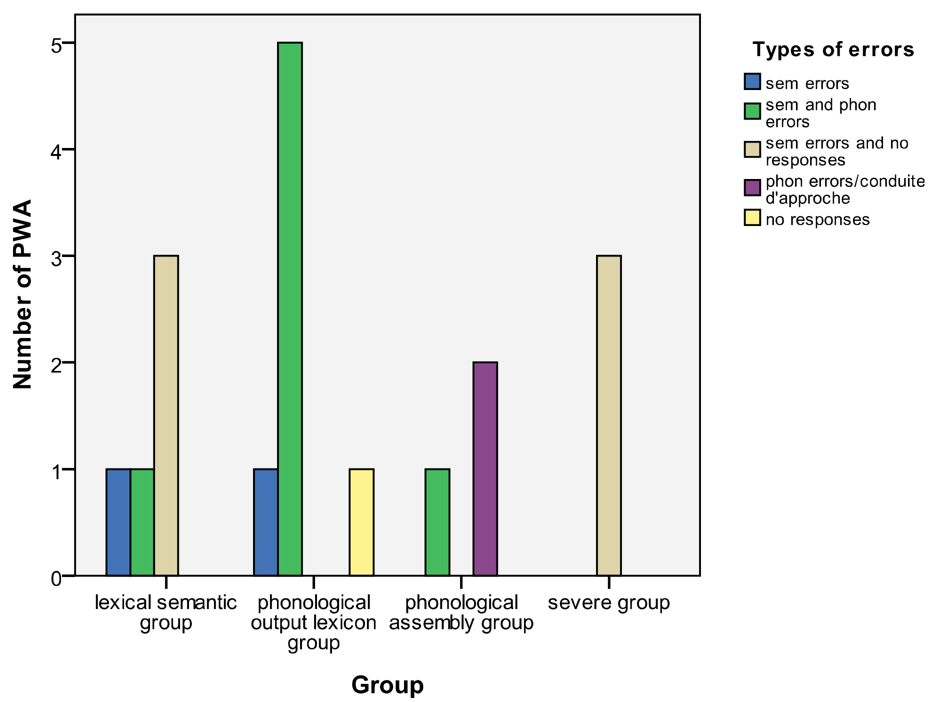
Main types of errors and level of breakdown/group

	Value	**df**	**P-value**
Pearson Chi-Square	21.61	12	.017

Chi-Square Tests

This implies that each level of breakdown may be said to be associated with particular types of errors. Table 8.24 and the Figure 8.5 display these associations. For example, the lexical semantic group tends to produce *semantic errors and no responses* mainly, the phonological output lexicon group tends to produce *semantic and phonological errors* mainly, the phonological assembly group tends to produce *phonological errors* mainly, and the severe group tends to produce *semantic errors and no responses*. This result may have important implications for differential diagnosis as evidence of particular types of errors may be said to indicate the impaired level in spoken word production. The Chi-Square test indicates that this association is significant (p-value = .017 as it is less than the 0.05 level of significance). However, one must be cautious about this result since the sample size (n = 18) is very small.

FIGURE 8.5



Association between PWA group and the main types of errors

Although the types of errors have been, overall, associated with particular groups of PWA, it is necessary to point out that:

- each level of breakdown (e.g. phonological output lexicon) also consists of individuals who produce other types of errors (e.g. semantic errors only) apart from the main type of errors (semantic and phonological errors),
- similar types of errors (e.g. semantic errors) tend to appear in all the levels of breakdown (either alone or in combination with other errors), and
- some exceptions may appear in the data as a result of individual differences (such as use of language, as in A12 who produces semantic errors despite being classified as having a phonological assembly deficit).

This means that error profiles alone may be an indicator in determining level of breakdown, but may not be conclusive. The results of this chapter will be discussed in the next chapter. Conclusions and recommendations for the use of the reduced set of the BNT on Maltese speakers will follow.

Chapter 9

Discussion 2

People with aphasia

9.0 INTRODUCTION

This research focused on picture naming performance in Maltese-speaking individuals. The first part reported picture naming accuracy and latency in healthy Maltese adults. The Boston Naming Test (BNT) was used to examine the effects of demographic and word variables on naming performance, and a short form of the BNT was proposed to suit the linguistic and cultural characteristics of Maltese adults. The results were reported in Chapter 4 and Chapter 5. The second part of the research focused on the performance of people with aphasia on the BNT with the aim of determining whether the proposed reduced set is an appropriate tool for assessing Maltese adults with naming difficulties. Chapter 8 presented the results of the second part of the research, while focusing on some important questions regarding the suitability of the test. This chapter will discuss the results related to the questions targeted in Chapter 8 together with reflections regarding the implications of these results.

9.1 SCORES

The scores obtained by PWA were compared to the scores obtained by the PwithoutA and conclusions regarding the most appropriate set of the BNT were reached. Complete and reduced sets of the BNT were compared and methods of scoring were investigated.

9.1.1 Complete and reduced set scores

The scores of the complete set were compared to the scores of the reduced set. A very small significant difference was found between the scores obtained by the PWA on these two sets. Although a significant difference was also found between scores of the PwithoutA on the BNT sets, the difference was much smaller in the PWA. The PWA were less able to name the items which were excluded from the reduced set, hence the small difference between scores of the sets. This leads to the conclusions that both sets provide similar information, and that the reduced set has the potential of producing the same or very similar results as the complete set, especially in PWA. The benefits of using a reduced set have already been highlighted in Chapter 6. Advantages of a short naming test include (i) retaining items which are more linguistically and culturally familiar than the rest of the items of the BNT, (ii) using a set of items which have over 70% name agreement, (iii) catering for individuals with possible attention problems (e.g. ageing individuals and patients with head injuries), (iv) limiting the chance of fatigue during the test administration, and (v) being time effective, while still being positively correlated with the complete set. Therefore, the proposed reduced set of 38 items may be said to be the most appropriate set for assessing Maltese adults with aphasia.

9.1.2 Strict and lenient scoring methods

The research on Maltese PwithoutA showed that lenient scoring is the preferred method for scoring the reduced set of the BNT. Despite using moderately high name agreement items only (70%) in the reduced set, there were still a number of items with alternative responses. These alternative responses included the English (or Maltese) translation equivalent for each item (when English (or Maltese) was not the target word), and other acceptable responses for the same item.

Analysis of the data obtained from PWA showed that strict and lenient scoring methods were positively correlated with each other, and that a small, but significant, difference between the two scoring methods was evident. The difference between the scores obtained when using strict and lenient scoring was very small in both PWA and PwithoutA. Although significant, the difference between the two types of scores obtained by the PWA (1.3%) was very small, implying that a lenient scoring method has the potential of offering very similar information to the strict method of scoring, and may be suggested as the preferred method for scoring the proposed reduced set of the BNT on Maltese speakers.

Since the test being proposed here was not originally designed and standardized on Maltese speakers, a number of items of the reduced set had alternative responses to the target name which had been determined from the dominant name produced during the first study on PwithoutA. These alternative responses or synonyms (e.g. synonym 'moxt' for target name 'pettne' (*comb*)) cannot be regarded as incorrect responses and must be given credit when scoring the BNT[1]. This method has been utilised and recommended by other researchers who investigated the use of the BNT on populations which differ from the original standardization population (Worrall et al 1995, Cruice et al 2000, Patricacou et al 2007). Nicholas et al (1989) suggested changes to scoring and administration of the BNT and they also emphasise the need to consider synonyms (e.g. *mouth organ* for *harmonica*) or commonly used responses (e.g. *lock* for *latch*) as correct.

A final reason for preferring the use of a lenient scoring method is the possibility of producing English responses when the target word is in Maltese, and vice versa. Since Malta is a largely bilingual country, with most of the population being exposed to at least two languages daily, both Maltese and English responses are expected during a naming test. This was evident during both parts of the research. Lenient scoring allows the tester to give credit to both English and Maltese translation equivalents. Kohnert et al (1998), in their study on the BNT on a bilingual population, also recommended this alternative method of scoring in order to give credit to responses from both languages. Therefore, the proposed reduced set of the BNT with lenient scoring is being presented as the preferred and most appropriate version for assessing naming in Maltese PWA.

9.2 DIFFERENTIATING BETWEEN INDIVIDUALS WITH IMPAIRED AND UNIMPAIRED NAMING

9.2.1 Scores

A significant difference was found between the mean scores obtained by the PWA and the PwithoutA on both sets of the BNT and when using both methods of scoring. This finding was promising, giving an indication of the test's ability to differentiate between impaired and unimpaired naming. However, a closer look at the scores showed a good degree of overlap between the scores obtained by the controls and the PWA.

[1] In the reduced set there are 5 items with synonyms; nearly all items have possible English/Maltese translation equivalents; only 6 items have just one response (e.g. helicopter); there are 15 cognates.

The PWA obtained lenient scores ranging from 3 to 28 (out of 38) on the reduced set and the respondents of the main study also obtained a wide variation of lenient scores, ranging from 15 to 38 on the same set. It became evident that nine individuals without aphasia were not simply outliers in the data as all of their scores were below the confidence interval for their age and education level and their scores tended to overlap with the scores obtained by the PWA. The score obtained on the reduced set of the BNT is, therefore, not enough to differentiate between impaired and unimpaired individuals as there is not a clear line between those who have been diagnosed with aphasia and those who are said to be neurologically unimpaired. Rather, there appears to be a graceful decline in naming accuracy even in the PwithoutA, indicating that score alone may not be a very good indicator of naming ability, especially in older people.

The *mean scores* obtained by other studies on different populations have already been reported in Chapter 6. In general, it was concluded that the performance of the Maltese PwithoutA was mostly similar to that of the healthy Greek sample (Patricacou et al 2007), possibly as a result of the geographical and cultural similarities. When the Maltese results were compared to other bilingual studies (e.g. Kohnert et al 1998; Roberts et al 2002), similarities in the mean scores also became apparent when samples were matched for age and education.

The researcher also attempted to investigate and compare the range of scores obtained by other studies on different populations to the present data on PwithoutA, in order to determine whether low BNT scores were also obtained by other studies on the BNT. Complete sets will be compared since it is not always possible to determine the range of scores obtained on the reduced set of other studies. Table 9.1 presents the mean scores, standard deviations as well as the range of scores obtained in different studies, including the data from the present study. Lenient scores are presented in all studies, except for the Dutch-speaking Belgian study. The English monolingual, the French/English bilingual and the Spanish/English bilingual data are derived from the same study (Roberts et al 2002) which was based in Canada.

It is apparent that the three bilingual studies reported here have similar mean scores and range of scores. The lowest scores obtained are 20, 21 and 26 by the Maltese/English bilinguals, the Spanish/English bilinguals and the French/English bilinguals respectively, while the monolingual Dutch-speaking Belgian and the Canadian English participants obtained lowest scores of 36 and 45 respectively. Although it is difficult to generalize from such small amounts of data, the studies presented here may indicate a tendency for bilinguals to obtain lower mean scores and for some bilinguals to obtain very low scores, as was found in the present study on PwithoutA. Yet, the Greek study also reported lower mean scores and a wide range of scores despite the fact that the

TABLE 9.1

	Dutch-speaking Belgian[2]	English mono-lingual[3]	Greek sample[4]	French/ English bilingual[5]	Spanish/ English bilingual[6]	Maltese/ English bilingual[7]
Mean score	51.9	53.9	42.8	41.4	43.9	39.2
SD	5.52	3.59	9.76	7.76	8.12	9.72
Mininum score	36	45	15	26	21	20
Maximum score	59	60	58	58	56	57

Scores obtained by PwithoutA on the complete set

Greek participants were all monolinguals. The results of the Greek study have been attributed to the socio-cultural characteristics of the Greek population (Patricacou et al 2007). Thus, the studies presented here suggest that low BNT scores have also been observed in other populations and are not only witnessed in Maltese speakers. Unfortunately, the above studies do not report scores of PWA from the same population and, therefore, it is difficult to say whether overlap between scores of PWA and PwithoutA is evident in these populations too.

To the researcher's knowledge there is no literature about BNT studies that compare and report the range of scores obtained by PWA and PwithoutA and that investigate the overlap in the range of scores seen in the samples of the present study. The studies generally report and compare the means and standard deviations only (e.g. Elman et al

[2] Mariën et al (1998) N = 200; age range: 55 – 91 years; educ range: 2 - 22 years.
[3] Roberts et al (2002) N = 42 Canadian monolinguals; age range: 20 – 52, educ range: 11 - 24 years.
[4] Patricacou et al (2007). N = 100; age range: 20 – 71+ years; educ range: 0 – 13+ years. Monolingual participants (confirmed by Patricacou (Oct 2010) through emails to the researcher).
[5] Roberts et al (2002) N = 49; age range: 25 – 55; educ range: 11 - 22 years.
[6] Roberts et al (2002) N = 32; age range: 29 – 54; educ range: 11-27 years.
[7] This research on PwithoutA. N = 60; age range: 20 – 85; educ range: 0 – 13+ years.

1992) showing that there is a significant difference between the mean scores, without referring to the presence or absence of overlap in scores. Since the lowest scores obtained by the healthy participants (15 - 22) of this study were below 1 and 2 standard deviations from the mean, it is evident that score alone does not always indicate naming impairment. It is necessary to look at different sources of information about naming performance, in order to make accurate identification of individuals with naming difficulties. The errors made during naming may, therefore, provide more information regarding differentiation between impaired and unimpaired naming.

9.2.2 Error profile

We have seen that individuals may obtain very low scores on naming tests in the absence of aphasia. On the other hand, PWA may obtain adequate scores, in spite of an error profile that is suggestive of a naming impairment (Storms et al 2004; Mansur et al 2002). The usefulness of examining error profiles will be investigated here.

9.2.2.1 Common and atypical errors

In addition to the nine types of errors (*common errors*) produced by PwithoutA, nine different types of errors were also produced by the PWA (see Table 8.8 in Chapter 8). These nine *atypical* errors, may be a salient difference between the error profiles of the unimpaired and the impaired participants of this research. This result may, therefore, be used to differentiate between individuals as the presence of any of these atypical errors would signal a risk of naming impairment.

Although *common errors* were more frequently produced by PWA and a significant difference was found between the mean number of errors produced by PWA and PwithoutA, some of the participants who had a milder word finding difficulty, such as participant A01, produced only a few errors. Participant A01 produced a total of 10 errors, of which four responses were commonly produced by PwithoutA (such as ‘horse’ for ‘unicorn’ and circumlocution for ‘stethoscope’). The other 6 errors were atypical errors and consisted of unrelated errors, perseveration and long hesitations. Therefore, rather than focusing on number of errors or number of error types, it is more useful to focus on the type of errors produced and to investigate whether atypical errors are present in an individual’s error profile. Despite the varied levels of severity (in terms of score) of the PWA of the present study, all of the 18 participants produced atypical errors, showing that they could be differentiated from PwithoutA, as a result of their deviant pattern of responses. Therefore, the number of particular types of errors may not be used to distinguish between people with impaired and unimpaired error profiles. Rather, it is the presence or absence of atypical errors that is the most determining factor in diagnosing an individual with a naming impairment. The following table lists the common and atypical errors produced during the current research.

TABLE 9.2

Common Errors	Atypical Errors
Coordinate	Multiword (MW)
Superordinate	Morphological alteration
Subordinate	Perseveration
Semantic association (S Assoc)	Unrelated
Circumlocution	Forgot
Phonologically related real word (PhRR)	Phonologically related non word (PhRN)
Don't know	Successful conduite d'approche
Visual Association/not recognized	Unsuccessful conduite d'approche
Self correction	Neologism

Common and atypical errors produced during the study

9.2.2.2 Semantic errors in PwithoutA

The PwithoutA who participated in this research produced semantic errors, no responses, visual associations and phonological errors, in this order of frequency. This is consistent with the findings of another study on healthy participants. In her study on word finding in normal ageing, Strauss Hough (2007) found that semantic errors were the most frequently produced picture naming error, followed by no responses, and visual and phonological errors. The PWA of the current study also produced these types of errors. However, *no responses* were the most frequent type of errors, followed by semantic errors, visual associations and phonological errors (see Table 9.3).

Tallberg (2005) and Mariën et al (1998) studied the type of errors produced by a sample of 111 Swedish and 200 Dutch-speaking Belgian healthy adults (on the BNT) respectively and also found that most of the errors were semantic in nature. None of the atypical errors listed in Table 9.2 appeared in the types of errors identified by Tallberg in her sample. On the other hand, the Belgian sample produced very few atypical errors, including perseverations (1 response or 0.06% of all errors produced), neologisms (3 responses or 0.19%) and empty words (1 response or 0.06%). The production of semantic errors in healthy samples has been especially reported in older people (Nicholas et al 1985) and

TABLE 9.3

	PWA	PwithoutA
Mean number of semantic errors[8]	6.56	3.20
Mean number of phonological errors[9]	2.17	0.18
Mean number of visual associations[10]	2.44	1.68
Mean number of no responses[11]	7.94	1.93

Mean number of errors produced by PWA and PwithoutA

Barresi et al (2000) also conclude that naming errors may be, in part, a result of a "semantic degradation" that occurs in healthy ageing. Albert et al (1988) also showed that semantic, as well as perceptual, errors tend to increase with age, but not lexical errors (phonologically related real words and phonologically related nonwords). The present study supports Mariën et al (1998)'s conclusions that both scores and types of errors should be investigated when analyzing naming ability, and that the presence of atypical, or "exceptional error types" (p. 462) indicates a high risk of impaired naming ability. The current research goes one step further in stating that error types are the most important indicator of naming impairment.

