Specific Language Impairment (SLI) revisited: evidence from a psycholinguistic investigation of grammatical gender abilities in Brazilian Portuguese-speaking children

Marisa Silveira

University College London

Thesis submitted in partial fulfilment of the degree of Doctor of Philosophy

I, Marisa Silveira, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKOWLEDGEMENTS

I will never be able to thank my supervisors enough for taking me as a student half way through this journey. Ad Neeleman and Shula Chiat have been wonderful and have provided the perfect combination of expertise for undertaking this multidisciplinary project.

A huge thank you to Maria Black, who has read so much of this manuscript and discussed so many parts of it that it would be fair to unofficially nominate her my third supervisor. I am truly grateful for her help.

I am also grateful to Letícia Sicuro Corrêa, my MA supervisor from Brazil, whose work I admire enormously. All the years I spent at PUC-Rio working with her profoundly shaped my way of thinking about language cognition.

I would also like to thank several people at the Division of Psychology and Language Sciences. Chris Donlan and Jane Maxim were always very supportive and extremely efficient when I needed their help. John Swettenham has been a great Graduate Tutor and I am very grateful for his support during the toughest months of this PhD. Sarah Alleemudder, Angela Pozzuto, Steve Newton and Gordon Craig have also provided support of different kinds.

I gratefully appreciate the financial support provided by the Brazilian Ministry of Education and their funding agency Capes, and by the UK higher education funding bodies and their Overseas Research Students award.

This work would not have been possible without the help of several schools and speech and language therapists in Rio de Janeiro, Brazil. Special thanks are due to Olivia Haeusler, Regina Amaral's Clinic, Escola Municipal Albert Schweitzer, Colégio da Providência and Centro Educacional Miraflores. Many thanks to the dozens of lovely children who took part in this study, their teachers and their parents.

Outside the academic world, life in London has been outstanding. I have had a fantastic time in the UK. Cristiane, Dulce, Erika, Laura, Luli, Marina, Thais, Yami, my other Brazilian friends, my neighbours at Goodenough College and my UCL

friends, thank you so much for your friendship, for listening to all my complaints about this PhD and for keeping me sane.

My warmest thanks go to my family. My parents and my brother have provided unconditional support throughout my academic career. They have also helped enormously during the long months of data collection in Rio. And, of course, my wonderful husband, who lived the PhD experience with me, providing constant emotional and intellectual support.

ABSTRACT

This thesis is about Specific Language Impairment (SLI) in children. Its aim is twofold: 1. To provide a theoretical analysis of the field of SLI and discuss its controversies from a novel angle and 2. to present the outcome of a behavioural study which evaluated abilities related to gender agreement in Brazilian Portuguese-speaking children with language impairment. In the first part, I develop a critique of the field, focusing on the multiplicity of conceptions of the term *language* embraced by different disciplines that study the disorder. A critical review of how the field of SLI has developed in recent decades reveals that the conceptual fluctuation in the use of the term *language* has, in many ways, impeded progress in the field. I claim that the only way SLI could be a valid category is if studies focus on basic language skills which, under typical conditions, are acquired spontaneously, without any formal instruction. In Part II, I report an experimental study carried out on the basis of the approach to SLI advocated in Part I. I present a series of experiments that explore the processing of grammatical gender agreement in the Determiner Phrase (DP), in a range of lexical, morphophonological and morphosyntactic conditions in Brazilian Portuguese. Participants were six children with language impairment and 60 typically developing children, including equal numbers from middle class and working class backgrounds. Results showed that gender agreement was very robust for the two groups of typically developing children but problematic for children with SLI, particularly when adjective agreement and gender assignment to novel nouns were involved. The pattern of errors observed and the theoretical discussion throughout Part II suggest that the processing of determiner/noun agreement is a different phenomenon from the processing of noun/adjective agreement, which is vulnerable in children with SLI. In addition, their difficulties with novel nouns suggest that they may require more exposure to input than typically developing children to acquire the gender of nouns.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
INTRODUCTION

This thesis is about Specific Language Impairment (SLI) in children. Its aim is twofold. The first part of this dissertation seeks to provide a critical analysis of the current picture of the research field of SLI. The second part of the thesis presents an experimental study with Brazilian Portuguese (BP) speaking children with language impairment which focuses on grammatical gender agreement within the Determiner Phrase.

The number of studies and researchers investigating the manifestations of SLI has grown enormously in the last decade. A reasonable level of agreement towards what characterises a child with SLI has been developed throughout the years: a child is commonly diagnosed as SLI if his/her process of language acquisition does not follow the normal pattern despite no other apparent cognitive or neurological disorder that may account for their language deficit (Leonard, 1998). Definitions of SLI typically specify that the child must have a substantial discrepancy between language ability and non-verbal IQ (Bishop, 1994). Contrary to many other disorders that affect language, the etiology of SLI is not yet known. The diagnosis of SLI is based mainly on exclusionary criteria: a child should not present any non-linguistic disorders despite the language delay.