9.2.2.3 Naming errors in PWA

Chapter 2 reported the types of errors that are typically produced by people with aphasia, including semantic errors, phonologically related real word errors, mixed semantic and phonological errors, visual errors, neologisms, unrelated errors, no-response errors, circumlocutions and other errors (Cuetos et al 2002). These errors have

[8] Semantic errors: coordinates, superordinates, subordinates, semantic associations and circumlocutions.
[9] Phonological errors include: phonologically related nonwords and real words, and conduite d'approche.
[10] Visual associations: visually related associations and instances when person admits that item is not recognized.
[11] No responses: don't know, forgot and instances when phonemic cues are given.

all been found in the present data. However, mixed errors were only minimally produced by the Maltese PWA.

Budd et al (2010) found that coordinate semantic errors were by far the most common type of errors in all their subjects (with primary progressive aphasia and acute post stroke aphasia). Although the present study identified semantic errors as very common types of errors in both PWA and PwithoutA, coordinate errors were not the most frequent errors produced. Circumlocutions, visual errors, don't know responses and semantic associations were more frequently produced than coordinate errors by the PWA of the present study.

Dell et al (1997) tested their interactive model of lexical access and concluded that naming impairments in aphasia reflect normal naming performance. They stated that aphasic naming may be explained by slight quantitative changes to the theory of normal performance and the error patterns of people with aphasia are simply an exaggeration of the error patterns observed in healthy individuals. Silkes et al (2004) also reached the latter conclusions when a group of controls were tested under temporal stress, stating that there is an aphasic-to-normal continuum of errors supporting the "aphasia-normalcy continuum hypothesis" (p.615). The results of the current study tend to disagree with the latter views. The errors produced by the PWA consist of a mixture of the nine commonly produced errors of PwithoutA, as well as nine other atypical errors which were produced *only* by PWA. This behaviour leads to the conclusion that the PWA of the current study are not simply producing an exaggerated amount of normal/common errors, but that they are behaving in very particular ways. Ruml et al (2005) explored error data from Italian PWA and refuted Dell's (continuity) thesis. They agree with Rapp and Goldrick (2000) that Dell et al (1997)'s theoretical model is too limited. The current data also appears to be in conflict with Dell et al's proposal since all the participants with aphasia (including those participants who had milder word finding difficulties), showed evidence of atypical errors in their error profiles, and not just an exaggeration of the common errors. This present work tends to agree with that of Ruml et al (2005) who prefer a dissociative thesis that allows for "a complex mixture of individual deficits to co-occur", and also for individual deficits to "occur in isolation"

(Ruml et al 2005, p. 158). This may serve as a better framework to account for the errors produced by the PWA of this study.

9.3 DIFFERENTIAL DIAGNOSIS

The final section of Chapter 8 looked into the capability of the reduced set of the BNT to identify the level of breakdown in spoken word production. Three sources of evidence were utilized for this purpose: effect of word variables, performance on additional tasks and the nature of the errors produced during picture naming (Whitworth et al 2005).

9.3.1 Effect of word variables

Regression analysis showed that the same effects of AoA and syllable length emerged in both PWA and PwithoutA as these two word variables remained significant predictors of naming performance. Other word variables (such as familiarity and imageability) were excluded from the regression model. These results replicate the findings of Nickels and Howard (1995) who used two different samples of PWA and 2 different sets of pictures. They found that, although different patterns emerged, AoA and (phoneme) length affected naming accuracy in both studies. It is possible that word length is influenced by cross-linguistic variation. While some studies have emphasised that it is phoneme length that affects naming, the present study on healthy Maltese and on Maltese people with aphasia identified syllable length as one of the significant predictors of naming accuracy. This may be explained by the fact that Maltese has a complex syllable structure (Comrie 2009) which has the potential of influencing naming, rather than phonemes,

Variable effects, in the current study, were not particularly predictive of level of breakdown. Logistic regression showed that the effects of AoA and length may not be reliable after all. Syllable effects appeared in only 2 participants and these length effects were not associated with phonological assembly deficits as expected. On the other hand, most of the PWA experienced the effects of AoA. It is possible that the particular material used, the number of items (38) and the small sample of 18 participants may be possible reasons for these results. A bigger set of items and a bigger sample might have produced bigger word variable effects.

9.3.2 Predictors of degree of anomia

The additional tasks of comprehension and repetition were initially used to identify a possible level of breakdown for each participant. Four main groups of PWA became clear when their performance on the additional tasks was analyzed. These groups consisted of a lexical semantic group, a phonological output lexicon group, a phonological assembly group and a severe group. The following sections intend to investigate whether the additional tasks of comprehension and repetition could be used to predict the degree (this section) and the nature of anomia (section 9.3.3).

Lambon Ralph et al (2002) used a set of semantic and phonological measures of impairment in their case series study on 21 PWA. They showed that the degree and nature of anomia can be predicted from the set of measures of semantic and phonological impairment. Martin et al (2006) also used a "multi-measurement approach" and found that, as a whole, their 6 semantic and 5 phonological input measures were good predictors of picture naming performance. The present study is very limited in that it only tested semantics and phonology by means of single tasks of spoken word-picture matching, and word and nonword repetition tasks respectively. Nonetheless, the information available from the 18 PWA of the current study will be used to establish whether the results of the additional tasks may predict the degree and nature of naming impairment.

In the present study, word and nonword repetition were positively correlated with each other, but comprehension was not found to be correlated with any of the repetition tasks. These results were also found by Lambon Ralph et al (2002). Positive correlations emerged between naming accuracy and comprehension, and between naming accuracy and real word repetition, but not between naming accuracy and nonword repetition (see section 8.3.2 in Chapter 8), implying that performance on additional tasks of semantics and phonology is related to the scores obtained on the naming test.

Although the additional tests of semantics and phonology were found to be related to score in the present study on PWA, the best predictor of naming accuracy was fluency naming. This result emerged following a stepwise multiple regression analysis. This analysis showed that fluency naming (the number of animal names produced in one minute) was the best predictor of naming accuracy on the reduced set of the BNT, and excluded the predictor variables comprehension and repetition. This implies that in this study on Maltese PWA degree of anomia may be more reliably predicted by the performance on a simple, fluency naming task, rather than by the comprehension and repetition tasks used in the current study.

Lambon Ralph et al (2002) used a principal components analysis and concluded that measures of semantic and phonological ability can accurately predict picture naming performance, but fluency naming was not included in their analysis. During the study on the Maltese PwithoutA, fluency naming data was not collected. Therefore, it is not possible to say whether fluency naming is also an important predictor of naming accuracy in unimpaired individuals. This data is not available from the Maltese authors of the LEST either. Other studies, such as that by Schmitter-Edgecombe et al (2000), on word finding in healthy ageing, reported weak positive correlations between fluency naming and BNT scores. On the other hand, a moderate correlation between fluency and naming has been found in other studies on PWA (Thompson and Heaton 1989 cited in Budd 2007). Kim et al (2011) also showed that fluency naming can discriminate between PWA and PwithoutA even when using a 30 second time span.

9.3.3 Predictors of the nature of anomia

Chapter 8 showed that, in general, the *main* types of errors produced were associated with the particular level of breakdown identified by the additional tasks (see Figure 8.5 and Table 8.24 in Chapter 8), such that PWA who were impaired on the comprehension test (lexical semantic group) tended to produce semantic errors and no responses. PWA who failed the repetition test (the phonological assembly group) tended to produce phonological errors. On the other hand, PWA who were not impaired on any of the additional tasks (phonological output lexicon group) mainly produced semantic and phonological errors. It is necessary to determine whether the additional measures of comprehension and repetition may really predict the type (nature) of errors produced.

Overall, as suggested by Bormann et al (2008), it is possible that the PWA who failed the comprehension test (the lexical semantic group 1) had difficulties with selecting and activating the specific item from the mental lexicon as a result of semantic competition. This resulted in semantic errors. On the other hand, it may be possible that when semantic competition was weaker, as a result of a smaller number of semantically related items, omissions, or no responses, tended to occur.

When individuals had difficulties with repetition (the phonological assembly group 3), it may be assumed that they experienced difficulties at the segmental level of spoken word production. Phonological errors were possibly a result of difficulties in selection and ordering of phonemes once the phonological word form had been correctly retrieved (e.g. Goldrick and Rapp 2002).

On the other hand, the group of people who had adequate comprehension and repetition skills (as assessed on the present additional tasks) still experienced semantic and phonological errors (group 2). As explained by Cox (2008), phonological errors possibly occurred as a result of incomplete availability of the word forms, while semantic errors possibly occurred as a result of impaired connections between lexical semantic and lexical phonological levels. Seven out of the eighteen subjects who participated in this part of the research were categorized into this group, offering support for the existence of a phonological output lexicon (POL) level. Similarly Caramazza and Hillis (1990) reported findings on two people with aphasia who produced semantic errors despite having adequate comprehension and better written than spoken naming. They argued that the errors produced emerged from an impaired POL. This is consistent with the findings made in the current study as 39% of the PWA were said to have a POL deficit.

9.3.3.1 Additional measures and types of errors

Using the types of errors as dependent variables and the results of the additional tests as predictor variables it was possible to test whether additional tests of semantics and phonology have the potential of predicting the nature of the naming impairment. Simple linear correlations were carried out between predictor variables comprehension of spoken words, repetition of words and fluency naming, and dependent variable number of semantic errors. There was no evidence of a significant correlation between any of these variables, not even between comprehension and semantic errors, showing that the present data did not show a relationship between the semantic measure (or any other measure) and the number of semantic errors produced.

Similar linear correlations were carried out using the same predictor variables and the number of phonological errors as dependent variable. A strong positive correlation became evident between fluency naming and the number of phonological errors produced ($r = .705$, $p < .01$). The phonological measures (nonword and real word repetition) were not found to be related to the number of phonological errors produced by the PWA of the current study. This implies that there is no relationship between the phonological measure and the number of phonological errors produced in this study. Therefore, the additional phonological and semantic measures/tests used in the current study cannot be used to predict the type or number of errors. The fluency naming test was the only measure found to be related to phonological errors.

The other types of errors (visual associations and no responses) were also investigated to identify a possible relationship between measures of semantic and phonological integrity and the types of errors. Once again, the only significant correlation that emerged was between the number of no responses and fluency naming ($r = -.519$, $p < .05$), implying that performance on fluency naming is negatively associated with the number of no responses. Therefore, the current data shows that poor performance on fluency naming predicts (i) worse anomia (in terms of score or number of correct responses), (ii) a greater number of phonological errors, and (iii) a larger number of no responses on the proposed reduced set of the BNT.

The above results are not consistent with the findings presented by Lambon Ralph et al (2002) and Martin et al (2006). Despite the fact that the main types of errors were somewhat associated with the level of breakdown that was identified through the performance on the additional tests, a direct relationship between the semantic test and the number of semantic errors produced was not evident, and similarly between the phonological test of repetition and the number of phonological errors produced. Therefore, in the present study, the nature of anomia cannot be directly predicted from the performance on other tests of semantics and phonology. It is worth noting that single spoken word-picture matching tasks may not always identify individuals with semantic deficits as the semantics involved in spoken word production is more complex than that which is assessed through word-picture matching (Lambon Ralph et al 2000; Nickels 2004; Howard and Gatehouse 2006). It is unclear whether the administration of other tests of semantic and phonological integrity in addition to the tasks used here would have produced different results. Martin et al (2006) insist that multiple semantic and phonological measurements should be used to assess PWA to determine whether naming performance may be predicted from these measures. Because each measure differs in level of difficulty, together they (the set of measures) "succeed in capturing variation at both the severe and mild ends of the spectrum" (p.164). Thus, a multi-measurement approach might have resulted in different findings in the present study.

In this research, fluency naming has been shown to play a more important role in predicting naming scores and type of errors. Unfortunately, normative data on category fluency (animal names, etc) is not yet available for Maltese speakers. Since fluency naming has been shown to play an important role in predicting the degree of anomia and in the production of phonological errors and no responses, it is recommended that such normative data should be obtained. The norms available from other studies on different populations cannot be used for Maltese speakers as different populations differ in linguistic and cultural characteristics. This need is highlighted by Elkadi et al (2006) who found that category fluency norms for Australian women differed from those of other populations.

9.3.3.2 Level of breakdown and *main* types of errors

The main (or most common) type of errors produced by each person with aphasia was somewhat associated with particular levels of breakdown in spoken word production (see Table 8.24 in Chapter 8). Three PWA of the lexical semantic group mainly produced *semantic errors and no responses*, five of the POL group produced *semantic and phonological errors*, two of the PA group produced *phonological errors* and all of the severe group produced *semantic errors and no responses*. Yet, other individuals in each group produced other types of errors (e.g. one person with POL breakdown produced *no responses* only). Therefore, although most of the individuals in each group produced similar errors, some individuals made different errors, implying that different individuals having the same level of breakdown may produce different types of errors. Thus, for example, different people in the lexical semantic group of the current study produced *semantic errors*; *semantic and phonological errors*; and *semantic errors and no responses*. This variability in performance within the same group may be the result of the variation in the severity of breakdown. Butterworth et al (1984, cited in Howard and Gatehouse 2006) give this explanation for the different types of errors that occur within the same level of breakdown. In fact, the three (of five) individuals who produced *semantic errors and no responses* obtained lower lenient scores on the reduced BNT and did worse on the comprehension test than the two individuals who produced *semantic errors* only and *semantic and phonological errors*.

On the other hand, different individuals who produce similar errors were found to have different levels of breakdown. In fact, *semantic errors* and *no responses* appear across all the different groups of the PWA in the current study, either as the main type of error or else in combination with other types of errors. In a review of the literature, Howard and Gatehouse (2006) show that it is often difficult to determine the level of breakdown from error analysis alone, since *semantic errors* and *no responses* tend to occur at every level of impairment, as was evident in the current study. Gonzáles Nosti et al (2008) categorized semantic errors in naming into three subtypes and concluded that different types of semantic errors arise from different levels of breakdown. They found that superordinate errors arise from the semantic level, semantic associations arise from the lexical level and coordinates arise from both levels of processing. An attempt to apply this finding to the current data proved futile as most of the PWA had a mixture of types of semantic errors. The presence of semantic errors in all levels of breakdown and the occurrence of a mixture of types of errors in different levels supports the idea of interactivity in spoken word production. The present data is not consistent with discrete models (such as Levelt's), which hold that semantic errors are more likely to occur at the semantic level. Rather, this research tends to support interactive models (such as Dell et al 1997) and the restricted interaction model (RIA of Goldrick and Rapp 2002), which propose that semantic errors may appear at any level of breakdown as a result of the strong competition imposed by semantic neighbours (Cox 2008).

Although each participant produced main types of errors associated with particular levels of breakdown, multiple errors types were also produced by the same individual. This provides further support for a degree of interactivity between levels of word production. Goldrick and Rapp (2002) use their RIA to account for multiple errors too. They provide evidence to show that breakdown at the lexical level and the phoneme level or a global lexical level breakdown may result in multiple error types (as well as the occurrence of more formal errors than non words). If the lexical level is impaired it confuses the subsequent level and, therefore, indirectly causes breakdown at the phoneme level too. Goldrick and Rapp explain that mixed lexical and phoneme level damage also results in a variety of errors. Since there is feedback from the phoneme level to the lexical level in their RIA, disruption of phonological information may also confuse the lexical level. Therefore, the present data appear to be consistent with interactivity between stages of spoken word production.

Each type of error may arise from different levels of breakdown (Caramazza and Hillis 1990; Nickels and Howard 1995; 2004), making it nearly impossible to reach conclusions regarding the source of impairment from error profiles alone. It is evident that differential diagnosis entails more than just looking at particular errors. Rather, it is necessary to look at the results of additional tasks as well as the main types or combination of errors produced. As seen above, variable effects, in this study, were not particularly predictive of level of breakdown since most PWA experienced age of acquisition effects and syllable length effects were not associated with phonological assembly deficits as expected. A multi measurement approach, as recommended by Martin et al (2006), to analyze phonological and semantic integrity by means of a variety of input and output measures, investigation of effect of miscues (Howard and Gatehouse 2006), variable effects (if possible), and error profiles should all be used to obtain a definite differential diagnosis. Therefore, the reduced set of the BNT should not be used on its own to make a differential diagnosis. Rather, together with other sources of evidence, an error profile based on this modified BNT contributes to identifying level of breakdown.

9.3.3.3 Exceptions and individual differences

Different PWA having the same level of breakdown in spoken word production are not a homogenous group of people (Nickels and Howard 1995). Each individual's error profile is a reflection of the highly complex interaction between the site and time post onset of neurological impairment, the reorganization of functions as in chronic compensation for difficulties, and the level of breakdown that may be impaired (Budd et al 2010). Each individual also brings with him/her a range of personal characteristics, including age, education, language, pre-morbid abilities and existing co-morbidities. Thus, this combination of different features inherent in each individual may be one of the reasons for the lack of direct association between one type of error and level of breakdown.

A particular exception in the current data is the error profile of participant A12. Her performance on the additional tasks showed that she had difficulties with repetition and adequate comprehension of single words. Although she appeared to have difficulties at the phonological assembly level, she tended to produce a number of semantic errors apart from phonological errors. On the reduced BNT, participant A12 produced most of her responses in English (63.2%) and it seemed that she was overall using English to name the pictures. Although she regarded herself as a Maltese speaker, she reported that she read everyday in English and spoke English often with her grandchildren. Nearly all (5 out of 7) of the semantic errors produced by A12 were in Maltese. Only two phonologically related non words were produced and these were phonologically related to the English equivalent (e.g. "pom" for *comb*). Therefore, it is possible that in her effort to avoid making further phonological errors in English, A12 retrieved words in Maltese, but made semantic errors in the process.

The effect of English responses was not analyzed in detail during the present study. The main issue regarding the use of English was related to the benefit of using lenient scoring, as this alternative method of scoring gave credit to both English and Maltese responses. Only two participants (A11 and A12) showed a large discrepancy between their strict and lenient scores. They both obtained much better lenient scores as they produced many English responses (44.4% and 63.2% respectively). Both of these participants were females, over 70 years of age and with less than 12 years of education.