English has been, undoubtedly, the most thoroughly studied language in the field. Studies have reported a range of language problems, such as low frequency in the use of embedded sentences, omission of determiners, prepositions, pronouns, plural forms, the genitive *–s*, as well as difficulties with third person singular marking, past morpheme *–ed*, auxiliary forms, reversible passives, wh-questions and argument structure (Clahsen & Almazan, 1998; Leonard, 1998; van der Lely & Battell, 2003; van der Lely, 1998). Researchers have also reported cases of children encountering difficulties with pragmatics, non-word repetition and word finding (Craig, 1991; McGregor and Leonard, 1995; Marshall, 2004).

Although much progress has been achieved in the field in the past few decades, there is still much controversy. Many issues referring to clinical and psychometric aspects of the field are still highly controversial, such as inclusion criteria and cut off scores on standardised tests. In addition, many different

hypotheses have been advanced in an attempt to explain SLI. These hypotheses usually fall into two broad groups: 'linguistic hypotheses' or 'processing hypotheses'. Much work has been carried out about the clinical and psychometric controversies that surround the field and about the hypotheses that have been formulated thus far. This dissertation aims to bring a different perspective into the discussion of the field. It provides an analysis of how SLI research has developed in the recent decades, focusing on the conceptual fluctuation in the use of the term *language* and how this has had an arguably negative impact on the field. On the one hand, one group of researchers works with a broad meaning of *language*, often used interchangeably with the term *communication*. On the other hand, other researchers share a much narrower view, namely that of language as a computational system plus a lexicon, a cognitive component which is acquired by young children without any formal instruction.

The different uses of the term *language* are investigated from a historical viewpoint and through an analysis of the main standardised tests used in SLI diagnosis, in an attempt to shed light into our understanding of some of the field's controversies.

Traditionally, research on SLI has been characterised by a polarisation of positions stemming from disciplines such as linguistics, psycholinguistics and developmental psychology. These related fields have often tended to take opposing approaches, and attempts to explain the nature of SLI have focused heavily on the distinction between 'processing' versus 'linguistic' accounts. Assumptions made on the basis of this distinction have even served as grounds for one of the most popular research questions in the field: 'is SLI a processing or linguistic deficit?'. Such a polarisation is, in my view, misleading, and the assumptions underlying it are not justified. Identifying the sort of language problems children with SLI encounter and describing them from the perspective of linguistic theory is very important, but it is incomplete. On the other hand, attempts to identify the sort of input processing problems children with SLI might have without considering linguistic models that spell out the sort of knowledge that needs to be acquired by the child during the process of language acquisition are essential, but not exhaustive either. The theoretical framework that guides this dissertation is one which seeks to integrate theories within generative linguistics and theories of language processing (Corrêa, 2006). This dissertation

adopts the view that an conciliatory approach to SLI is crucial for a better understanding of the disorder.

In addition to discussing general aspects of the SLI field from the theoretical perspectives mentioned above, the current dissertation aims to investigate gender agreement within the Determiner Phrase in (Brazilian) Portuguese SLI. As Rodrigues (2006) points out, the study of agreement is of considerable importance, as it raises crucial questions about how different sources of information are retrieved and kept in memory during language production; how the flow of information unfolds throughout processing and to what extent a syntactic formulator operates independently from other components of the production system. Furthermore, studies of agreement processing are crucial to a better understanding of SLI, as they can help us pinpoint where agreement errors may occur.

Moreover, Portuguese is a language that has been little studied thus far in the field of SLI. Silveira (2002)[1] provided an extensive preliminary evaluation of characteristics of the disorder in this language. Besides Silveira (op. cit.), to my knowledge, only three other studies have been carried out on the manifestations of SLI in Portuguese. Macacchero (2004) investigated the use of functional categories of tense and aspect by two Brazilian children with SLI. Haeusler (2005) investigated argument omission in three Brazilian children with SLI. And, finally, Hermont (2005) has further explored the issue concerning tense and aspect with a single case study.

Gender is considered the most puzzling of the grammatical categories (Corbett, 1991). Grammatical gender involves two or more items sharing a feature, controlled by the Noun. It is not a universal feature of human languages and the languages that do have a gender system manifest it in different ways: the type of agreement relations that are overtly expressed in the elements other than the Noun varies across languages. In the Portuguese Determiner Phrase, grammatical gender is manifested in determiners, nouns and many, but not all, adjectives. As regards acquisition, a puzzling phenomenon occurs. While gender systems can pose major problems for second language learners and even

[1] Dissertation which was part of a large research project at the Psycholinguistic and Language Acquisition Laboratory (LAPAL) at Pontifícia Universidade Católica at Rio de Janeiro and was financed by FAPERJ (Foundation for Research Funding of the State of Rio de Janeiro). The project aimed to devise the language abilities test referred to as MABILIN (see chapter 6).

advanced speakers continue to often make mistakes over the years, young children tend to acquire the gender system of their native language without any major problems and make hardly any errors.