However, they both read everyday in English and used English often with their grandchildren.

9.4 CONCLUSION

This chapter has attempted to discuss the results obtained from the study on 18 PWA. The reduced set of the BNT has been shown to be suitable for assessing Maltese individuals with suspected naming impairments and lenient scoring is the preferred method of scoring for this population.

The suitability of the test to differentiate between impaired and unimpaired individuals has been investigated. Although the mean scores obtained by the PWA and the PwithoutA were shown to be significantly different from each other, a degree of overlap in scores became apparent. This was caused by 9 unimpaired individuals who obtained very low scores and whose scores were similar or worse than the scores obtained by some PWA. For this reason, error profiles were suggested as better indicators of naming impairment and the presence of atypical errors to be the most important tool for identifying impaired individuals.

Differential diagnosis was also investigated to determine whether the test may help in identifying the level of breakdown in spoken word production. Lambon Ralph et al (2002)'s findings, that measures of semantic and phonological integrity may predict the degree and nature of anomia, were applied to the current study. It was concluded that, in the current study, fluency naming is the best predictor of naming accuracy, of the number of phonological errors and the number of no responses. The findings were not consistent with Lambon Ralph et al (2002)'s conclusions that comprehension and repetition tasks predict the degree and nature of anomia. It became evident that, while some people who failed the comprehension test produced semantic errors and some people who failed the repetition test produced phonological errors, other people produced semantic and phonological errors despite adequate performance on the tests of comprehension and repetition. This finding supports the idea of a Phonological output lexicon (POL) in spoken word production. Breakdown at this level is demonstrated by adequate performance on semantic and phonological measures in spite of semantic and phonological errors in naming.

Finally, it was concluded that types of errors alone cannot be used to make differential diagnosis, as similar errors occur in different levels of breakdown and different types of errors appear in the same level of breakdown. The mixture of error types that appeared in the same level of breakdown is consistent with interactive theories of spoken word production as different levels interact, provide feedback to and influence each other to produce different types of errors within the same level of breakdown. Thus, as shown by various researchers (Whitworth et al 2005; Howard and Gatehouse 2006; Budd et al 2010), differential diagnosis requires a multi measurement approach, including a variety of sources of evidence, ranging from error profiles, variable effects, performance on additional tests, and response to phonemic cues and miscues.

Chapter 10

Conclusion and recommendations

10.0 INTRODUCTION

This study investigated naming performance of Maltese speakers on the Boston Naming Test. During the first part of the research, data was collected from 60 healthy individuals. The study focused on the influence of demographic variables and word variables on naming accuracy, naming latency and the number of English responses produced. A modified BNT was proposed following the results obtained. The second part of the research, then, collected data from 18 people with aphasia and aimed to determine the suitability of the proposed test for assessing Maltese people at risk of naming impairment.

10.1 VARIABLES AFFECTING NAMING PERFORMANCE

The research on healthy participants found that the level of education was the strongest predictor of naming accuracy and the number of English responses produced. Age also played an important role as it emerged as the strongest predictor of reaction time with evidence of a significant decline in both naming accuracy and naming latency over 66 years of age.

Of all the word variables investigated, age of acquisition (AoA) was found to be the most significant predictor of both naming accuracy and naming latency. Finally, an interesting interaction between word variable, AoA, and demographic variables, education and age, became evident. While age emerged as the most significant predictor of naming accuracy on *early* AoA words, years of education was found to be the most significant predictor of naming accuracy on *late* AoA words.

A modified BNT for Maltese speakers was proposed according to the findings of this part of the study. A reduced set of 38 items and an alternative method of scoring (lenient scoring) were suggested. The reduced set retained only those items which had at least 70% name agreement. Normative data stratified by age and education was presented in Chapter 6.

10.2 IDENTIFICATION OF INDIVIDUALS WITH NAMING IMPAIRMENT

The second part of the research focused on the performance of Maltese PWA on the proposed reduced set of the BNT. The suitability of the proposed test for assessing Maltese speakers at risk of naming impairments was analyzed. Analysis during this part of the study aimed at determining whether the proposed test is able to differentiate between impaired and unimpaired individuals, and whether the test contributes to differential diagnosis, to identify level of breakdown in impaired individuals. Additional tests of comprehension and repetition were administered on the PWA to acquire further information about semantic and phonological abilities.

Mean scores of PWA and PwithoutA (controls) were found to be significantly different. However, a number of older controls obtained very low scores, such that their scores overlapped with those of the PWA. This overlap in the range of scores of the two samples indicated that scores alone do not always provide sufficient information to differentiate between impaired and unimpaired individuals. Other studies which compare BNT scores of people with aphasia and controls only report and compare means and standard deviations, and fail to compare the range of naming scores obtained by the two samples. In the present study, comparison of the *range* of scores, and not comparison of *means,* indicated that scoring alone is not enough to diagnose word retrieval difficulties.

It was concluded that the main differentiating factor between impaired and unimpaired individuals is the error profile. People with aphasia were found to produce atypical errors, which were not apparent in any of the response patterns of the PwithoutA. While common errors were produced by both samples, the atypical errors were produced only by the PWA. Therefore, error profiles should always be drawn for each individual who is tested on the modified BNT. The presence of atypical errors in one's error profile should serve as an important marker of naming impairment.

Therefore, this study has shown that, apart from comparing scores to the normative data, it is necessary to analyse error profiles as the occurrence of atypical errors is a determining factor in diagnosing a person with a naming impairment.

10.3 FLUENCY NAMING

The data obtained from PWA was used to identify correlations between additional tests of semantics and phonology and the BNT. BNT scores were found to be positively correlated with fluency naming, with auditory comprehension of single words, and with real word repetition. Fluency naming emerged as the most important predictor of naming accuracy on the modified BNT.

While scores were correlated with the results of the additional tests of comprehension and repetition, correlation analysis failed to identify any relationship between particular types of errors and the additional tests. Contrary to the findings of other studies (e.g. Lambon Ralph et al (2002) and Martin et al (2006)), the present research failed to find a relationship between tests of semantic integrity and semantic errors, and also between tests of phonological integrity and phonological errors. However, it should be noted that only one test of comprehension of spoken words and one test of repetition of words and nonwords were used as additional tests during the current study. A multi measurement approach, using a range of additional tests including analysis of written text, investigation of person's ability to access semantic and conceptual information, etc., as recommended by the above mentioned researchers, might have produced different results.

Fluency naming, once again, emerged as an important predictor as it was found to be related to two types of errors. Moderately strong correlations appeared between fluency naming and the number of *phonological errors* and the number of *no responses*.

Therefore, a simple fluency naming task might indicate the probability of a naming impairment.

10.4 DIFFERENTIAL DIAGNOSIS

Particular types of errors were noted to be associated with particular levels of breakdown. The lexical semantic level was overall related to the presence of *semantic errors and no responses*, the phonological output lexicon level was mainly associated with *semantic and phonological errors* and the phonological assembly level was generally associated with *phonological errors*.

However, a degree of overlap was also apparent here. Different types of errors were found to occur in the same level of breakdown and similar errors occurred in different levels of breakdown. Thus, it was concluded that, although error profiles may contribute valuable information to making a differential diagnosis (to determine the level of breakdown in spoken word production), on their own, they do not provide sufficient information to be conclusive. Many other sources of evidence may be required. These include additional tasks of semantic and phonological integrity, such as, but not limited to, comprehension and repetition tests, investigation of word variable effects on naming, response to cues and miscues, etc..

10.5 RECOMMENDATIONS FOR MALTESE BNT

Another important aim of the present study was to provide practical and clinical suggestions regarding the use of the BNT on Maltese speakers. These suggestions were based on the findings that emerged from the study and included recommendations regarding the material that may be used, the method of administration of the test, the method of scoring of the test and the method of interpreting the results obtained. An adapted Maltese Boston Naming Test (MBNT) booklet was created (see Chapter 11). The latter includes all the necessary information required for clinicians to administer, score and interpret the test on Maltese speakers. The reduced set of the BNT with lenient scoring is being suggested as the most suitable adaptation of this popular test for Maltese individuals.

A number of modifications to the original test were necessary to accommodate cultural and linguistic differences of different populations. The reduced test retained only those items which were found to have at least 70% naming agreement. The advantages of a shorter form of the BNT have also been highlighted. The same sequence of item presentation as the original BNT has been retained in order to keep the same sequence used during the standardization process. A lenient method of scoring is also being suggested to give credit to alternative, but acceptable responses, including synonyms and English/Maltese translation equivalents. Normative data stratified for age and education and information regarding common errors (produced by both impaired and unimpaired individuals) and atypical errors (produced by impaired individuals only) are available.

Interpretation of results obtained on the modified set entails not only comparing lenient scores to the available normative data, but also most importantly analyzing error profiles. The presence of atypical errors in one's error profile is a good indication of naming impairments.

All the necessary information about administration, scoring and interpretation of results obtained from the reduced set of the BNT are included in the Maltese Boston Naming Test (MBNT) booklet, which was created for this purpose (Chapter 11).

10.6 FURTHER RESEARCH AND LIMITATIONS

The present study provides the foundation for further research in the area of spoken word production in Maltese speakers. To the researcher's knowledge, no other work on the area preceded the present one. It would be interesting to see whether future research would replicate the findings of this project, and whether a different sample, and possibly different material, would also produce similar results.

Since the results obtained in this project are based on a cross-sectional study of individuals aged between 20 and 85 years, cohort effects may have been present. The older individuals who took part in this study undoubtedly had different experiences from the younger individuals. As a result of the socioeconomic changes that occurred in Malta over the past decades, the ageing individuals of today (especially females) had different past experiences from those of young generations, as many more opportunities for further education are available today. Therefore, in order to avoid cohort effects, a future project may take the form of a longitudinal study, spanning more than four years[1], in order to follow the naming ability of particular individuals as they age. Furthermore, it would be interesting to see if a larger sample (i.e. more than 60 people) would produce the same results with regards to the effect of demographic and word variables and also with regards to the normative data obtained here.

Since it has been found that common medical conditions (such as hypertension and Diabetes Mellitus, (see Chapter 6)) may affect cognitive abilities and lexical retrieval, it is necessary to record detailed personal data from each participant. This data should also include information regarding the cognitive and physical lifestyle of each participant. This is essential as psychogenic factors may account for the large variation in the performance of ageing individuals on cognitive tests, and may need to be analysed with regards to naming performance. It would also be useful to study the suitability of the recommended version of the BNT in the assessment of individuals with other neurological conditions, such as dementia. Considering the growing population of ageing individuals in Malta, it is advisable to invest in more research concerning older age groups.

The impact of bilingualism on naming ability in Maltese PWA was not investigated in depth during this study. More specific research on the effect of bilingualism on naming may be another potential area of research. A study of the predictor/s of the production of English responses in Maltese speakers should also produce interesting results regarding naming in bilinguals.

[1] In a four-year longitudinal study, Cruice et al (2000) reported only a weak effect of age on naming performance.

As seen earlier in the chapter, the study on PWA was limited in that it did not use a multi measurement approach. It is not comprehensive enough in terms of the variety of additional tests that were administered on the PWA. No written text was presented to test comprehension or phonology through reading and the PWA were not tested for central semantic deficits. Further, miscues were not used during the study. These would have produced important results as miscueing has been shown to lead to semantic errors when the lexical semantic level is impaired. Thus, a future study on naming in PWA should include all the relevant tests for a comprehensive approach. A similar study using a different naming test might also produce interesting results, which may or may not be similar to the present study.

Another problem might be the number of PWA participating in the second part of the research as it may be too small to generalize the findings of the present study. The numbers in each group associated with the different levels of breakdown was very small. Therefore, a similar study with a bigger sample of PWA may reach more conclusive results regarding predictors of degree and nature of anomia, as well as about differential diagnosis.

Since fluency naming emerged as such an important predictor of naming ability in the present study, it would be useful to obtain normative data on category fluency for Maltese speakers as this is not yet available.

10.7 CONCLUSION

The present study is another confirmation that neuropsychological tests should not be used across populations unless they are adapted and standardised on each particular population. Language and cultural differences undoubtedly affect naming performance since different populations may not be familiar with the items included in the original test. In fact, low mean scores may be obtained by some studies on different populations (e.g. the Maltese and Greek studies) in comparison to other populations who may be more familiar with the test items. Furthermore, it is evident that bilingual and monolingual populations require a different method of scoring on such tests. When administering a naming test on bilingual individuals it is proposed that alternative responses (including translation equivalents) should be given credit as correct responses, especially when the test being used had been originally created for a different population. On the other hand, it is necessary to note that when using this lenient scoring method, it may not be possible to decide whether alternative responses are being used as a result of word finding difficulties or not.

Normative data on the specific population must take into account the demographic variables that affect naming performance in that particular population. The norms provided in this study are stratified for age and education, since these two variables were found to be significant predictors of naming accuracy and naming latency. Word variables should also be taken into account, since the material that is used also influences naming performance.

This study has provided a Maltese version of the BNT which may be used by clinicians working with Maltese-speaking individuals with language and/or cognitive difficulties. Apart from being modified and standardised to suit the needs of the Maltese, it is also a shorter test than the original BNT, with the added benefit of being less time consuming, and more appropriate for individuals who may fatigue easily and who may have attention difficulties.

This study has filled a gap in neuropsychological tests available for Maltese speakers. The test may now be said to be culturally and linguistically appropriate for this population. Population specific items, methods of scoring and interpretation of results are all suggested to satisfy the needs of the particular population. Yet, as shown above, there are many other areas that require research. Additional sources of information such as tests of semantic and phonological integrity need to be developed specifically for this population. Fluency naming, which emerged as a very important predictor of naming ability in the current study, must also be studied to obtain normative data.

In conclusion, the study on PWA has confirmed that the proposed test should serve as a useful clinical tool to assess PWA. Lenient scores, and especially error profiles, should be used to identify individuals with naming impairments, and a variety of additional tests, together with error profiles should be used to make a differential diagnosis. The following chapter provides clear and detailed information regarding administrating, scoring and interpreting results of the Maltese BNT. Clinicians should adhere to the instructions given for the sake of consistency, and should fill the booklet with all the relevant information and results about each person tested.

Chapter 11

Recommendations for using the BNT in Maltese

11.0 INTRODUCTION

This chapter is intended to provide recommendations for using the BNT with Maltese speaking individuals. The recommendations are based on the results obtained from the present research on a group of healthy individuals and a group of people with aphasia. This chapter aims to present:

- a modified version of the BNT that is more culturally and linguistically appropriate than the original BNT;
- normative data, including Means, Standard Deviations and Confidence Intervals, stratified for age and education, for lenient scoring;
- a list of target responses and alternative responses for lenient scoring;
- a list of 9 common errors (errors produced by both PWA and PwithoutA) and 10 atypical errors (errors produced by PWA only) in order to show which type of errors occur within the normal population and which errors were produced by PWA when using the modified version of the BNT;
- additional information about age of acquisition and syllable length of each item included in the reduced set of the BNT, since these item variables were found to be the most important predictors of naming accuracy in the controls;
- a comprehensive record booklet for the modified version of the BNT in Maltese. This record booklet[1] was designed with the aim of including all the relevant information that is necessary to interpret the results of the modified BNT both quantitatively and qualitatively (see section11.2).

11.1 A MODIFIED VERSION OF THE BNT

The modified version of the BNT that is being recommended for Maltese speakers retained only those items which were found to be linguistically and culturally

[1] The record booklet may be found at the end of this chapter.

appropriate for the Maltese population. It consists of the BNT items which had at least 70% name agreement (on the dominant/target name) when administered on a sample of healthy Maltese individuals during a preliminary study. Although it was not calculated, the name agreement percentage in lenient scoring was inevitably larger. None of the original items were replaced with new items and no new items were added to the BNT items. The modified version simply removed those items which had less than 70% name agreement. This 38-item version is referred to as the reduced set of the BNT, and it has been found to be highly correlated with the complete set. The reduced set also shows the age and education effects on naming that were found in the complete set, with the added advantage of ending up with a shorter naming test which may be useful with particular populations (e.g. individuals with attention difficulties).

The items in the original BNT are graded in difficulty, from pictures which are likely to be named accurately to pictures which may be more difficult to name. This sequence was derived from the North American sample which was used during the collection of the original normative data. Since the original sequence of the BNT items was used for standardization purposes during the current study on Maltese speakers, it has been decided to retain this same original order in the modified BNT for Maltese speakers, rather than presenting a new sequence according to the frequency of correct responses produced by the PwithoutA

The modified test that is being proposed here should serve as a culturally and linguistically appropriate tool for the assessment of naming skills in Maltese speakers. It should also be an easy and practical tool to administer and should provide the clinician with clear information about the naming skills of the individual being tested in a short time. Clear, step by step instructions for test administration are included in the Record Booklet in order to use consistent instructions with anyone who is being assessed. Both quantitative and qualitative information should be recorded in order to make the best interpretation of the results. The response form has been specifically designed to fill in this information easily. The quantitative results should then be matched to the normative data and the qualitative results should serve as valuable information for a thorough interpretation of the individual's naming skills.

It must be emphasised that picture naming tests are the most frequently used tasks to assess spoken word production (Herbert et al. 2008; Goral et al. 2007). Apart from the number of correct responses, a thorough assessment of lexical retrieval should also analyse:

- the type of errors – what is the relationship between the target response and the error made? (e.g semantic, phonological, visual);
- the variables that affect naming performance (e.g. word length, imageability, age of acquisition);
- the response to semantic and phonemic cues;
- the performance on different tasks, including comprehension, repetition, reading and writing.

From Nickels (2001b) and Whitworth et al (2005).