The thesis is organised as follows: chapter 2 presents a thorough analysis of the research field of SLI, carried out from a wide perspective, in an attempt to understand the controversies that surround the field in an effective manner. Historical features of SLI research are addressed and the validity of the term 'Specific Language Impairment' itself is evaluated. Two alternative scenarios for the field are introduced. I argue that an approach to the field which is theoretically motivated and which focuses on SLI as a disorder of basic language skills which, under typical conditions, are acquired spontaneously, without any formal instruction, is the only way SLI could be considered a valid category.

The second part of the thesis is subdivided into the following chapters: in chapter 3, the theoretical issues that underlie the behavioural study subsequently presented are discussed and a proposal on how to conciliate linguistics and psycholinguistics based on the work of Marr (1982) is presented. In chapter 4, grammatical gender is thoroughly reviewed. First, I define the phenomenon of grammatical gender in human languages and look at how generative linguistics has been dealing with several aspects of gender, from its characteristics in Brazilian Portuguese to the different theoretical accounts of gender agreement available in the literature. In chapter 5, I carry out a review of how research on SLI, typical language acquisition and adult language processing has been investigating gender agreement. In chapter 6, I describe the criteria and measures that were taken in recruiting participants and the profile of the children who took part in the experimental study, which is reported in chapter 7. Finally, chapter 8 recapitulates the main issues addressed throughout the thesis and presents an evaluation of the outcome of the study.

PART I
CRITIQUE OF SLI

Chapter 2
ANALYSIS OF THE FIELD

2.1 Introduction

SLI research has made considerable progress in the last couple of decades. Interest in the disorder has grown enormously, researchers with different backgrounds started to look at SLI, and the number of investigations carried out in an attempt to understand the phenomenon has greatly increased. Nevertheless, there is still much controversy in the field. Most of the debate, however, is concentrated in controversies surrounding clinical and psychometric issues. For example, there is no consensus about which tests and which cut off scores to use with potential SLI cases, nor about which cognitive components to include in test batteries. In addition, given the heterogeneity typically observed in groups of children diagnosed with SLI, some researchers question the existence of a single disorder and argue in favour of different subgroups of SLI. Furthermore, researchers disagree with respect to the preservation of non-verbal abilities in children with SLI, given extensive data suggesting weaknesses in areas of functioning that fall outside language cognition (Leonard, 1998). In this chapter, I will discuss the controversies mentioned above in a different way. I will address the research field of SLI in a wider perspective, bringing together views from various disciplines in an attempt to understand these controversies more effectively. In particular, I will argue that at least part of the dispute in the field originates from the variability of interpretation of the term *language*: researchers from different disciplines and backgrounds are discussing SLI studies as if they were dealing with the same phenomenon but this does not seem to be the case. I will show that different conceptions of the term *language* are present in these studies. On the one hand, a large number of researchers work with a broad definition of *language*, often used interchangeably with the term *communication.* For instance, these researchers validate, as language problems, difficulties such as those of adequately placing an utterance in a social context or failing to provide the correct answer to a narrative comprehension task. On the other hand, other researchers work with a much narrower definition of *language*, according to which children with SLI should typically present difficulties such as producing well-formed sentences and establishing grammatically relevant distinctions.

These distinct conceptions of *language* have major implications for several other aspects of SLI research, contributing to the creation of a blurred picture. In the following pages, I present a brief historical overview of SLI research and the main characteristics of the disciplines that have been studying its manifestations, showing how the different notions of *language* impact on the way research is carried out. Through a detailed analysis of some popular standardised tests used in SLI studies and an evaluation of the theoretical assumptions that serve as a basis for these tests, I will show that the SLI 'label' is not just heterogeneous, as some researchers claim, but a label which is being used to group together children encountering difficulties that are largely unrelated to each other. I will end the chapter proposing two potential alternative scenarios for the field of SLI. I will argue in favour of a moderate version of a narrow approach to SLI, claiming that this approach is potentially more productive (at least as a starting point of any investigation) and the only alternative for SLI to be a truly valid category. Future investigations will be able to select between these two scenarios.

2.2 Historical overview of SLI research and conceptions of language according to different disciplines that study its manifestations

Traditionally, children with language impairment have been studied from a clinical perspective. Studies date to as early as mid-nineteenth century (Wilde,1853; Benedikt, 1865; Waldenburg, 1873 and others, apud Leonard, 1998), when physicians would report cases of children whose 'language abilities' were impaired and whose 'non-verbal intelligence' seemed to be intact. The terminology used through the decades varied largely. Terms such as *delayed speech development*, *congenital aphasia*, *infantile aphasia*, *developmental dysphasia* have been frequently used in the past. Towards the end of the 20th century, terms with a neurological connotation such as *congenital aphasia* and *developmental aphasia*, for example, become outdated and start to be used to refer to deficits caused by cerebral damage. The term *Specific Language Impairment* begins to be used more widely. SLI is currently a very popular research topic among investigators in speech and language therapy, special education, psychology and linguistics in some countries.

It is clear from reading the literature coming from clinical settings that the term *language* is conceived in a broad sense, close to what the term *communication* generally conveys. In other words, *language* is conceived as a set of abilities used to interact and communicate. For instance, in a paper

published by the American Speech-Language-Hearing Association, Culatta and colleagues (1983) use the terms *communicative performance* and *language* interchangeably. In the abstract, the authors state that their study "investigated the use of a story retelling task as a mechanism for screening *integrated communicative performance*". Further on in the text, they use the term *language* instead: "The purpose of this study was to determine the feasibility of story retelling as a *language* screening device" (my italics). Another example of the broad definition of *language* can be found in Bloom (1991): "... knowing language includes knowing what one can and cannot say to different people in different circumstances and even knowing when one can talk or not talk at all". Under this conception of *language*, SLI studies may include pragmatic, social interaction and other abilities related to how language is used socially: "Pragmatic capabilities are typically included in most diagnostic schemes of language impairment" (Tomblin et al, 1996). As will become clear throughout this chapter, this broad use of the term *language* is also often employed together with terminology referring to educational settings.

In a later stage in the history of SLI research, scholars coming from the discipline of linguistics, in particular those involved with Chomsky's generative theory, were attracted to the SLI field. Generative theory paved the way for the cognitive revolution that took place in the 1960s, which changed the manner in which the human mind was investigated. The notion of modularity of mind was introduced, and the possibility of mental phenomena arising from the operation of multiple distinct processes, rather than a single undifferentiated one, started to be explored (see Fodor, 1983; Barrett & Kurzban, 2006). A lot of effort was put into the investigation of potential *selective disorders*, such as different types of aphasia (Grodzinsky, 1990) and the cases of *savants* (Smith & Tsimpli, 1995). In this context, SLI seemed to provide very appealing evidence for the hypothesis that the mind is organised in semi-independent modules. An approach to SLI within the generative linguistics framework naturally leads to a narrower interpretation of the term *language* than in the clinical and educational contexts. The term *language* is used to refer to an internal component of the mind/brain (sometimes called "internal language" or "I-language"), composed by the lexicon and a computational system common to human beings (Hauser, Chomsky & Fitch, 2002), which is responsible for generating phrases. Moreover, the term is used, within this context, to refer to language as a basic human capacity, put into use by any typical child with enough exposure to a language community. Within

generative linguistics, *language* refers to knowledge which is acquired by any typically developing child naturally, without any formal instruction. Pragmatics and social interactive skills are not considered to be part of the language domain.

Explicit examples of how differently *language* is conceived within SLI research can often be found in the literature. The work of Clahsen, for instance, is very much oriented to the generative linguistics framework and, thus, his interpretation of *language* and, consequently, his definition of SLI is a narrow one: "SLI is a condition of disordered or delayed language acquisition which is characterized by severe problems in the normal development of morphosyntax in subjects who did not seem to have any clear non-linguistic deficits" (Clahsen, University of Essex webpage). Van der Lely and colleagues take a similar approach, as notions of generative linguistics have been incorporated in their research and the focus of their investigations is on aspects relating to phonology, morphology and syntax (van der Lely, 2005a). Both Clahsen's and Van der Lely's work attempt to provide a hypothesis for SLI making direct use of linguistic constructs formulated within generative linguistic theory, i.e. linguistic models are directly used with the intention of explaining the manifestations of SLI in children. On the other hand, whilst some investigators recognise the possibility that pragmatic problems might be secondary to more strictly linguistic problems, others seem to view pragmatics as an important line of research in SLI: "Systematic examination of the pragmatic skills of children with SLI who differ in receptive language skills may be warranted" (Craig & Evans, 1993: 779). An even looser meaning of the term *language* is found in Evans (2001). The author, when describing the characteristics of the profile of children with SLI, states that their "difficulties range from deficits in vocabulary and word-finding, to impairments in morphology, syntax, pragmatics, nonverbal and verbal working memory, slower verbal and nonverbal processing, and deficits in speech perception" (p. 40).

Researchers' views towards the current state of SLI research vary enormously. Some investigators are aware of the controversial issues that surround the field and recognise that SLI is not an established fact: "it appears that many persons involved with children with language impairments have accepted the concept of specific language impairment as an established fact, rather than recognizing that it more accurately represents an hypothesis in need of testing and validation" (Aram et al, 1993: 582). A similar view is present in Tomblin et al (1996: 126): "… the issue of what aspects of language should be

tested remains. There are no established guidelines concerning the specific areas of language that should or must be examined within the diagnosis of SLI in children". Nevertheless, many researchers conduct their studies taking for granted the existence of SLI as such and simply use the diagnostic tools and exclusionary criteria available in the literature. It is, therefore, common to find statements in the literature which reflect this lack of questioning: "The prevalence of SLI is about 7%" (Leonard, 1998: 3); "SLI affects about 7% of the population" (van der Lely, 2004: 119); "SLI has a genetic component …" (van der Lely, 2005b: 13). These statements are not necessarily incorrect, but they rest on the assumption that SLI is an agreed phenomenon.