The modified BNT that is being suggested here offers the possibility of comparing scores to normative data that is stratified for age and education, as well as identifying the presence of irregular errors and the number of common errors. Variable effects and response to cues may also be investigated and noted down in the record booklet. Other additional tasks (e.g. comprehension of spoken and written words, repetition and tests of central semantic deficits) may also be required to perform in-depth analysis of an individual's skills and to reach conclusions regarding differential diagnosis

11.2 DESCRIPTION OF THE RECORD BOOKLET

The record booklet of the Maltese modified version of the BNT includes all the relevant details that are necessary in order to be able to assess Maltese speakers' ability to name the 38 pictures of the reduced set of the BNT. The instructions and record booklet described here were based on the work produced by other researchers, including Kaplan et al (1983) and Swinburn et al (2004), as well as on the findings made during the present study. The booklet consists of:

- clear instructions that should be followed consistently by clinicians during administration of the BNT;
- a form to record relevant personal and medical information about the individual being assessed;
- clear instructions that need to be read by the clinician to the individual being assessed;
- a response form to record the response for each item, the quantitative lenient score, the manner of response (to record any delays or self corrections), the type of response/error (e.g. phonological or semantic association), and to record any cues (semantic and/or phonemic) that were necessary for each item;
- normative data that is stratified for age and education;
- a description of the types of responses that may be produced, with examples, together with a list of common and atypical error data from the study;

Information about each of the above sections is described below.

11.2.1 Instructions for test administration

The instructions for administering the Maltese BNT are intended for clinicians administering the test. It is necessary that instructions are consistently adhered to in order to ensure that the results are reliable.

11.2.2 Record of personal information

Clinicians must record salient information about the individual being assessed in order to have a clear overview of the features which may affect naming performance. Both demographic and medical details should be recorded on the respective form.

11.2.3 Instructions for individuals being assessed

The instructions given to individuals being tested should be consistently used with everyone. These instructions are similar to those used during the study. A Maltese translation of the following instructions is also available in a separate record booklet.

I am going to show you a set of pictures. I would like you to name each picture:

- *with a **single word**,*
- *with the **first name** that comes to mind, and*
- *as **quickly and accurately** as possible.*

11.2.4 The response form

The response form is designed in such a way that both quantitative and qualitative analysis may be carried out. It provides an opportunity to record responses and to score the test at the same time as the test is being administered. These response pages should also provide a clear overview of the naming skills and difficulties of the person being assessed as it provides both quantitative (lenient scores) and qualitative (manner of response, error profile, word variable effects) information of naming performance. Lenient scoring is being recommended as it has been found to be highly correlated with strict scoring, and because it allows other acceptable responses (including alternative responses and English translation equivalents) to be scored as correct. Lenient scoring allows the clinician to give credit to responses which are not target responses, but which may still be regarded as accurate responses (i.e. both target and alternative responses). This lenient scoring method has been recommended in proposed revisions to the original BNT (Nicholas et al 1989) and in bilingual contexts, as well as when using a test which was originally created on a diverse linguistic and cultural population. Target responses may be either in Maltese or in English, depending on which name was found to be the dominant response during the study. Alternative responses, therefore, include synonyms and the English or Maltese translation equivalent of the target response. The following is a description of the terms (headings) used in the response form.

No. refers to the item number of the modified version of the BNT in Maltese. There are 38 items/pictures in the reduced version of the test. The sequence of presentation of the items is not based on the frequency of correct responses obtained during the study. Rather, the items are presented in the original order of the BNT, to follow standardization procedures.

BNT No. refers to the original BNT item number.

Target response refers to the response that was used by the majority of the participants in the study (the dominant response). Some target responses are in English since this was the trend observed during the study (e.g. 'mushroom' for *faqqiegħ)*. English names are listed in italics. Strict scoring (i.e. responses which match target responses only) is not recommended.

Alternative responses refer to other acceptable responses which were not the target response and will be used for lenient scoring. These alternative responses include items in English or Maltese (depending on which was the target response) and synonyms. A lenient score of 1 point is awarded for any response which matches the target or the alternative responses. Any other response should be scored as incorrect and the type of error should be noted.

AoA refers to the Age of Acquisition of the words included in the reduced set of the BNT. Early (E) AoA refers to words which were acquired early in life, while Late (L) AoA refers to words which were learned later in life. Since this was found to be the most influential predictor word variable for naming performance in this study, it is useful to know which words have early AoA and which have late AoA. Since **Syllable Length** also remained significantly related to naming performance in regression analysis, the number of syllables is also specified. The number listed in the syllable length columns refers to the number of syllables of each target response. In order to identify whether there is a variable effect two subsets of words are presented towards the end of the booklet (see page 296). These two short lists include 10 early AoA words and 10 late AoA words. Each list has the same number of short and long syllable words.

In the **Response** column a transcription of the actual response should be written.

In **Manner of Response** it is possible to note down any hesitation or delay (**D**) in responding, as well as self corrections (**Sc**). These notes enable the clinician to make a qualitative as well as a quantitative analysis of the results obtained.

It is possible to acquire an **Error Profile** of each individual by recording the type of error/s made. A description of the different types of responses or errors is presented at the end of the record booklet. The different types of errors should be analyzed to differentiate between common errors (which are produced by both impaired and unimpaired individuals) and atypical errors (which are only apparent in the error profiles of persons with naming impairments). The presence of atypical errors in one's response profile is a good indication of naming impairments.

The **Lenient Score** column should be used to record the score obtained for each item. Lenient scoring is recommended in order to give credit to all acceptable responses. A **Lenient Score** is derived by scoring 1 point for all the responses which match either the target response or one of the acceptable responses listed in the alternative responses column, responses which follow a semantic cue, and self corrections. A correct response following a phonemic cue should **not** be scored as correct in order to follow standardization procedures. Any phonemic and/or semantic cues provided must be recorded in the final column (**PH/S cue**).

11.2.5 Normative data

Normative data stratified by age and education is presented in the record booklet. This consists of means, standard deviations and confidence intervals for 2 age groups and 2 levels of education. Since a decline in overall naming performance was evident from 66 years of age, the normative data is stratified into 2 age groups - a younger age group (20 – 65 years of age) and an older age group (66 + years of age). The education categories also reflect the methodology followed during the study and consists of a lower education category (0 – 12 years of education) and a higher education category (13 or more years of education). Norms are presented for lenient scoring since this is the recommended scoring procedure. Mean Scores (M) and Standard Deviations (SD) for the Modified/Maltese Boston Naming Test (MBNT) are available. Further, 68% confidence intervals (CI), consisting of upper and lower limits, are also provided to account for the probability that the score lies within one standard deviation away from the mean score.

11.2.6 Response categories (Error Profile)

This section describes the types of responses or errors that were produced during the study on individuals with impaired and unimpaired naming. The Record booklet includes two types of information regarding response categories. The first table lists the common types of errors and the atypical types of errors for the clinician to list the number of each type of error. Different researchers have utilised different error categories to classify the responses in their data. The list presented in the final table has been adapted from the categories used by various authors to reflect the type of responses that occurred throughout the present study. Examples accompany each description.

11.3 CONCLUSION

The recommendations described in the above sections should provide a clear framework for administering an adapted version of the BNT on Maltese speakers and for interpreting the results obtained. The modifications proposed should ensure that the method of administration of the test is both culturally and linguistically appropriate for Maltese speakers, and that the results are interpreted according to the findings made during the present study. The comprehensive record booklet (see following pages) has been especially created with the aim of offering a means of recording as much information as possible about an individual's naming skills. The booklet also includes specific quantitative and qualitative data that was obtained during the study. This additional information about word variables and error categories is a necessary component for qualitative analysis, interpretation of results and for differentiating between impaired and unimpaired word retrieval ability. Differential diagnosis of level of breakdown in naming impairment may utilize results obtained from the naming test (types of errors produced), information from additional tests of semantic and phonological integrity (such as spoken and written comprehension, repetition and reading), analysis of word variable effects, response to cues and miscues, etc..

BOSTON NAMING TEST

FOR MALTESE SPEAKERS

MBNT

The first part of this research was funded by a Malta Government Scholarship (MGSS grant MST 2008-20), while the second part was made possible by means of a scholarship awarded by the University of Limerick EU Fee Waiver for the Academic Year 2010/11. This research was part of a doctoral research degree at the faculty of Education and Health Sciences, University of Limerick, Limerick, Ireland. The researcher is indebted to Professor Sue Franklin (research supervisor) for her support and guidance throughout the research project.

Ritienne Grima (2011)

INSTRUCTIONS FOR TEST ADMINISTRATION

The following instructions are intended for clinicians administering the test. It is necessary that instructions are consistently adhered to in order to ensure that the results are reliable.

- The modified BNT for Maltese speakers should be administered in its entirety in one session whenever possible.
- In the case of failure to respond correctly the test should be discontinued following 8 consecutive incorrect responses.
- The personal details and medical history of the individual being assessed should be recorded by completing the corresponding form of the booklet.
- The instructions given to individuals being tested should be consistently used with everyone. These instructions are similar to those used during the study.
- Each picture is presented one by one to the individual being assessed.
- A time limit of **12** seconds is allowed before attempting to provide a phonemic cue.
- If visual misinterpretation of particular pictures is suspected, a semantic cue may be provided. If the response is still not correct, a phonemic cue may, then, be provided.
- The response form of this booklet should be used during administration of the test, and should be used for recording all the data obtained from individuals being tested.
- All cues given must be noted by means of writing Ph (phonemic cue) or S (Semantic cue) in the respective column (Ph/S cue) of the record booklet. If the individual benefits from a phonemic or a semantic cue, by producing a correct response, this should also be noted by a √ next to Ph or S.

- All responses should be recorded verbatim in the space provided on the response form (Response).
- A lenient score of 1 point should be awarded for every response that matches a target or an alternative response. This should be marked in the column labeled *lenient score* in the response form.
- Delayed (D) responses (i.e. responses which take longer than 12 seconds) and self corrections (Sc) should also be noted in the respective column (Manner of responses).
- A correct response that follows a semantic cue in the case of visual misinterpretation should be regarded as correct and scored accordingly as long as the semantic cue is noted down (S cue). Self corrections (without phonemic cues) should also be scored as correct as long as a note is made (Sc).
- Correct responses that follow a phonemic cue should **not** be scored as correct. It is necessary to follow this scoring procedure in order to match the scoring methodology used during the study.
- Correct responses that follow conduite d'approche should **not** be scored as correct, as these are considered atypical errors, which appeared only in PWA.
- The type of errors that occur must be noted in the column provided (Error profile). The error categories are listed at the end of the response form with examples for each category.
- At the end of the test administration the total lenient score should be worked out by adding all the points. The scores may, then, be matched to the normative data.
- The number of delayed responses, self corrections, semantic and phonemic cues given and the total for each type of error should be added and recorded in the Error profile. The Error profile and the other qualitative notes made during the assessment should provide additional information for interpretation of the individual's naming skills and difficulties.

MODIFIED BOSTON NAMING TEST FOR MALTESE SPEAKERS

Name _______________ I.D. no. _____________ DOB _________ Age___ Date ________________

Address ____ Tel. no. _____________________

Level of education _____________ Language/s spoken daily _______________________________________

Years of education _____________ Language/s heard daily ____

Marital status _________ Lives with ____________ Independence _________ Family support ________

Handedness ___________ Ability to read & write prior to illness _____ Current ability to read & write _____

Occupational history ____

Medical diagnosis ____________________________ Current SLP diagnosis _________________________

Medical history: Hypertension Diabetes Mellitus Previous CVA Seizures CHF/IHD Asthma COPD
Psychiatric disorders Visual/Auditory difficulties

Date of admission to hospital OR
Date of neurological insult ______________________

Progress of illness/side affected _________________
____

CT/MRI brain imaging reports ____

I am going to show you a set of pictures. I would like you to name each picture *with a* ***single word****, with the* ***first name*** *that comes to mind, and as* ***quickly and accurately*** *as possible.*

RESPONSE FORM

No.	BNT No.	Target response	Alternative responses	AoA[1]		Syllable length[2]		Response	Manner of Response		Error Profile	Lenient Score	Ph/S cue
				E	**L**	**1-2**	**3-4**		**D**	**Sc**			
1	1	sodda	*bed*	E		2							
2	2	siġra	*tree*	E		2							
3	3	lapes	*pencil*	E		2							
4	4	dar	*house*	E		1							
5	5	suffara	*whistle*	E			3						
6	6	mqass	*scissors*	E		1							
7	7	pettne	moxt, *comb*	E		2							
8	8	fjura	*flower*	E		2							
9	9	serrieq	*saw*		L	2							
10	10	*toothbrush*	xkupilja tas-snien	E		2							
11	11	ħelikopter		E			4						
12	12	xkupa	mikinsa, *broom*	E		2							
13	13	qarnita	*octopus*	E			3						
14	14	*mushroom*	faqqiegħ	E		2							

[1] E = Early AoA (0-6 years of age), L = Late AoA (6+ years of age).
[2] 1-2 = 1-2 syllable words; 3-4 = 3-4 syllable words.

	BNT No.	Target response	Alternative responses	AoA		Syllable length		Response	Manner of Response		Error Profile	Lenient Score	Ph/S cue
				E	**L**	**1-2**	**3-4**		**D**	**Sc**			
15	15	spalliera	*hanger*	E			3						
16	17	ġemel	*(camel)*	E		2							
17	18	maskra/maskla	*(mask)*	E		2							
18	20	bank	*(bench)*	E		1							
19	22	bebbuxu	*snail*	E			3						
20	23	vulkan	*(volcano)*	E		2							
21	25	dart			L	1							
22	27	globu	id-dinja, (*globe)*		L	2							
23	31	*rhinoceros*	(rinoċeronte)		L		4						
24	33	iglù			L	2							
25	35	*dominoes*	(domino)		L		3						
26	36	kaktus			L	2							
27	37	*escalator*			L		4						
28	38	arpa	*(harp)*		L	2							
29	40	ħabbata	*knocker*		L		3						
30	42	*stethoscope*	(stetoskopju)		L		3						
31	43	piramida	*(pyramid)*		L		4						
32	45	*unicorn*	(unikornu)		L		3						
33	46	lembut	*funnel*		L	2							
34	50	kumpass	*(compass)*		L	2							

	BNT No.	Target response	Alternative responses	AoA		Syllable length		Response	Manner of Response		Error Profile	Lenient Score	Ph/S cue
				E	**L**	**1-2**	**3-4**		**D**	**Sc**			
35	52	*tripod*	(tripied)		L	2							
36	58	*palette*	(palette), tavlozza		L	2							
37	59	*protractor*			L		3						
38	60	*abacus*	(abaku)		L		3						
Total Lenient Score													

NORMATIVE DATA FOR THE REDUCED SET OF THE MALTESE BNT

	AGE GROUP 1 20 – 65 YEARS			AGE GROUP 2 66 + YEARS		
	MBNT LENIENT SCORE			MBNT LENIENT SCORE		
	M	**SD**	**CI**	**M**	**SD**	**CI**
0 – 12 YEARS OF EDUCATION	29.1	3.1	27.7 – 30.4	23.9	4.4	22.8 – 25.1
13 OR MORE YEARS OF EDUCATION	34.2	3.1	33.2 – 35.2	28.5	9.3	26.8 – 30.2

ERROR PROFILE (state number of …)			
Number of semantic cues given (**S**)		Number of self corrections (**Sc**)	
Number of phonemic cues given (**Ph**)			
Common Errors Errors produced by people with and without aphasia		**Atypical Errors** Errors produced by people with aphasia only	
Coordinate		Multiword (MW)	
Superordinate		Morphological alteration	
Subordinate		Perseveration	
Semantic association (S Assoc)		Unrelated	
Circumlocution		Forgot	
Phonologically related real word (PhRR)		Phonologically related non word (PhRN)	
Don't know		Successful conduite d'approche	
Visual Association/not recognized		Unsuccessful conduite d'approche	
Self correction		Neologism	
Correct following a phonemic cue		Incorrect following a phonemic cue	

EARLY AOA WORDS			
No.	**Target response**	**Syllable length**	**Response**
1	sodda	short	
2	siġra	short	
3	lapes	short	
4	dar	short	
6	mqass	short	
5	suffara	long	
11	ħelikopter	long	
13	qarnita	long	
15	spalliera	long	
19	bebbuxu	long	

LATE AOA WORDS			
No.	**Target response**	**Syllable length**	**Response**
9	serrieq	short	
21	dart	short	
22	globu	short	
24	iglù	short	
26	kaktus	short	
25	*dominoes*	long	
23	*rhinoceros*	long	
27	*escalator*	long	
29	ħabbata	long	
31	piramida	long	

Comments

DESCRIPTION OF EACH TYPE OF ERROR

Coordinate	A response that is derived from the same semantic category as the target word, e.g. ajruplan / ħelikopter (*aeroplane / helicopter).*	**Successful conduite d'approche**	Repeated sounds, syllables or word attempts which result in successfully producing the target word, e.g. kum, ku, kumpass *(compass).*
Superordinate	A response that is more vague and general than the target word, e.g. braxx / *toothbrush* (*brush / toothbrush).*	**Unsuccessful conduite d'approche**	Repeated sounds, syllables or word attempts which do not result in successfully producing the target word, e.g. pila, pilalida, pilapida / piramida *(pyramid).*
Subordinate	A response that is more specific than the target word, e.g. villa / dar (*villa* / *house*).	**Don't know**	Person admits that s/he does not know the name for the item.
Semantic Association	A response that is semantically related to the target word, e.g. saħħara / xkupa *(witch / broom).*	**Forgot**	Person admits that s/he does has forgotten the name for the item or cannot retrieve it at the moment.
Circumlocution	A response which correctly describes a target word, e.g. torqod fiha / sodda (*you sleep in it / bed).*	**Self correction**	Immediate self correction of an incorrect response.
Multiword	Empty connected speech that fails to describe the target word, e.g. ilbieraħ sajjartha / qarnita *(I cooked it yesterday / octopus).*	**Morphological alteration**	Morphological alteration to the target word, e.g. bebbux / bebbuxu *(snails / snail).*
Phonologically related real word (PhRR)	A real word that is phonologically related to the target word, e.g. lapida / piramida (*memorial stone / pyramid*).	**Perseveration**	Repeated productions of an earlier response.
Phonologically related non-word (PhRN)	A nonword that is phonologically related to the target word, e.g. splaxrum / *mushroom.*	**Visual Association**	A response that is visually associated to the target word/picture, e.g. magg / maskra *(mug / mask – BNT picture),* or else person admits that s/he cannot recognize the item.
Unrelated	A real word that is not related[3] semantically and/or phonologically to the target word, e.g. kejk / dar (cake / house).	**Phonemic cue (correct)**	Correct response following a Phonemic Cue.
Neologism	A nonword that is not related to the target word, e.g. kim / ħelikopter (*helicopter).*	**Phonemic cue (incorrect)**	Incorrect response following a Phonemic Cue.