2.3 SLI within the context of the study of Language Acquisition

In order to understand the origins of the variability surrounding the interpretation of the term *language* and its impact on the development of research on language impairments, one needs to get acquainted with how the field of language acquisition (henceforth LA) in general unfolded in the past decades. In many respects, studies on SLI reflect the different theoretical assumptions that oriented the work on LA in the late decades of the 20th century. The aim of the current section is, hence, to identify, in the recent history of LA, those elements which can potentially shed some light on the controversies that surround the study of SLI. This section is, by no means, a comprehensive and detailed historical panorama of LA. The discussion that follows draws heavily on the survey and analysis conducted by Corrêa (1999).

In the 1960s, the field of linguistics saw a major change in its underlying principles: Chomsky's generative grammar incorporated a cognitive approach to the field that differed from the descriptive approach that was predominant at the time. Under this new perspective, generative grammar was committed to describing linguistic knowledge of adult speakers, as well as postulating a model of linguistic knowledge that can be acquired by any child in normal circumstances: "Thus the goal was to write grammars of languages that corresponded to those that were represented in the brain/minds of the speakers rather than simply elegant descriptions of linguistic patterns" (Fromkin, 1997: 4). In order to account for the fact that any normal children can acquire the language of their community easily and rapidly, as well as relatively uniformly across different languages, Chomsky proposes that children are born with an innate

capacity for acquiring a language, i.e., with a device that would interact with input of a particular language during the process of language acquisition.

The Linguistic Revolution (and, more generally, the Cognitive Revolution) that took place at that time had a major impact on the research area of LA. The "logical problem of language acquisition", as formulated by generative linguistics, received attention from different disciplines.

As Corrêa (1999) points out, the main controversy surrounding the investigations on LA refers to how much information about the grammar of human languages could be attributed to a predetermined biological program characteristic of our species. Such controversy reveals itself in a variety of topics: in the way the relationship between language development and cognition is tackled, in the manner the relationship between syntax and semantics is conceived in the development of language, the form hypotheses about the necessary interpersonal and environmental conditions for language acquisition are put forward, and the nature of the proposed acquisition procedures.

In the 1970s, within generative linguistics, what is known as 'learnability theory' looked at LA from an abstract point of view. The focus was on the properties of formal systems that could be identified by what came to be called "Language Acquisition Device", i.e., the aim was to formulate models of grammars that were potentially suited for acquisition by a newborn. Parallel to this abstract perspective, a less formal approach was also in place, which posited hypotheses directly based on data from the linguistic performance of children. This approach proved problematic, for example, because data were taken as directly reflecting the language competence of children, without taking into consideration any elements of performance systems (Corrêa, op. cit.).

Early in the 1980s, progress within learnability theory led generative linguistics to propose a model of universal grammar (UG) which was formulated in terms of universal principles and parameters to be set. According to this proposal, human languages are ruled by a set of invariable principles common to human kind, accounting, therefore, for the universality of human languages. Variability, on the other hand, is treated in terms of parameters that need to be set, i.e., variables that are assigned a positive or negative value during the process of language acquisition (Chomsky, 1981, 1986). On the basis of the

principles and parameters theory, the ‘problem of language acquisition’ came to be seen as a matter of ‘parameter setting’ and lexical acquisition. A note about the lexicon needs to be made: the acquisition of new lexical items can last until adulthood, even if it has its peak at around 2 years of age. It is, thus, possible to conjecture that lexical acquisition is a process which is ruled by different factors than the ones operating in the acquisition of syntax and phonology (Corrêa, op. cit.). Following the introduction of the principles and parameters approach, a second generation of language acquisition studies under the framework of generative linguistics took place. Based mainly on spontaneous data, researchers put their efforts into testing hypotheses about what characterises the principles of UG and what is involved in the setting of parameters.

Within developmental psychology, research on LA took a more concrete path. Brown (1973, apud Corrêa, op. cit.) and Bloom (1970, 1973, apud Corrêa, op. cit.), for example, provided the field with longitudinal accounts of the acquisition process of English, somewhat resuming the old tradition of baby diary studies of Stern and Stern (1907, apud Corrêa, op. cit.) and Leopold (1939-49, apud Corrêa, op. cit.), for example. The longitudinal accounts of the 1970s, however, were linguistically more informed than the diaries of the first half of the century. Nevertheless, linguistic theory in those diary accounts was used mainly as a descriptive tool for the presentation of data, and not as a hypothesis about the nature of LA course of development. Furthermore, in the use of generative linguistics as a descriptive tool, it was not unusual to witness the use of theoretical terms introduced within linguistics with a very distinct, rather misleading, reading in developmental psychology, triggering a conceptual fluctuation that hampered the debate. For example, many developmental psychologists (including Wiig and Semel, co-authors with Secord of the CELF test, still widely used nowadays in the identification of children with SLI), use Chomsky’s ‘surface structure’ and ‘deep structure’ concepts in a misleading way, claiming the former reflects the syntactic properties of a sentence and the latter reflects its meaning (see Crystal (1997) for accurate definitions of the terms). I will come back to this issue further in the chapter, when I discuss the validity of standardised tests used with children.

Unlike generative linguistics, developmental psychology over this period was crucially child-centred: “A developmental perspective assumes that children play an active part in acquiring language. They are, in effect, ‘the agents of their

own development" (Bloom, 1991[2]). In addition, it is possible to say that the rationalistic assumptions present in the proposals of generative linguistics and the idea of a modular biological basis to account for the initial state of human cognition did not trigger a positive reaction from developmental psychologists, who belonged to a long empiricist tradition that oriented the field, towards a generalist view of the mind of newborns and argued for the idea that language was necessarily dependent upon the development of cognition in general. The quotes below, from Bloom (1991) are good illustrations of this viewpoint: "The human infant begins life in an essentially global and undifferentiated state" (p. 3); "Semantics, syntax and discourse are not separate for children learning language. Rather, they learn these aspects of language together, from the beginning. For these reasons, the studies in this book have an integrative perspective and address the contact among semantics, syntax, and discourse in the course of acquisition" (p. 23); "Language development, in this view, follows from and depends upon conceptual development in a logical way - as traditionally argued by Piaget (1954) …" (p. 44).

Translations of Piaget's works on cognitive development began to arrive in the United States from Europe. At the same time, Chomsky's work would travel in the opposite direction, to a Europe where a long Functionalist tradition within linguistics and a strong empiricist tradition within philosophy were in place. In this context, the relation between language and cognition, on the one hand, and between language and communication (or social interaction), on the other hand, took a major part in the discussions of psychologists, linguists and philosophers (Corrêa, op. cit.).

The different approaches and viewpoints that came to guide the research on LA at that time placed the field amidst many controversies. As Corrêa (op. cit.) points out, besides the more obvious debate about innateness, other discussions built around the focus (or the lack of) on syntax and semantics: generative linguistics of the 60s concentrated its LA studies on syntactic aspects, while developmental psychologists put their efforts into examining concepts and semantic relations expressed in the child's speech and put forward the hypothesis of semantics preceding syntax in the course of language development. Additional controversies surrounded the discussion regarding the

[2] Lois Bloom's book *Language Development from Two to Three* was published in 1991, but contains work originally published between 1970 and 1989.

linguistic input which is available to the child in the first stages of LA. On the one hand, generative linguistics put forward the 'poverty of the stimulus' argument, and, on the other hand, developmental psychologists concentrated on characterising what came to be known as 'child-directed speech' or 'motherese'. Another controversy concerned the type of data investigated: while some focused on language development starting with speech signal discrimination within the first few days of birth, others viewed language development on the basis of pre-linguistic communicative or pragmatic abilities[3].

It was within the controversial context outlined above that the contemporary multidisciplinary study of SLI had its start. Indeed, we will see, further in this chapter, that the unfolding of the research field of SLI reflects, at least partially, the development of the field of language acquisition more generally. Many of the language assessments tests that are still used nowadays were originally conceived and published during the 70s and 80s, within the controversial context discussed above. Additionally, we will see that many of the misunderstandings and misconceptions that pertain to the study of SLI can be explained (but not necessarily justified) from a historical perspective. Such misunderstandings and misconceptions have, in many instances, impeded progress in the field.

2.4 Diagnostic procedures of SLI: do they lead to a consistent group?

The different conceptions of *language* impact on a very basic issue, namely, procedures for diagnosing children with SLI and selection criteria for inclusion in SLI studies. The citation that follows is a good example of the problems surrounding this issue: "Standardised test scores serve only as the starting point. The work of actually describing and explaining these children's language functioning must then begin. In large part because this more detailed analysis is to be conducted, the standardized language tests used as inclusionary criteria can be quite broad in scope" (Leonard, 1998: 11). Needless to say, standardised test scores do not provide a complete picture of the language profile of any child. Nonetheless, only a highly informed and theoretically-oriented diagnostic procedure will pave the way for good and clarifying experimental studies. Therefore, the weight given to experimental studies should not be so heavy as Leonard's statement suggests. More attention needs to be given to diagnostic

[3] See Corrêa (1999) for references illustrating the types of work mentioned here.

tools of SLI. For example, the decision as to whether or not to include pragmatic skills assessment should be based on theoretical assumptions regarding cognitive demands and the way language and cognition in general are structured, not on methodological shortcomings, as mentioned by Tomblin and colleagues: “Due to the methodological difficulties of testing pragmatics in a norm-referenced setting … a pragmatic dimension was not included in the EpiSLI diagnostic scheme ...” (Tomblin et al. , 1996: 1287). Tomblin and colleagues recognise that pragmatics is not an area of primary language deficit for SLI population, but the very fact that they considered including pragmatic abilities in their diagnostic scheme is alerting.