[3]An unrelated word or nonword must share less than 50% of the phonemes of the target word (Nickels 2001).

References

REFERENCES

Albert, M. L., Spiro, A., Sayers, K. J., Cohen, J. A., Brady, C. B., Goral, M. and Obler, L. K. (2009) 'Effects of Health Status on Word Finding in Aging', *Journal of the American Geriatrics Society,* 57(12), 2300-2305.

Abulatebi, J. and Green, D. (2007) 'Bilingual speech production: The neurocognition of language representation and control', *Journal of Neurolinguistics,* 20, 242-275.

Agrell, B. and Dehlin, O. (1998) 'The Clock-Drawing Test', *Age and Ageing* [online], 27, 399-403, available http://www.findarticles.com/p/articles/mi_m2459/is_n3_v27/ai_211326 56/ [accessed 25 Aug 2008].

Alario, F. X. and Ferrand, L. (1999) 'A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability and age of acquisition', *Behaviour Research Methods, Instruments and Computers,* 31(3), 531-552, available: http://ludovic.ferrand.googlepages.com/my publications [accessed 27 Mar 2008].

Alario, F. X., Costa, A. and Caramazza, A. (2002) 'Frequency effects in noun phrase production: Implications for models of lexical access', *Language and Cognitive Processes,* 17(3), 299-319.

Alario, F. X., Ferrand, L., Laganaro, M., New, B. Frauenfelder, U. H. and Segui, J. (2004) 'Predictors of picture naming speed', *Behaviour Research Methods, Instruments, & Computers*, 36(1), 140-155, available: http://search.yahoo.com/search;_ylt=A0geu.MbX4dHpmQA05ZXNyoA? [accessed 11 Jan 2008].

Albert, M. S., Heller, H. S. and Milberg, W. (1988) 'Changes in naming ability with age', *Psychology and Aging,* 3(2), 173-178.

Allegri, R. F., Mangone, C. A., Fernandez Villavicencio, A., Rymberg, S., Taragano, F. E. and Baumann, D. (1997) 'Spanish Boston Naming Test Norms', *The Clinical Neuropsychologist,* 11(4), 416-420.

Almeida, J., Knobel, M., Finkbeiner, M. and Caramazza, A. (2007) 'The locus of the frequency effect in picture naming: When recognizing is not enough', *Psychonomic Bulletin & Review,* 14 (6), 1177-1182, available: http%3a//www.wjh.Harvard.edu/¬jalmeida/R196B_MS_Almeida.pdf [accessed 5 Jan 2008].

Almkvist, O., Adveen, M., Henning, L. and Tallberg, I.M. (2007) 'Estimation of premorbid cognitive function based on word knowledge: the Swedish Lexical Decision Test (SLDT)', *Scand J Psychol,* 48(3), 271-279.

Aquilina, J. (1987-1990) *Maltese-English dictionary* (vols. 1-2), Malta: Midsea Books Ltd.

Aquilina, J. (1999-2000) *English-Maltese dictionary* (vols. 1-4), Malta: Midsea Books Ltd.

Azuma, T., Bayles, K. A., Cruz, R. F., Tomoeda, C. K., Wood, J. A., McGeagh, A. (1997) 'Comparing the difficulty of Letter, Semantic and Name Fluency Tasks for Normal Elderly and Patients with Parkinson's disease', *Neuropsychology,* 1(4): 488-497.

Barker-Collo, S. L. (2001) 'The 60-item Boston Naming Test: Cultural bias and possible adaptations for New Zealand', *Aphasiology,* 15(1), 85-92.

Barker-Collo, S. L. (2007) 'Boston Naming Test performance of older New Zealand adults', *Aphasiology,* 21(12), 1171-1180.

Barresi, B. A., Nicholas, M., Tabor Connor, L., Obler, L. K. and Albert, M. L. (2000) 'Semantic Degradation and Lexical Access in Age-Related Naming Failures', *Aging, Neuropsychology and Cognition*, 7(3), 169-178.

Belke, E. and Meyer, A. S. (2007) 'Single and multiple object naming in healthy ageing', *Language and Cognitive Processes,* 22(8), 1178-1211.

Bialystok, E. (2001) *Bilingualism in development: Language, literacy, and cognition,* New York: Cambridge University Press.

Bialystok, E. (2006) 'Bilingualism and cognitive development: Evidence for changes in attentional control', accepted for conference *Language acquisition and bilingualism: Implications for a multilingual society,* May 2006.

Bialystok, E. and Feng, X. (2009) 'Language Proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults', *Brain and Language,* 109, 93-100.

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

Bialystok, E., Craik, F. I. M. and Klein, R. (2004) 'Bilingualism, aging, and cognitive control: Evidence from the Simon task', *Psychology and Aging,* 19(2), 290-303.

Bialystok, E., Craik, F. I. M. and Luk, G. (2008a) 'Lexical access in bilinguals: Effects of vocabulary size and executive control', *Journal of Neurolinguistics,* 21 (6), 522-538.

Bialystok, E., Craik, F. I. M. and Luk, G. (2008b) 'Cognitive control and lexical access in younger and older bilinguals', *Journal of Experimental Psychology - Learning, Memory and Cognition,* 34 (4), 859-873.

Bialystok, E., Craik, F. I. M. and Ryan, J. (2006) 'Executive control in a modified antisaccade task: Effects of aging and bilingualism', *Journal of Experimental Psychology - Learning, Memory and Cognition,* 32(6), 1341-1354.

Biedermann, B., Ruh, N., Nickels, L. and Coltheart, M. (2008) 'Information Retrieval in Tip of the Tongue States: New Data and Methodological Advances', *Journal of PsycholinguisticResearch,* 37, 171-198.

Bird, H., Franklin, S. and Howard, D. (2001) 'Age of Acquisition and imageability ratings for a large set of words, including verbs and function words', *Behaviour Research Methods, Instruments, & Computers,* 33 (1), 73-79, available: http://a465.g.akamai.net/f/465/1984/1d/www.ingentaconnect.com/images/pdf_button.gif [accessed 17 Jan 2008].

Bird, H., Howard, D. and Franklin S. (2000) 'Why is a verb like an inanimate object? Grammatical category and semantic category deficits', *Brain and Language,* 72(3), 246-309.

Bird, H., Howard, D. and Franklin, S. (2003) 'Verbs and nouns: the importance of being imageable', *Journal of Neurolinguistics,* 16, 113-149.

Bormann, T., F. Kulke, C-W. Wallesch & G. Blanken. (2008). 'Omissions and semantic errors in aphasic naming: Is there a link?' *Brain and Language*, 104, 24 – 32.

Brincat, J. M. (2000) *Il-Malti: Elf Sena ta' Storja*, Malta : Pubblikazzjonijiet Indipendenza.

Brodaty, H. and Moore, C. M. (1997) 'The Clock Drawing Test for Dementia of the Alzheimer's type: a comparison of 3 scoring methods in a memory disorders clinic', *International Journal of Geriatric Psychiatry,* 12, 619-627.

Brutt-Griffler, J. and Varghese, M. (2004). *Bilingualism and language pedagogy.* Clevedon, UK : Multilingual Matters Ltd.

Budd, A. (2007) *Boston Naming Test with Latencies (BNT-L)* unpublished thesis (Ph.D.), University of North Texas, Texas.

Budd, M. A., Kortte, K., Cloutman, L., Newhart, M., Gottesman, R. F., Davis, C., Heidler-Gary, J., Seay, M. W. and Hillis, A. E. (2010) 'The nature of naming errors in primary progressive aphasia versus acute post-stroke aphasia', *Neuropsychology,* 24(5), 581 – 589.

Burns, A., Lawlor, B. and Craig, S. (2004) *Assessment Scales in Old Age Psychiatry,* 2nd ed., London: Martin Dunitz.

Cabeza, R., Nyberg, L. and Park, D. C. (2005) 'Cognitive Neuroscience of Aging: Emergence of a New Discipline', in Cabeza, R., Nyberg, L. and Park, D. C. (eds.) *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging,* New York: Oxford University Press, 3-15.

Camilleri, A. (1995) *Bilingualism in education,* Heidleberg: Julius Groos Verlag.

Camilleri, L. (2009) ‘Statistical model frameworks’ *M Stor Newsletter* 1(3). Malta: Statistics & Operations Research Department, University of Malta.

Caramazza, A. and Hillis, A. E. (1990) ‘Where do semantic errors come from?’, *Cortex,* 26, 95-112.

Carroll, J. B. and White, M. N. (1973) ‘Word frequency and age-of-acquisition as determiners of picture-naming latency’, *Quarterly Journal of Experimental Psychology,* 25, 85-95.

Casas, R., Calamia, M. and Tranel, D. (2008) ‘A screening test of English naming ability in bilingual Spanish/English speakers’, *Journal of clinical and experimental neuropsychology,* 30(8), 956-966.

Catling, J. C. and Johnston, R. A. (2008) ‘The varying effects of age of acquisition’, *The Quarterly Journal of Experimental Psychology,* 62(1), 50-62.

Christensen, H., Korten, A. E., Jorm, A. F., Henderson, A. S., Jacomb, P. A., Rodgers, B. and Mackinnon, A. J. (1997) ‘Education and decline in cognitive performance: compensatory but not protective’, *International Journal of Geriatric Psychiatry,* 12, 323-330.

Christoffels, I. K., DeGroot, A. M. B. and Kroll, J. F. (2006) ‘Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency’, *Journal of Memory and Language,* 54, 324-345.

Christoffels, I. K., Firk, C. and Schiller, N. O. (2007) ‘Bilingual language control: An event-related brain potential study’, *Brain Research,* 1147, 192-208.

Comrie, B. (2009) 'Maltese and the world atlas of language structures' in Comrie, B., Fabri, R., Hume, E., Mifsud, M., Stolz, T. and Vanhove, M. (eds.) *Introducing Maltese Linguistics - Studies in language companion series*, Vol. 113. The Netherlands: John Benjamins Publishing Company, 3-11.

Connor, L. T., Spiro, A., Obler, L. K. and Albert, M. L. (2004) ‘Change in Object Naming Ability During Adulthood’, *Journal of Gerentology,* 59B (5), 203-209.

Cortese, M. J. and Fugett, A. (2004) ‘Imageability ratings for 3,000 monosyllabic words’, *Behaviour Research Methods Instruments and Computers,* 36(3), 384-387.

Cortese, M. J. and Khanna, M. M. (2007) ‘Age of Acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words’, *Quarterly Journal of Experimental Psychology,* 60 (8), 1072-1082.

Costa, A. (2005) ‘Lexical access in bilingual production’, in Kroll, J. F. and DeGroot, A. M. B. (eds.), *Handbook of Bilingualism: Psycholinguistic Approaches,* New York: Oxford University Press, 308-325.

Costa, A., Caramazza, A. and Sebastian-Galles, N. (2000) 'The Cognate Faciliation Effect: Implications for Models of Lexical Access', *Journal of Experimental Psychology: Learning, Memory and Cognition,* 26 (5), 1283-1296.

Costa, A., Hernández, M. and Sebastián-Gallés, N. (2008) 'Bilingualism aids conflict resolution: evidence from the ANT task', *Cognition,* 106(1), 59-86.

Costa, A., Santesteban, M. and Ivanova, I. (2006) 'How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional', *Journal of Experimental Psychology and Learning,* 32, 1057-1074.

Cox, P.P.H. (2008) *Recovery of naming in aphasia,* unpublished thesis (M.A.), Utrecht University, Utrecht.

Cruice, M., Worrall, L. E. and Hickson, L. M. H. (2000) 'Boston Naming Test results for healthy older Australians: A longitudinal and cross-sectional study', *Aphasiology,* 14 (2), 143-155.

Cuetos, F., Aguado, G. and Caramazza, A. (2000) 'Dissociation of Semantic and Phonological Errors in Naming', *Brain and Language,* 75, 451-460.

Cuetos, F., Aguado, G., Izura, C. and Ellis, A. W. (2002) 'Aphasic naming in Spanish: predictors and errors', *Brain and Language,* 82 (3), 344-365.

Cuetos, F., Glez-Nosti, M., Barbón, A. And Brysbaert, M. (2011) 'SUBTLEX-ESP: Spanish word frequencies based on film subtitles', *Psicológica,* 32, 133-143.

D'Amico, S., Devescovi, A. and Bates, E. (2001) 'Picture naming and lexical access in Italian children and adults', *Journal of Cognition and Development*, 2 (1), 71-105, available: http://crl.ucsd.edu/¬bates/papers/pdf/from-meiti/22-DAmico%20er%20al.pdf [accessed 14 Dec 2007].

Damian, M. F., Bowers, J. S., Stadthagen-Gonzalez, H. and Spalek, K. (2010) 'Does word length affect speech onset latencies when producing single words?' *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(4), 892-905.

De Leon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A. and Hillis, A. E. (2007) 'Neural regions essential for distinct cognitive processes underlying picture naming', *Brain,* 130, 1408 - 1422.

Delia, K. L., Ellul Mercer, E., Fiorini, H. and Ameen, A. (2010) *Language of the Elderly Screening Test (LEST)*, poster presented at the 28th World Congress of the International Association of Logopedics and Phoniatrics (IALP), Athens, Greece.

Dell, G. S., & O'Seaghdha, P. G. (1992) 'Stages of lexical access in language production', *Cognition*, 42, 287–314.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M. and Gagnon, D. A. (1997) 'Lexical access in aphasic and nonaphasic speakers', *Psychological Review,* 104(4), 801-838.

Elkadi, S., Clark, M. S., Dennerstein, L., Guthrie, J. R., Bowden, S. C. and Henderson, V. W. (2006) 'Normative data for Australian midlife women on category fluency and a short form of the Boston Naming Test', *Australian Psychologist*, 41(1), 37-42.

Ellis, A. W. and Morrison, C. M. (1998) 'Real Age-of-Acquisition Effects in Lexical Retrieval', *Journal of Experimental Psychology: Learning, Memory and Cognition,* 24(2), 515-523.

Ellul, S. (1978) *A Case Study in bilingualism: code-switching between parents and their pre-school children in Malta*. Cambridge: Huntington Publishers.

Elman, R. J., Klatzky, R. L., Dronkers, N. F., and Wertz, R. T. (1992) 'Phonologic priming and picture naming in aphasic and normal subjects, *Clinical Aphasiology*, 21, 255–262.

Evrard, M. (2002) 'Ageing and Lexical Access to Common and Proper Names in Picture Naming', *Brain and Language,* 81, 174-179.

Farmer, A. (1990) 'Performance of normal males on the Boston Naming Test and The Word Test', *Aphasiology,* 4(3), 293-296.

Feyereisen, P., Van der Borght, F. and Seron, X. (1988) 'The operativity effect in naming: a re-analysis', *Neuropsychologia,* 26(3):401-15.

Figueredo Balthazar, M. L., Cendes, F. and Pereira Damasceno, B. (2008) 'Semantic Error Patterns on the Boston naming Test in Normal Aging, Amnestic Mild Cognitive Impairment, and Mild Alzheimer's Disease: Is there Semantic Disruption?', *Neuropsychology,* 22(6), 703-709.

Finkbeiner, M., Gollan, T. H. and Caramazza, A. (2006) 'Lexical access in bilingual speakers: What's the (hard) problem?', *Bilingualism: Language and Cognition,* 9(2), 153–166.

Fisher, N. J., Tierney, M. C., Snow, G. W. and Szalai, J. P. (1999) 'Odd/Even Short Forms of the Boston Naming Test: Preliminary GeriatricNorms', *Clinical Neuropsychologist*, 13 (3), 359-364.

Folstein M.F., Folstein S.E. and McHugh P.R. (1975) 'MiniMental state: A practical method for grading the cognitive state of patients for the clinician', *Psychiatric Research*, 12, 189-98.

Forster, K. I. & Forster, J. C. (2003) 'DMDX: A Windows display program with millisecond accuracy', *Behavior Research Methods, Instruments, & Computers,* 35(1), 116-124.

Franklin, S. (1989) ‘Dissociations in auditory word comprehension – evidence from nine fluent aphasic patients’, *Aphasiology,* 3, 189-207.

Franklin, S., Buerk, F. and Howard, D. (2002) ‘Generalized improvement in speech production for a subject with reproduction conduction aphasia’, *Aphasiology,* 16, 1087-1114.

Franklin, S., Howard, D. and Patterson, K. (1995) ‘Abstract Word Anomia’, *Cognitive Neuropsychology*, 12, 549-566.

Gilhooly, K. J. and Logie, R. H. (1980) ‘Age-of-acquisition, imagery, concreteness, familiarity and ambiguity measures for 1,944 words’, *Behaviour Research Methods & Instrumentation,* 12 (4), 395-427, abstract available: http://doi.apa.org/?uid=1981-25426-001 [accessed 4 Jan 2008].

Goldrick, M. and Rapp, B. (2002) ‘A restricted interaction account (RIA) of spoken word production: The best of both worlds’, *Aphasiology*, 16(1/2), 20-55.