The matter of diagnostic tools presents yet more problems. As it is widely known, SLI diagnosis depends heavily on exclusionary conditions, which is not ideal: “One of the banes of professionals who diagnose SLI is that it is a diagnosis based as much on exclusion as on inclusion” (Leonard, 1998: 10). Another set of criteria refers to the discrepancy between the child’s achieved language status and some standard of expectation for the child’s language status. Even if some common guidelines have been reached in the past decades, there is still quite a lot of variation and studies differ considerably with respect to cutoff scores in standardized tests, selection of tests, selection of typically developing control groups, etc.

Regarding non-verbal abilities test scores, many studies state that children need to obtain a nonverbal IQ score of at least 85 in order to be considered SLI (Leonard, 1998), but some researchers define 80 as their cut off score (Aram, Morris & Hall, 1993). In addition to scoring relatively high on non-verbal tests, children need to obtain a relatively low score on language tests, showing a “gap” between the two domains. Transforming something that is a continuum into discrete categories is not an easy task. As noted by Tomblin et al (1996: 1285), “the size of the discrepancy between language achievement and chronological or mental age expectations necessary for determination of SLI has usually been arbitrarily set.”.

Further, a variety of IQ tests is used by researchers, such as the Raven’s (missing segment completion task, Raven, 2003), the WISC (Wechsler Intelligence Scale for Children, Wechsler, 1974), the K-BIT (Kaufman Brief Intelligence Test, Kaufman & Kaufman, 1990) and the CMMS (Columbia Mental

Maturity Scale, Burgemeister, Blum & Lorge, 1972). A wide range of 'language' tests is also used depending on the research group carrying out the study, such as the CELF (Clinical Evaluation of Language Fundamentals, Semel, Wiig & Secord, 1995), the Peabody (Dunn, 1965, a vocabulary test), the BPVS (British Picture Vocabulary Scale, Dunn, Dunn, Whetton & Burley, 1997, a vocabulary test), and the TROG (Test for Reception of Grammar, Bishop, 2003). Considering the measurement errors[4] that may affect these tests and the natural variation that results from the fact that different tests are used with different children, employing those tools for a diagnosis of SLI does not seem to provide a very reliable population to be used in investigations. As we will see in chapter 6, one of the children recruited for this study, CA, provides an example of the instability of diagnosis of SLI based on current criteria. CA was assessed with two different non-verbal tests: she passed one, but failed the other. Therefore, it is possible to say that different studies may include quite distinct populations, yet findings across studies are compared as evidence regarding a supposedly agreed category of SLI.

Next, I discuss one of the most widely cited papers in the SLI literature, namely Tomblin and colleagues' 1997 article in the *Journal of Speech, Language and Hearing Research*. The discussion of the paper will lead to an analysis of some of the main language assessment tests used in the clinical context in the English-speaking world and elsewhere via translations. In particular, I look into the validity of the tests and propose that they are not evaluating what they claim to be evaluating.

2.5 Prevalence of SLI – an illustration of the importance of definition and operationalisation

The critical discussion of the identification of SLI raises issues for the percentage of children estimated to be affected by the disorder. The estimates reported in the study of Tomblin et al (1997) are widely cited in the literature. According to this study, around 7% of school age children in the USA have SLI. If this is correct, SLI is a frequent condition, affecting at least 1 child per standard size classroom

[4] These tests are administered by people (teachers, therapists, etc.) and children in the test population are normally assessed in different environments (e.g. schools). Hence, it is likely that both some specific characteristics of the 'interviewer' (e.g. different degree of training) or of the environment (e.g. more or less noisy schools) may introduce a systematic error of measurement. Accordingly, the degree of measurement error may not be negligible, with this possibly biasing the test results.

(around 30 students). However, a look at the studies by Tomblin and colleagues, with special attention to the types of tasks employed by them, reveals some drawbacks. In Tomblin et al (1996), the authors seek to provide a rationale for a 'valid and reliable diagnostic system for Specific Language Impairment' (labeled as EpiSLI system). They employed five composite scores representing norm-referenced performance in what they consider three domains of language (vocabulary, grammar, and narration) and two modalities (comprehension and production). Those children with two or more composite scores below - 1.25 standard deviations were considered as children with language impairment. Two language measures were used: 1) the Test of Language Development (TOLD-2:P) (Newcomer & Hammil, 1991), to 'provide measures of receptive and expressive vocabulary (picture vocabulary, oral vocabulary) and sentence usage (grammatic comprehension, sentence imitation, grammatic completion)'; 2) 'narrative comprehension and production screening test' of Culatta et al. (1983).