Gollan, T. H. and Acenas, L. A. R. (2004) ‘What is a TOT? Cognate and Translation Effects on Tip-of-the-Tongue States in Spanish-English and Tagalog-English Bilinguals’, *Journal of Experimental Psychology: Learning, Memory and Cognition,* 30 (1), 246-269.

Gollan, T. H. and Silverberg, N. B. (2001) ‘Tip-of-the-tongue states in Hebrew-English bilinguals’, *Bilingualism: Language and Cognition,* 4, 63-84.

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I. and Jernigan, T. L. (2007) ‘The bilingual effect on Boston Naming Test performance’, *Journal of the International Neuropsychological Society,* 13, 197-208.

Gollan, T. H., Montoya, R. I. and Werner, G. A. (2002) ‘Semantic and letter fluency in Spanish-English bilinguals’, *Neuropsychology,* 16(4), 562-576.

Gollan, T. H., Montoya, R. I., Cera, C. and Sandoval, T. C. (2008) ‘More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis’, *Journal of Memory and Language,* 58(3), 787-814.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C. and Morris, S. K. (2005) ‘Bilingualism affects picture naming but not picture classification’, *Memory and Cognition,* 33(7), 1220-1234.

González Nosti, M., Rodríguez Ferreiro, J. and Cuetos Vega, F. (2008) ‘Variabilidad en los errores semánticos producidos por pacientes con daño cerebral’, *Psicothema*, 20(4), 795-800, abstract available: http://academic.research.microsoft.com/Publication/11294099/variabilidad-en-los-errores-sem-nticos-producidos-por-pacientes-con-da-o-cerebral[accessed 4 July 2011].

Goral, M., Spiro, A. III., Albert, M. L., Obler, L. K. and Connor, L. T. (2007) ‘Change in Lexical Retrieval Skills in Adulthood’, *The Mental Lexicon* 2(2), 215-238.

Graves, R. E., Bezeau, S. C., Fogarty, J., Blair, R. and Bruyere, E. (2004) 'Boston Naming Test Short Forms: A Comparison of Previous Forms with New Item Response Theory Based Forms', *Journal of Clinical and Experimental Neuropsychology,* 26 (7), 891–902.

Grech, H. and Dodd, B. (2008) 'Phonological acquisition in Malta: A bilingual language learning context', *International Journal of Bilingualism,* 12 (3), 155-171.

Grossman, M., McMillan, C., Moore, P., Ding, L., Glosser, G., Work, M. and Gee, J. (2004) 'What's in a name: voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration', *Brain,* 127, 628-649.

Hanley, J. R., Dell, G. S., Kay, J. and Baron, R. (2004) 'Evidence for the involvement of a nonlexical route in the repetition of familiar words: A comparison of single and dual route models of auditory repetition', *Cognitive Neuropsychology,* 21(2/3/4), 147-158.

Hawkins, K. A. and Bender, S. (2002) 'Norms and the relationship of Boston Naming Test performance to vocabulary and education: A review', *Aphasiology,* 16(12), 1143-1153.

Heinik, J., Solomesh, I., Lin, R., Raikher, B., Goldray, D., Merdler, C. and Kemelman, P. (2004) 'Clock Drawing Test – Modified and Integrated Approach (CDT-MIA): Description and Preliminary Examination of Its Validity and Reliability in Dementia Patients referred to a specialized Psychogeriatric Setting', *Journal of Geriatric Psychiatry and Neurology,* 17(2), 73-80.

Herbert, R., Hickin, J., Howard, D., Osborne, F. and Best, W. (2008) 'Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia?', *Aphasiology,* 22(2), 184-203.

Hernandez, A. E. and Kohnert, K. J. (1999) 'Aging and language switching in bilinguals', *Aging Neuropsychology and cognition,* 6(2), 69-83.

Hillis, A. and Caramazza, A. (1991) 'Mechanisms for accessing lexical representations for output: Evidence from a category specific semantic deficit', *Brain and Language,* 40, 106-144.

Hobson, V. L., Hall, J. R., Harvey, M., Spering, C., Cullum, M., Lacritz, L., Massman, P., Waring, S. and O'Bryant, S. (2009) 'The Boston Naming Test: Calculation and Utilization of an Estimated 60-Item Score from 30- and 15-Item Administrations, *Archives of Clinical Neuropsychology,* 24(5), 462-462.

Hodgson, C. and Ellis, A. W. (1998) 'Last in, First to Go: Age of Acquisition and Naming in the Elderly', *Brain and Language,* 64, 146-163.

Howard, D. (2000) 'Cognitive neuropsychology and aphasia therapy: the case of word retrieval' in Papathanasiou, I., ed., *Acquired Neurogenic Communication Disorders: A Clinical Perspective*, London: Whurr Publishers.

Howard, D. and Gatehouse, C. (2006) 'Distinguishing semantic and lexical word retrieval deficits in people with aphasia', *Aphasiology,* Volume: 20 Issue: 9-11 Pages: 921-950.

Howard, D., Nickels, L., Coltheart, M. and Cole-Virtue, J. (2006) 'Cumulative semantic inhibition in picture naming: experimental and computational studies', *Cognition,* 100 (3), 464-482.

Ivanova, I. and Costa, A. (2008) 'Does bilingualism hamper lexical access in speech production?', *Acta Psychologica,* 127(2), 277-288.

Iyer, G., Saccuman, C., Bates, E., and Wulfeck, B. (2001) 'A study of Age of acquisition ratings in adults', *Centre for Research in Language Newsletter,* 13(2), available: http://crl.ucsd.edu/newsletter/13-2/article.html [accessed 11 Jan 2008].

Janyan, A. (2007) 'Lateralized object naming: Effects of Image Agreement, Name Agreement, AoA and Animacy', *The 2nd European Cognitive Science Conference,* (848-853), New York, NY: Lawrence Erlbaum Associates, Inc.

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

Kavé, G. (2005) 'Standardization and norms for a Hebrew naming test', *Brain and Language,* 92, 204-211.

Kim, H. and Na, D. L. (1999) 'Normative data on the Korean version of the Boston naming test*. J Clin Exp Neuropsychol,* 21, 127-133.

Kim, H., Kim, J., Kim, D. Y. and Heo, J. (2011) 'Differentiating between aphasic and nonaphasic stroke patients using semantic verbal fluency measures with administration time of 30 seconds', *European Neurology,* 65(2), 113 -117.

Kiran, S. and Tuchtenhagen, J. (2005) 'Imageability effects in normal Spanish-English bilingual adults and in aphasia: Evidence from naming to definition and semantic priming tasks', *Aphasiology,* 19(3-5), 315-327.

Kittredge, A. K., Dell, G. S., Verkuilen, J. and Schwartz, M. F. (2008) 'Where is the effect of frequency in word production? Insights from aphasic picture-naming errors.' *Cognitive Neuropsychology,* 25(4), 463-492.

Knobel, M., Finkbeiner, M. and Caramazza, A. (2008) 'The many places of frequency: Evidence for a novel locus of the lexical frequency effect in word production', *Cognitive Neuropsychology*, 25(2), 256-286.

Kohnert, K. J. (2004) ‘Cognitive and cognate-based treatments for bilingual aphasia: A case study’, *Brain and Language,* 91(3), 294-302.

Kohnert, K. J., Hernandez, A. E. and Bates, E. (1998) ‘Bilingual Performance on the Boston Naming Test: Preliminary Norms in Spanish and English’, *Brain and Language,* 65, 422-440.

Kuipers, J-R. and La Heij, W. (2009) ‘The limitations of cascading in the speech production system’, *Language and Cognitive Processes,* 24(1), 120-135.

Laine, M. and Martin, N. (1996) ‘Lexical retrieval deficit in picture naming: implications for word production models’, *Brain and language,* 53, 283-314.

Lambon Ralph, M. A., Moriarty, L. and Sage, K. (2002) ‘Anomia is simply a reflection of semantic and phonological impairments: Evidence from a case-series study’, *Aphasiology*, 16(1) 56 – 82.

Lambon Ralph, M. A., Sage, K. and Roberts, J. (2000) ‘Classical anomia: a neuropsychological perspective on speech production’, *Neuropsychologia*, 38(2), 186-202.

Levelt, W. J. M. (2001) ‘Spoken word production: A theory of lexical access’, *Proceedings of the National Academy of Sciences of the United States of America,* 98(23), 13464-13471.

Levelt, W. J. M. (2002) ‘Picture naming and word frequency: Comments on Alario, Costa and Caramazza’, *Language and Cognitive Processes,* 17(6), 663-671.

Levelt, W. J. M., Roelofs, A. and Meyer, A. S. (1999) ‘A theory of lexical access in speech production’, *Behavioural and brain Sciences,* 22(1), 1-75.

Levelt, W. J.M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T. and Havinga, J. (1991) ‘Normal and Deviant Lexical Processing: Reply to Dell and O’Seaghdha’, *Psychological Review,* 98(4), 615-618.

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

Li, S. C. (2005) ‘Neurocomputational perspectives linking neuromodulation, processing noise, representational distinctiveness and cognitive aging’ in Cabeza, R., Nyberg, L. and Park, D. C. (eds.), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging,* New York: Oxford University Press, 354-379.

Lindeboom, J. and Weinstein, H. (2004) ‘Neuropsychology of cognitive ageing, minimal cognitive impairment, Alzheimer’s disease, and vascular cognitive impairment’, *European Journal of Pharmacology,* 490, 83-86.

Lin-Kiat Yap, P., Ng, T., Niti, M., Yeo, D. and Henderson, L. (2007) ‘Diagnostic Performance of Clock Drawing Test by CLOX in an Asian Chinese Population’, *Dementia and Geriatric Cognitive Disorders,* 24: 193-200.

Mack, W. J., Freed, D.M., Williams, B.W. and Henderson V. W (1992) ‘Boston Naming Test: shortened versions for use in Alzheimer's disease’, Journal of Gerontology, 47(3), 154-158.

MacKay, A., Connor, L. T. and Storandt, M. (2005) ‘Dementia does not explain correlation between age and scores on Boston Naming Test’, *Archives of Clinical Neuropsychology,* 20 (129-133).

Madden, D. J., Whiting, W. L. and Huettel, S. A. (2005) ‘Age-related changes in neural activity during visual perception and attention’ in Cabeza, R., Nyberg, L. and Park, D. C. (eds.), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging,* New York: Oxford University Press, 157-185.

Manos, P. J. (1998) ‘10-Point Clock Test Screens for Cognitive Impairment in Clinic and Hospital Settings’, *Psychiatric Times* [online], 15(10), available: http://www.psychiatrictimes.com/display/article/10168/54386? [accessed 20 Aug 2008].

Mansur, L. L., Radanovic, M., Rüegg, D., Zanotto de Mendonça, L. and Scaff, M. (2002) ‘Descriptive study of 192 adults with speech and language disturbances’, *Sao Paulo Med J/Rev Paul Med,* 129(6), 170-174.

Mansur, L. L., Radanovic, M., Araujo G., Taquemori, L. Y. snf Greco, L. L. (2006) ‘Boston Naming Test: performance of Brazilian population from São Paulo’, *Pró-Fono Revista de Atualização Científica,*18 (1), available: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010456872006000100003&Ing=en&nrm=iso> [accessed 15 Sept 2007].

Mariën, P., Mampaey, E., Vervaet, A., Saerens, J. and De Deyn P. P. (1998) ‘Normative Data for the Boston Naming Test in Native Dutch-Speaking Belgian Elderly’, *Brain and Language,* 65, 447-467.

Martin, N., Schwartz, M. F., Kohen, F. P. (2006) ‘Assessment of the ability to process semantic and phonological aspects of words in aphasia: a multi-measurement approach’, *Aphasiology*, 20(2-4), 154-166.

Mathuranath PS, G. A. (2007) ‘Community-based naming agreement, familiarity, image agreement and visual complexity ratings among adult Indians’, *Ann Indian Acad Neurol,* 10(2), 92-99, available: http://www.annalsofian.org/text.asp?2007/10/2/92/33216 [accessed 12 Dec 2007].

Mesulam, M. (2008) ‘Representation, inference, and transcendent encoding in neurocognitive networks of the human brain’, *Annals of Neurology,* 64(4), 367–378.

Morrison, C. M., & Ellis, A. W. (2000) 'Real age of acquisition effects in word naming and lexical decision', *British Journal of Psychology*, 91, 167-180.

Morrison, C. M., Chappell, T. D. and Ellis, A. W. (1997) 'Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables, *Quarterly Journal of Experimental Psychology,* 50, 528-559.

Morrison, C. M., Ellis, A. W. and Quinlan, P. T. (1992) 'Age of acquisition, not word frequency, affects object naming, not object recognition', *Memory & Cognition,* 20(6), 705-715.

National Statistics Office (2006) *Report on Census of Population and Housing 2005,* Vol. 1., Malta: National Statistics Office, available: http://www.census2005.gov.mt/repo rts/Census2005_Vol.1.pdfx [accessed 20 Sept 2008].

Neils, J., Baris, J. M., Carter, C., Dell'aria, A. L., Nordloh, S. J., Weiler, E. and Weisiger, B. (1995) 'Effects of Age, Education, and Living environment on Boston Naming Test Performance', *Journal of Speech and Hearing Research,* 38, 1143-1149.

Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G. and Parrazzo, S. A. (1989) 'Revised administration and scoring procedures for the Boston Naming Test and norms for non-brain-damaged adults', *Aphasiology,* 3(6), 569-580.

Nicholas, M., Obler, L., Albert, M. and Goodglass, H. (1985) 'Lexical retrieval in healthy aging', *Cortex,* 21(4), 595-606.

Nickels, L. (2000) 'A sketch of the cognitive processes involved in the comprehension and production of single words', available: http://www.maccs.mq.edu.au/~ lyndsey /model available: [accessed 2 Feb 2009].

Nickels, L. (2001) 'Producing spoken words', in Rapp, B., ed., *The handbook of Cognitive Neuropsychology,* New York: Psychology Press, 291-320.

Nickels, L. (2002) 'Theoretical and methodological issues in the cognitive neuropsychology of spoken word production', *Aphasiology,* 16(1/2), 3-19.

Nickels, L. A. (1995) 'Getting it right? Using aphasic naming errors to evaluate theoretical models of spoken word production', *Language and Cognitive Processes*, 10, 13-45.

Nickels, L. and Howard, D. (1995) 'Aphasic naming: what matters', *Neuropsychologia*, 33 (10), 1281–1303.

Nickels, L. and Howard, D. (2004) 'Dissociating effects of number of phonemes, number of syllables, and syllable complexity on word production in aphasia: It's the number of phonemes that counts', *Cognitive Neuropsychology*, 21(1), 57-78.

Nickels, L.A. (2004) 'How much semantics is enough? The relationship between comprehension and production in aphasia.' in Murdoch, B.E., Goozee, J., Whelan, B.

M. and Docking, K., eds., *Proceedings of the 26th World Congress of the International Association of Logopedics and Phoniatrics.* International Association of Logopedics and Phoniatrics.

Nyberg, L. (2005) 'Cognitive Training in Healthy Aging: A cognitive neuroscience perspective' in Cabeza, R., Nyberg, L. and Park, D. C. (eds.), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging,* New York: Oxford University Press, 309-321.

Patricacou, A., Psallida, E., Pring, T. and Dipper, L. (2007) 'The Boston Naming Test in Greek: Normative data and the effects of age and education on naming', *Aphasiology,* 21(12), 1157-1170.

Pilgrim, L.K., Moss, H.E., & Tyler, L.K. (2005) 'Semantic processing of living and nonliving concepts across the cerebral hemispheres', *Brain and Language, 94,* 86-93.

Predictive Analytics Software Statistics (PASW) for Windows, Release version 18.0, © SPSS, Inc., 2009, Chicago, IL, www.spss.com.

Randolph, C., Lansing, A. E., Ivnik, R. J., Munro Cullum, C. and Hermann, B. P. (1999) 'Determinants of Confrontation Naming Performance', *Archives of Clinical Neuropsychology*, 14(6), 489-496.

Rapp, B. and Goldrick, M. (2000) 'Discreteness and interactivity in spoken word production', *Psychological Review,* 107, 460-499.

Rapp, B. and Goldrick, M. (2006) 'Speaking words: Contributions of cognitive neuropsychological research', *Cognitive Neuropsychology,* 23(1), 39-73.

Raz, N. (2005) 'The Aging Brain observed in vivo: differential changes and their modifiers' in Cabeza, R., Nyberg, L. and Park, D. C. (eds.), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging,* New York: Oxford University Press, 19-57.

Rivera Mindt, M., Arentoft, A., Kubo Germano, K., D'Aquila, E., Scheiner, D., Pizzirusso, M., Sandoval, T. C. and Gollan, T. H. (2008) 'Neuropsychological, Cognitive, and Theoretical Considerations for Evaluation of Bilingual Individuals', *Neuropsychological Review*, 18, 255-268.

Roberts, P. M., Garcia, L. J., Desrochers, A. and Hernandez, D. (2002) 'English Performance of proficient bilingual adults on the Boston Naming Test', *Aphasiology,* 16(4-6), 635-645.

Roelofs, A. (2004) 'Seriality of phonological encoding in naming objects and reading their names', *Memory and Cognition,* 32, 212-222.

Ross, T. P., Lichtenberg, P. A. and Christensen, B. K. (1995) 'Normative data on the Boston Naming Test for elderly adults in a demographically diverse medical sample', *Clinical Neuropsychologist,* 9(4), 321-325.

Royall, D. R., Cordes, J. A. and Polk, M. (1998) ‘CLOX: an executive clock drawing task’, *Journal of Neurology, Neurosurgery and Psychiatry,* 64, 588-594.

Royall, D. R., Espino, D. V., Polk, M. J., Verdeja, R., Vale, S., Gonzales, H., Palmer, R. R. and Markides, K. P. (2003) ‘Validation of a Spanish translation of the CLOX for use in Hispanic samples: the Hispanic EPESE Study’, *International Journal of Geriatric Psychiatry,* 18, 135-141.

Royall, D.R., Espino, D.V., Polk, M. J., Palmer, R. F. and Markides, K. S. (2004) ‘Prevalence and patterns of Executive Impairment in Community dwelling Mexican Americans: results from the Hispanic EPESE Study’, *International Journal of Geriatric Psychiatry,* 19, 926-934.

Rugg, M. D. and Morcom, A. M. (2005) ‘The relationship between brain activity, cognitive performance and aging: the case of memory’, in Cabeza, R., Nyberg, L. and Park, D. C. (eds.), *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging,* New York: Oxford University Press, 132-154.

Ruml, W., Caramazza, A., Capasso, R. and Miceli, G. (2005) ‘Interactivity and continuity in normal and aphasic language production’, *Cognitive Neuropsychology,* 22(2), 131-168.

Samton, J. B., Ferrando, S. J., Sanelli, P., Karimi, S., Raiteri, V. and Barnhill, J. W. (2005) ‘The Clock Drawing Test: Diagnostic, functional, and neuroimaging correlates in older medically ill adults’, *The Journal of Neuropsychiatry and Clinical Neurosciences,* 17(4), 533-540.

Schmitter-Edgecombe, M., Vesneski, M. and Jones, D. W. R. (2000) ‘Aging and Word-Finding: A Comparison of Spontaneous and Constrained Naming Tests’, *Archives of Clinical Neuropsychology,* 15(6), 479 – 493.

Schramm, U., Berger, G., Müller, R., Kratzsch, T., Peters, J. and Frölion, L. (2002) ‘Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population’, *International Journal of Geriatric Psychiatry,* [abstract available online], 17(3), 254-260, available: http://www3.interscience.wiley.com/journal/91015759/abstract [accessed 13 Sept 2008].

Schriefers, H., Meyer, A. S., and Levelt, W. J. M. (1990) ‘Exploring the time course of lexical access in language production: Picture-word interference studies’, *Journal of Memory and Language,* 29, 86-102.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S. and Sobel, P. (2006) ‘A case-series test of the interactive two-step model of lexical access: Evidence from picture naming’, *Journal of Memory and Language,* 54, 228–264.

Sciriha, L. (1997) ‘One Country, Two Languages’ in Pascoe, R. and Ronayne, J. (eds.) *Malta: A siege and a journey.* Melbourne: Victoria University of Technology.

Sciriha, L. and Vassallo, M. (2006) *Living Languages in Malta,* Malta: Print It Printing Services.

Seigerschmidt, E., Mösch, E., Siemen, M., Förstl, H. and Bickel, H. (2002) 'The Clock Drawing Test and Questionable Dementia: Reliability and Validity', *International Journal of Geriatric Psychiatry,* [abstract available online], 17(11), 1048-1054, available:http://www3.interscience.wiley.com/journal/99016129/abstract?CRETRY=1&SRET [acessed 9 Sept 2008].

Serracino-Inglott, M. (2000) *Id-dizzjunarju Malti u teżawru ta' Malti mħaddem*, Malta: Merlin Library Ltd.

Shah, J. (2001) 'Only time will tell: Clock drawing as an Early Indicator of Neurological Dysfunction', *P & S Medical Review* [online], 7(2), 30-34, available: http://www.cumc.columbia.edu/news/review/pdf/p30-34.pdf [accessed 25 Aug 2008].

Shulman, K. I. (2000) 'Clock-Drawing: Is it the Ideal Cognitive Screening Test?', *International Journal of Geriatric Psychiatry,* 15, 548-561.

Shulman, K. I. and Feinstein, A. (2003) *Quick Cognitive Screening for Clinicians: Mini Mental, Clock Drawing and Other Brief Tests,* London: Taylor and Francis.

Shulman, K. I., Shedletsky, R. and Silver, I.L. (1986) 'The Challenge of Time: Clock Drawing and Cognitive Function in the Elderly', *International Journal of Geriatric Psychiatry,* 1: 135-140.

Silkes, J. P., McNeil, M. R. and Drton, M. (2004) 'Simulation of aphasic naming performance in non brain damaged adults', *Journal of Speech, Language & Hearing Research*, 47, 610-623.

Sörös P., Cornelissen, K., Laine, M. and Salmelin, R. (2003) 'Naming actions and objects: cortical dynamics in healthy adults and in an anomic patient with a dissociation in action/object naming', *Neuroimage,* 19(4), 1787-1801.

Stadthagen-Gonzalez, H. and Davis, C. J. (2006) 'The Bristol norms for age of acquisition, imageability, and familiarity', *Behaviour Research Methods,* 38(4), 598-605.

Statistical Package for the Social Sciences (SPSS) for Windows, Release version 16.0, © SPSS, Inc., 2007, Chicago, IL, www.spss.com.

Storms, G., Saerens, J. and De Deyn, P. P. (2004) 'Normative data for the Boston Naming Test in native Dutch-speaking Belgian children and the relation with intelligence', *Brain and Language*, 91(3), 274-281.

Strauss Hough, M. (2007) 'Incidence of Word Finding Deficits in Normal Aging', *Folia Phoniatrica et Logopaedica,* 59, 10-19.

Strauss, E., Spreen, O. and Sherman, E. M. (2006) *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary,* New York: Oxford University Press.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K. Herron, D., Iyer, G., Jacobsen, T. and Bates, E. (2003) 'Timed picture naming: Extended norms and validation against previous studies', *Behaviour Research Methods, Instruments, & Computers,* 35(4), 621-633.

Tallberg, I. M. (2005) 'The Boston Naming Test in Swedish: Normative data', *Brain and Language,* 94, 19-31.

Tombaugh, T. N. and Hubley, A. M. (1997) 'The 60-item Boston Naming Test: Norms for cognitively intact adults aged 25 to 88 years', *Journal of Clinical and Experimental Neuropsychology,* 19(6), 922-932.

Tsang, H-L. and Lee, T. M. C. (2003) 'The effect of ageing on confrontational naming ability', *Archives of Clinical Neuropsychology*, 18, 81-89.

Tyler, M. D., Tyler, L. and Burnham, D. K. (2005) 'The delayed trigger voice key: An improved analogue voice key for psycholinguistic research', *Behavior Research Methods,* 37 (1), 139-147.

Van Gorp, W., Satz, P., Klersch, M. E. and Henry, R. (1986) 'Normative data on the Boston Naming Test for a group of normal older adults', *Journal of Clinical and Experimental Neuropsychology,* 8, 702-705.

Von Gunten, A., Ostos-Wiechetek, M., Brull, J., Vaudaux-Pisquem, I., Cattin, S. and Duc, R. (2008) 'Clock-Drawing Test Performance in the Normal Elderly and Its Dependence on Age and Education', *European Neurology,* [abstract available online], 60(2), available: http://content.karger.com/ProduketeDB/produkte.asp?Aktion=ShowAbstract&ArtikeIN [accessed 9 Sept 2008].

Wahl, H.-W. and Kruse, A. (2003) 'Psychological gerontology in Germany: recent findings and social implications', *Ageing and Society,* 23, 131-163.

Welch, L. W., Doineau, D., Johnson, S. and King, D. (1996) 'Educational and gender normative data for the Boston naming test in a group of older adults', *Brain and Language,* 53(2), 260-266.

Whitworth, A., Webster, J. and Howard, D. (2005) *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia,* East Sussex: Psychology Press.

Wierenga, C. E., Benjamin, M., Gopinath, K., Perlstein, W. M., Leonard, C. M., Gonzalez Rothi, L. J., Conway, T., Cato, M. A., Briggs, R. and Crosson, B. (2008) 'Age-related changes in word retrieval: Role of bilateral frontal and subcortical networks', *Neurobiology of Aging,* 29, 436-451.

Williams, B.W., Mack, W. and Henderson, V. W (1989) Boston Naming Test in Alzheimer's disease, *Neuropsychologia*, 27(8),1073-1079.

Williams, V. G., Bruce, J. M., Westervelt, H. J., Davis, J. D., Grace, J., Malloy, P. F. and Tremont, G. (2007) ‘Boston naming performance distinguishes between Lewy body and Alzheimer’s dementias’, *Archives of Clinical Neuropsychology,* abstract available: http://www.lewybodydementia.org/forum/viewtopic.php?t=460&view=previous [accessed 10 Oct 2007].

Wilshire, C. E. (2008) ‘Cognitive neuropsychological approaches to word production in aphasia: Beyond boxes and arrows’, *Aphasiology,* 22(10), 1019-1053.

Wilshire, C. E. and Coslett, H. B. (2000) ‘Disorders of Word Retrieval in Aphasia: Theories and Potential Applications’, in Nadeau, S. E., Gonzalez-Rothi, L. J. and Crosson, B. A., eds., *Aphasia and Language: Theory to Practice,* New York: The Guilford Press, 82-107.

Worrall, L. E., Yiu, E. M-L., Hickson, L. M. H. and Barnett, H. M. (1995) ‘Normative data for the Boston Naming Test for Australian elderly’, *Aphasiology,* 9(6), 541-551.

Yoon, C., Feinberg, F. and Gutchess, A. H. (2006) ‘Pictorial Naming Specificity across Ages and Cultures: A Latent Class Analysis of Picture Norms for Younger and Older Americans and Chinese’, *Gerentology International Journal of Experimental, Clinical and Behavioural Gerentology,* 52(5). Abstract available: http://content.karger.com/produktedb/produkte.asp?typ=fulltext&file=GER2006052005295 [accessed 20 Jan 2008].

Zammit Mangion, J. (2000) *L-istorja tal-Edukazzjoni f’Malta*, Malta: PIN.

Zec, R. F., Burkett, N. R., Markwell, S. J. and Larsen, D. L. (2007a) ‘A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test’, *The Clinical Neuropsychologist,* 21, 587-616.

Zec, R. F., Burkett, N. R., Markwell, S. J. and Larsen, D. L. (2007b) ‘Normative Data stratified for age, education, and gender on the Boston Naming Test’, *The Clinical Neuropsychologist,* 21, 617-637.

Appendices

Appendix A

Response form used during both phases of the study

BOSTON NAMING TEST

RESPONSE FORM

No.	English words	Target and Alternative responses*	Responses	Strict score	Lenient score	Comments
1	bed	**sodda**				
2	tree	**siġra**				
3	pencil	**lapes**				
4	house	**dar**				
5	whistle	**suffara**				
6	scissors	**mqass**				
7	comb	**pettne**, moxt				
8	flower	**fjura**				
9	saw	**serrieq**				
10	toothbrush	***toothbrush***, xkupilja tas-snien, xkupilja				
11	helicopter	**ħelikopter**				
12	broom	**xkupa**, mikinsa				
13	octopus	**qarnita**				
14	mushroom	***mushroom***, faqqiegħ				
15	hanger	**spalliera**				
16	wheelchair	***wheelchair***, siġġu tar-roti				
17	camel	**ġemel**				
18	mask	**maskra**				
19	pretzel	***pretzel***				
20	bench	**bank**				
21	racquet	***racquet****, tennis racquet,* paletta tat-tennis, paletta, racquet tat-tennis				
22	snail	**bebbuxu**				
23	volcano	**vulkan**				
24	seahorse	***seahorse****,* żiemel tal-baħar				
25	dart	**dart**(s)				
26	canoe	**kenura**, *canoe*, dgħajsa				
27	globe	**globu**, id-dinja				
28	wreath	**girlanda**, ***wreath***, kuruna				
29	beaver	***beaver***, kastur				

* The target name is listed first (in bold). All English responses were accepted as alternative responses.

No.	English words	Target and Alternative responses	Responses	Strict score	Lenient score	Comments
30	harmonica	***mouth organ***, orgni tal-halq, *harmonica*, organett, fisarmonika				
31	rhinoceros	***rhinoceros,*** rinoċeronte				
32	acorn	***acorn,*** ġandra, ċippitatu				
33	igloo	**iglù**				
34	stilts	***stilts,*** strippi, trampli				
35	dominoes	***dominoes,*** dominò				
36	cactus	**kaktus**				
37	escalator	***escalator***				
38	harp	**arpa**				
39	hammock	***hammock,*** benniena, branda				
40	knocker	**ħabbata**				
41	pelican	**pelikan**, *pelican,* pellikan				
42	stethoscope	***stethoscope,*** stetoskopju				
43	pyramid	**piramida**, *pyramid*				
44	muzzle	**sarima**, *muzzle*				
45	unicorn	***unicorn,*** unikornu				
46	funnel	**lembut**, *funnel*				
47	accordian	***accordian,*** akkordjin, armonju				
48	noose	**ingassa**, ħabel, *noose*				
49	asparagus	***asparagus,*** *asparagu*				
50	compass	**kumpass**, *compass*				
51	latch	***lock***, lukkett, sokra, serratura				
52	tripod	***tripod,*** tripied				
53	scroll	**skroll**, manuscritt, parċmina				
54	tongs	***tongs***				
55	sphinx	***sphinx***, sfinġi				
56	yoke	**madmad**, *yoke*				
57	trellis	**kannizzata**, xeblieka, *trellis*				
58	palette	***palette,*** paletta, tavlozza				
59	protractor	***protractor***				
60	abacus	***abacus***, abaku				

Appendix B

Subjective ratings booklet

Subjective ratings for Age of Acquisition, Familiarity and Imageability

Thank you for accepting to participate in this study. Your time and contribution is extremely appreciated. This part of my study aims to acquire subjective ratings for three variables and to analyse their possible effect on picture-naming performance. The three variables that are being studied here include Age of Acquisition, Familiarity and Imageability.

You will be presented with a list of words which you will rate for each of the above variables according to a particular scale.

Please read the definitions and instructions carefully before you begin. You are kindly requested to be as accurate and honest as possible in your ratings. However, please do not spend too much time on each word. If you have any difficulty please don't hesitate to ask me before you begin.

Please fill in the following details:

Age :	Male/Female :
Area of residence :	Native language(s) :
Name of course :	Year :

AGE OF ACQUISITION

Definition

Age of acquisition (AoA) refers to the age at which words are learned in spoken or written form. By means of this study one aims to determine the approximate age for which a set of words have been acquired.

Instructions

Please rate each word depending on when you think you learned and/or understood it in spoken or written form, **even if** you did not use, read or write it at the time. Please mark the appropriate box according to the age/scale below. The scale is visible at the top of the word list for easy reference (see overleaf). Words that have been learned at an early age should be assigned a low rating. On the other hand, please mark value of *7* for words that you learnt later in life. Please mark **X** for words that you do not know.

	1	2	3	4	5	6	7
AGE (in years)	0 – 2	2 - 4	4 - 6	6 - 8	8 – 10	10 - 12	13+

AGE OF ACQUISITION

		1	2	3	4	5	6	7
		0 – 2 years	**2 – 4 years**	**4 – 6 years**	**6 – 8 years**	**8–10 years**	**10-12 years**	**13+ years**
1	vulkan							
2	kaktus							
3	rinoċeronte							
4	globu tad-dinja							
5	vleġġa							
6	serrieq							
7	*racquet*							
8	żiemel tal-baħar							
9	*escalator*							
10	*pretzel*							
11	*mushroom*							
12	*palette*							
13	ħelikopter							
14	stetoskopju							
15	*yoke*							
16	moxt							
17	asparagu							
18	*seahorse*							
19	spalliera							
20	prinjolata							
21	paletta							
22	*accordion*							
23	*protractor*							
24	dar							
25	sega							
26	lukkett							
27	kunċertina							
28	ingassa							
29	mqass							
30	tripied							
31	siġra							
32	tavlozza							
33	qarnita							
34	kastur							

AGE OF ACQUISTION								
		1	2	3	4	5	6	7
		0 – 2 years	**2 – 4 years**	**4 – 6 years**	**6 – 8 years**	**8–10 years**	**10-12 years**	**13+ years**
35	kannizzata							
36	abaku(s)							
37	girlanda							
38	sfinġi							
39	kwadrant							
40	maqfel							
41	grafometru							
42	bebbuxu							
43	*stirrups*							
44	iglù							
45	akkordjan							
46	*stethoscope*							
47	vara							
48	qubbajt							
49	*hammock*							
50	suffara							
51	piramida							
52	*toothbrush*							
53	lembut							
54	*acorn*							
55	serratura							
56	orgni tal-ħalq							
57	sokra							
58	*asparagus*							
59	madmad							
60	*sphinx*							
61	ġemel							
62	*beaver*							
63	fjura							
64	sarima							
65	mikinsa							
66	pettne							
67	unikornu							
68	faqqiegħ							
69	żebbiġija							
70	ġandra							

AGE OF ACQUISTION								
		1	2	3	4	5	6	7
		0 – 2 years	**2 – 4 years**	**4 – 6 years**	**6 – 8 years**	**8–10 years**	**10-12 years**	**13+ years**
71	organett							
72	sfraġ							
73	globu							
74	*dart(s)*							
75	spraġ							
76	skroll							
77	id-dinja							
78	*unicorn*							
79	ħabbata							
80	xkupa							
81	trampli							
82	figolla							
83	rakketta							
84	knisja							
85	stripep							
86	kumpass							
87	xkupilja tas-snien							
88	branda							
89	*harmonica*							
90	siġġu bir-roti							
91	lapes							
92	maskra							
93	*tongs*							
94	*stilts*							
95	kenura							
96	sodda							
97	skalatur							
98	fisarmonika							
99	pellikan							
100	dominò(s)							
101	luzzu							
102	*pelican*							
103	*wheelchair*							
104	bank							
105	arpa							
106	akkordjin							

FAMILIARITY

Definition

Familiarity is defined as the extent to which one comes in contact with or thinks about a particular word. By means of this study one aims to acquire subjective familiarity ratings for a set of words.

Instructions

Please rate each word for degree of familiarity, depending on how often you think you come in contact with or think about the particular word.

A 5-point scale will be used (see below). Please mark the appropriate box according to the scale. The scale is visible at the top of the word list for easy reference (see overleaf). Number 1 on the scale should be chosen for words that are very unfamiliar, i.e. for items that are never seen, heard or used; number 2 should be chosen for items that one rarely sees, hears or uses; number 5 will be chosen for items that are very familiar and that are encountered nearly every day; etc.. Don't worry if you do not know the meaning of particular words – just rate them depending on how often you come in contact with them. Further, there may be some words which you have used or heard more often than you have seen. Or there may be other words which you have seen more often than you have used or heard. In such cases, always give the highest rating of the three.

Please feel free to use any value from 1 to 5 on the scale when rating the words.

1	**2**	**3**	**4**	**5**
very unfamiliar	**unfamiliar**		**familiar**	**very familiar**
for words that you have **never** seen, heard or used	for words that you have seen, heard or used at least once, but only **rarely**		for words that you have **often** seen, heard or used but not nearly every day	for words that you see, hear or use **nearly every day**

FAMILIARITY

		1	2	3	4	5
		very unfamiliar	**unfamiliar**		**familiar**	**very familiar**
1	skalatur					
2	kumpass					
3	xkupilja tas-snien					
4	fjura					
5	trampli					
6	sega					
7	xkupa					
8	żiemel tal-baħar					
9	maqfel					
10	paletta					
11	figolla					
12	lapes					
13	suffara					
14	faqqiegħ					
15	*beaver*					
16	*stethoscope*					
17	*palette*					
18	lukkett					
19	globu					
20	sarima					
21	lembut					
22	*yoke*					
23	iglù					
24	kaktus					
25	girlanda					
26	spalliera					
27	maskra					
28	kenura					
29	dar					
30	*tongs*					
31	*pretzel*					
32	unikornu					
33	skroll					
34	vleġġa					

Familiarity						
		1	**2**	**3**	**4**	**5**
		very unfamiliar	**unfamiliar**		**familiar**	**very familiar**
35	*mushroom*					
36	*sphinx*					
37	rakketta					
38	kwadrant					
39	qarnita					
40	kannizzata					
41	sokra					
42	arpa					
43	madmad					
44	orgni tal-ħalq					
45	sfraġ					
46	knisja					
47	dominò(s)					
48	*asparagus*					
49	kunċertina					
50	qubbajt					
51	*seahorse*					
52	rinoċeronte					
53	siġra					
54	siġġu bir-roti					
55	id-dinja					
56	piramida					
57	sodda					
58	*racquet*					
59	mikinsa					
60	*stirrups*					
61	żebbiġija					
62	*acorn*					
63	bebbuxu					
64	serrieq					
65	pettne					
66	ħabbata					
67	*wheelchair*					
68	*hammock*					
69	prinjolata					
70	vara					

Familiarity						
		1	**2**	**3**	**4**	**5**
		very unfamiliar	**unfamiliar**		**familiar**	**very familiar**
71	asparagu					
72	kastur					
73	stripep					
74	*unicorn*					
75	Luzzu					
76	*protractor*					
77	abaku(s)					
78	grafometru					
79	mqass					
80	branda					
81	stetoskopju					
82	*toothbrush*					
83	*Stilts*					
84	sfinġi					
85	ħelikopter					
86	*dart(s)*					
87	serratura					
88	*pelican*					
89	fisarmonika					
90	moxt					
91	ġandra					
92	akkordjin					
93	pellikan					
94	tavlozza					
95	akkordjan					
96	ġemel					
97	spraġ					
98	*harmonica*					
99	vulkan					
100	bank					
101	*escalator*					
102	ingassa					
103	globu tad-dinja					
104	organett					
105	tripied					
106	*accordion*					

IMAGEABILITY

Definition

Imageability refers to the degree to which an item evokes mental images. Words differ in their capacity to arouse mental images of things and events. Words that are said to arouse a sensory experience (mental picture or sound) very quickly and easily are said to be high in imageability. Other words may arouse a mental picture only with some difficulty (or after a long delay). These words are said to be low in imageability. Furthermore, the higher the imageability, the greater the number of different images that are aroused when the word is heard. For example, the word *apple* would probably arouse a mental picture very easily, and a number of different images may be evoked (therefore, having high imageability). By means of this study one aims to acquire subjective imageability ratings for a set of words.

Instructions

Please rate each word for degree of imageability, depending on how easy and quickly it is for the particular word to arouse a mental picture.

A 5-point scale will be used (see below). Please mark the appropriate box according to the scale. The scale is visible at the top of the word list for easy reference (see overleaf). Number 1 on the scale should be chosen for words that are low in imageability and number 5 for words with high imageability (i.e. words that arouse a mental picture very quickly and easily).

Please feel free to use any value from 1 to 5 on the scale when rating the words.

1	**2**	**3**	**4**	**5**
low imageability				**high imageability**
for words that are **difficult** to arouse a mental picture and only a **few** images may be evoked				for words that are **easy** to arouse a mental picture and **many** images may be evoked

IMAGEABILITY

		1	2	3	4	5
		low imageability				**high imageability**
1	siġġu bir-roti					
2	xkupilja tas-snien					
3	Abaku(s)					
4	Soda					
5	vleġġa					
6	suffara					
7	dar					
8	rħis					
9	*palette*					
10	kif					
11	pettne					
12	ġemel					
13	mqass					
14	organett					
15	maqfel					
16	maskra					
17	arpa					
18	ġandra					
19	*seahorse*					
20	qarraq					
21	*yoke*					
22	żebbiġija					
23	żiemel tal-baħar					
24	qarnita					
25	stripep					
26	*dart(s)*					
27	*hammock*					
28	sfraġ					
29	figolla					
30	*asparagus*					
31	faqqiegħ					
32	skalatur					
33	sega					
34	*tongs*					

Imageability						
		1	**2**	**3**	**4**	**5**
		low imageability				**high imageability**
35	minn					
36	tripied					
37	liberu					
38	*protractor*					
39	*beaver*					
40	tavlozza					
41	nibex					
42	tattika					
43	*pelican*					
44	spalliera					
45	twemmin					
46	skroll					
47	luzzu					
48	rinoċeronte					
49	bżonn					
50	branda					
51	sarima					
52	blaff					
53	bebbuxu					
54	globu tad-dinja					
55	asparagu					
56	paletta					
57	ingassa					
58	ħajjar					
59	vara					
60	fjura					
61	knisja					
62	*toothbrush*					
63	siġra					
64	lembut					
65	sokra					
66	qubbajt					
67	inqas					
68	kwadrant					
69	*harmonica*					
70	jaf					

Imageability						
		1	**2**	**3**	**4**	**5**
		low imageability				**high imageability**
71	xkupa					
72	fisarmonika					
73	*pretzel*					
74	*sphinx*					
75	kannizzata					
76	unikornu					
77	vulkan					
78	*accordion*					
79	dominò(s)					
80	lukkett					
81	u					
82	saħta					
83	pellikan					
84	stetoskopju					
85	kaktus					
86	sprieġ					
87	ftaħar					
88	ħabbata					
89	tiegħi					
90	*stilts*					
91	akkordjan					
92	*racquet*					
93	lapes					
94	*wheelchair*					
95	ħjiel					
96	kunċertina					
97	*stirrups*					
98	*acorn*					
99	*mushroom*					
100	id-dinja					
101	*stethoscope*					
102	trampli					
103	kastur					
104	kumpass					
105	bank					

Imageability						
		1	**2**	**3**	**4**	**5**
		low imageability				**high imageability**
106	mikinsa					
107	iglù					
108	*escalator*					
109	madmad					
110	grafometru					
111	ħelikopter					
112	*unicorn*					
113	serrieq					
114	prinjolata					
115	rakketta					
116	piramida					
117	kenura					
118	akkordjin					
119	globu					
120	fatt					
121	orgni tal-ħalq					
122	serratura					
123	sfinġi					
124	girlanda					
125	fejn					
126	moxt					

Appendix C

Language History questionnaire

LANGUAGE HISTORY QUESTIONNAIRE

Participant Code no. :	

Pattern of language acquisition*

	M	Me	E	Em	ME
Which language were you **exposed to*** on a daily basis during your pre-school years (**0 – 4 years**)?					
Which language did you **use** on a daily basis during your pre-school years (**0 – 4 years**)?					
Which language were you **exposed to*** on a daily basis during your primary school years (**5 – 10 years**)?					
Which language did you **use** on a daily basis during your primary school years (**5 – 10 years**)?					
Which language were you **exposed to*** on a daily basis during your secondary school years (**11+ years**)?					
Which language did you **use** on a daily basis during your secondary school years (**11+ years**)?					
Which language were you **exposed to*** on a daily basis during your early adult years (**18+ years**)?					
Which language did you **use** on a daily basis during your early adult years (**18+ years**)?					

* This refers to the language used in one's environment, in family and social contexts, and including the language heard on TV.

Language of instruction at school†

	YES	NO
Did you attend an English-speaking school during your primary-school years (5 – 10 years)?		
Did you attend an English-speaking school during your secondary-school years (11+ years)?		

* Sciriha & Vassallo (2006) describe the productive language of the Maltese by using the following taxonomy: (1) Maltese only (**M**), (2) Maltese but with some English words (**Me**), (3) English only (**E**), (4) English but with some Maltese words (E**m**), or (5) Roughly equal degrees of Maltese and English (**ME**).

† Language of instruction may be English or Maltese depending on whether one attends a state school, church school or a private school.

Self-rating of English language proficiency

	1	2	3	4	5	6
	No knowledge at all	Few words or sentences	Can follow a simple conversation or speech	Average ability	Nearly as fluent as a native speaker	Fluent like a native speaker
How well do you understand English?						
How well do you speak English?						

Preferred language in different situations

Which language/s do you prefer (feel most comfortable) using in the following situations?		Maltese	English	Both
* e.g. with family and friends ** e.g. during an interview, addressing a group of people, during mass etc.	casual* speaking			
	casual* listening			
	formal** speaking			
	formal** listening			
	reading			
	writing			
	in general			

Preferred language at home & at work

Which language/s do you use everyday *at home* in the following situations?		Maltese	English	Both
	speaking			
	listening			
	reading			
	writing			
	in general			

Which language/s do you use everyday *at work* in the following situations?		Maltese	English	Both
* e.g. with colleagues in casual conversation ** e.g. during an interview, addressing a group of people, formal conversation relating to work, etc.	casual* speaking			
	casual* listening			
	formal** speaking			
	formal** listening			
	reading			
	writing			
	in general			

Appendix D

Personal details form

QUESTIONNAIRE FOR ADULT VOLUNTEERS

Participant Code no. :	Age:
Gender :	Area of Residence:
Native Language/s:	Handedness:
Level of Education :	**Reading habits** :
Attended primary school	Reads everyday
Attended secondary school	Reads several times a week
Attended post-secondary school	Reads once a week
Attended tertiary education	Rarely reads
Acquired postgraduate qualification	Never reads
Past occupation/s:	Present occupation :
Medical history	

Appendix E

Test of Auditory comprehension of spoken words

FEHIM IL-LINGWA MITKELLMA

Isem ____________________ Data ___________

ID ____________________ Kodiċi __________

	STAMPA TIR	TFIXKIL FONOLOĠIKU	TFIXKIL SEMANTIKU	TFIXKIL MHUX RELATAT	MOD TAT-TWEĠIB	MARKI[1]
P	**mutur**	kulur	rota	weraq	D K R	0 1 2
1	**ġemel**	ġebel	ġiraffa	pinzell	D K R	0 1 2
2	**nemel**	nemex	farfett	mejda	D K R	0 1 2
3	**dar**	ħmar	knisja	larinġa	D K R	0 1 2
4	**xemgħa**	sema	sulfarini	mqass	D K R	0 1 2
5	**vann**	fann	xarabank	siġġu	D K R	0 1 2
6	**ħelu**	felu	ċikkulata	ċavetta	D K R	0 1 2
7	**karozza**	karrotta	mutur	fjura	D K R	0 1 2
8	**sinna**	pinna	xofftejn	sellum	D K R	0 1 2
9	**riga**	rixa	lapes	qalziet	D K R	0 1 2
10	**bolla**	molla	envelop	beritta	D K R	0 1 2
11	**umbrella**	kannella	inċirata	*cushion*	D K R	0 1 2
12	**fellus**	karus	bajda	barmil	D K R	0 1 2
13	**pinzell**	pinzetta	żebgħa	vit	D K R	0 1 2
14	**ħut**	but	xibka	rutella	D K R	0 1 2
15	**ġiżirana**	banana	ċurkett	ħanżir	D K R	0 1 2
16	**bieb**	ktieb	ċavetta	gomma	D K R	0 1 2
17	**ħalib**	salib	jogurt	ċoff	D K R	0 1 2
18	**trakk**	plakk	gaffa	ġobon	D K R	0 1 2
19	**kelb**	deheb	qattus	tieqa	D K R	0 1 2
20	**borma**	borża	taġen	termometru	D K R	0 1 2
21	**bezzun**	beċċun	perżut	*mobile*	D K R	0 1 2
22	**qamar**	baħar	stilla	fjura	D K R	0 1 2
23	**bettieħa**	bettija	ħawħa	ilma	D K R	0 1 2
24	**nagħla**	pala	żiemel	larinġa	D K R	0 1 2
25	**kitla**	xitla	te	basket	D K R	0 1 2
26	**arloġġ**	papoċċ	ċappetta	pizza	D K R	0 1 2
27	**tazza**	mazza	platt	xkupilja	D K R	0 1 2
28	**pudina**	komodina	*trifle*	torċ	D K R	0 1 2
29	**mopp**	kopp	xkupa	*paper clip*	D K R	0 1 2
30	**żarbun**	żarmuġ	kalzetta	važun	D K R	0 1 2

[1] 0 = tweġiba żbaljata; 1 = tweġiba tajba wara repetizzjoni (R), ikkoreġa lilu nnifsu (K), dewmien aktar minn 5 sekondi (D); 2 = tweġiba tajba

Appendix F

Tests of Repetition

RIPETIZZJONI TA' KLIEM

Isem ___________________ Data __________

ID ___________________ Kodiċi _________

	KELMA TIR	TUL TAL-KELMA (f'sillabi)		RISPOSTA	MOD TAT-TWEĠIB	MARKI[1]
P	te	1			D K R	0 1 2
1	il	1			D K R	0 1 2
2	demm	1			D K R	0 1 2
3	xogħol	1			D K R	0 1 2
4	ħin	1			D K R	0 1 2
5	frott	1			D K R	0 1 2
6	blat	1			D K R	0 1 2
7	tina	2			D K R	0 1 2
8	sogħla	2			D K R	0 1 2
9	gawwi	2			D K R	0 1 2
10	ħabel	2			D K R	0 1 2
11	bitħa	2			D K R	0 1 2
12	festa	2			D K R	0 1 2
13	karfus	2			D K R	0 1 2
14	dħaħen	2			D K R	0 1 2
15	krakar	2			D K R	0 1 2
16	lampuka		3		D K R	0 1 2
17	kamoxxa		3		D K R	0 1 2
18	kompjuter		3		D K R	0 1 2
19	sarvetta		3		D K R	0 1 2
20	kampjuni		3		D K R	0 1 2
21	kannestru		3		D K R	0 1 2
22	iljunfant		3		D K R	0 1 2
23	granmastru		3		D K R	0 1 2
24	minjiera		3		D K R	0 1 2
25	fidiferru		4		D K R	0 1 2
26	garaxxijiet		4		D K R	0 1 2
27	biċċerija		4		D K R	0 1 2
28	irnexxielhom		4		D K R	0 1 2
29	sabuhili		4		D K R	0 1 2
30	fetaħhomlu		4		D K R	0 1 2

[1] 0 = tweġiba żbaljata; 1 = tweġiba tajba wara repetizzjoni (R), ikkoreġa lilu nnifsu (K), dewmien aktar minn 5 sekondi (D);2 = tweġiba tajba

RIPETIZZJONI TA' KLIEM FITTIZJU

Isem ___________________ Data __________

ID ___________________ Kodiċi _________

	KELMA TIR	TUL TAL-KELMA (f'sillabi)		RISPOSTA	MOD TAT-TWEĠIB	MARKI[2]
P	pi	1			D K R	0 1 2
1	uċ	1			D K R	0 1 2
2	lusa	2			D K R	0 1 2
3	ħatol	2			D K R	0 1 2
4	ripsel	2			D K R	0 1 2
5	gulirett		3		D K R	0 1 2
6	seltaqi		3		D K R	0 1 2
7	biġowa		3		D K R	0 1 2
8	parjuke		3		D K R	0 1 2
9	faropexu		4		D K R	0 1 2
10	tikamozza		4		D K R	0 1 2

[2] 0 = tweġiba żbaljata; 1 = tweġiba tajba wara repetizzjoni (R), ikkoreġa lilu nnifsu (K), dewmien aktar minn 5 sekondi (D);2 = tweġiba tajba

Appendix G

Personal details and medical history form

MODIFIED BOSTON NAMING TEST FOR MALTESE SPEAKERS

Name ____________ I.D. no. ____________ DOB ________ Age____ Date ____________

Address ______________________________ Tel. nos. ____________

Level of education ____________ Language/s spoken daily ______________________

Years of education ____________ Language/s heard daily ______________________

Marital status ________ Lives with ________ Independence ________ Family support ________

Handedness ________ Ability to read & write prior to illness ______ Current ability to read & write ______

Occupational history ______________________________

Medical diagnosis ____________________ Current SLP diagnosis ____________________

Medical history: Hypertension Diabetes Mellitus Previous CVA Seizures CHF/IHD Asthma COPD

Psychiatric disorders Visual/Auditory difficulties

Date of admission to hospital OR Progress of illness/side affected ____________

Date of neurological insult ____________ ____________________

CT/MRI brain imaging reports ______________________________