2.5.1 The Test of Language Development (TOLD)

An evaluation of TOLD's rationale and types of stimuli selected for its subtests reveals inconsistencies that lead to a debatable 'language' assessment measure. The authors state, at the start of the test's manual, that they "chose to use a linguistic model as a theoretical base" (p. 1). However, they do not specify what they mean by the use of the term 'linguistic'. Instead, they explain that they "did not adhere to any specific theoretical perspective" (p. 1), but incorporated "the contributions of a variety of esteemed linguists and psycholinguists" (p. 2). Newcomer and Hammil then provide a list of works they claim have contributed to constructing the TOLD, which includes several publications of Chomsky from between 1957 and 1981, but also works that belong to a distinct tradition, such as Lois Bloom's and Roger Brown's publications, which posit some assumptions about the nature of language acquisition that are incompatible with those of Chomsky's work (see pages 21 and 22 of the current thesis for a quote which illustrates Bloom's view on language). Indeed, towards the end of the test's manual, the authors provide some clear indication about their assumptions regarding language and its relation to the other cognitive domains: "Language is an important aspect of general intelligence. This is evidenced by the fact that most tests of intelligence incorporate many language elements into their contents and formats" (Newcomer & Hammil, 1991: 111). This shows that Newcomer and Hammil's 'use' of Chomsky's work is likely to reflect a tendency at the time the test was originally published (1977), to attempt to use generative linguistics as a

descriptive tool for dealing with data, not as a hypothesis about the nature of language development. Therefore, it is not clear what the authors mean by 'linguistic' when they say they chose to use a linguistic model as a theoretical base.

Moreover, Newcomer and Hammil explain that the words used in their vocabulary subtests were selected from a corpus which was compiled on the basis of frequency in written sources. Considering that the authors say that the TOLD is suitable for testing children as young as 4 years old and state, several times throughout the manual, that the TOLD tests oral language abilities, selecting words on the basis of written sources seems very inappropriate. Further, not only did the authors use lexical items from a written corpus, but they did so from a list prepared by Thorndike and Lorge in the 1940s. Surprisingly, Newcomer and Hammil used this corpus on the TOLD without any questioning, and Tomblin et al use the TOLD for their EpiSLI system without any questioning either. Knowing the source of the lexical items used in the vocabulary subtests, it is then unsurprising that many of them are formal and dependent on schooling. Indeed, Thorndike and Lorge's corpus was created as a resource for elementary and high school teachers in the United States (cf. Bauman, 1996).

Next I discuss some of the subtests of the TOLD in an attempt to illustrate the problems just raised and point out some other potential difficulties the test presents.

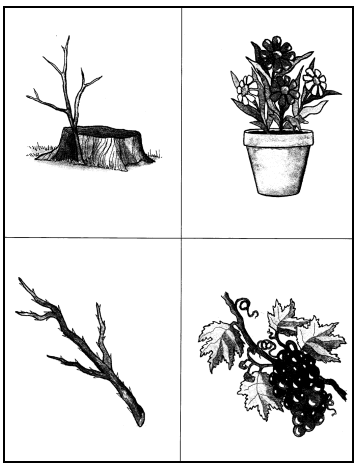
2.5.1.1 The TOLD's vocabulary subtests

The subtest *picture vocabulary* is a simple picture pointing task, in which the child needs only to point to one of the four pictures that best represents the meaning of a word spoken by the examiner. It includes nouns, verb, adjectives and prepositions. A major problem of this subtest is that a large number of the lexical items are highly dependent on formal schooling or on world knowledge that must be gained through particular life experience since the items are of low frequency in spoken language. For example, 'infirm' and 'abode' are very formal items which are most likely very rare in oral language whilst 'salmon' is a noun whose learning might be dependent on how much fish the child is exposed to, something that is potentially variable. Therefore, an incorrect response is not necessarily indicative of a deficit in the child's lexicon, since it is plausible that some teenagers or even some adults without any language impairment might not

know the meaning of these lexical items. Considering that many of the lexical items selected for this subtest can vary according to the child's world knowledge or access to formal schooling (which do not follow a standard developmental pattern), as exemplified by the items cited above, the validity of the subtest *picture vocabulary* as a task reflecting basic language skills is questionable.

The subtest *picture vocabulary* has additional potential problems. The authors deliberately include a few adjectives on their list of lexical items, but problems in the selection of pictures to accompany the adjectives nullify their selection to test the grammatical category 'adjective'. For example, the adjective 'floral' is used, but the picture that represents it is not adequate for testing the knowledge of an adjective. As figure 1, reproduced from the TOLD shows below, on this occasion, the child needs to select a flower pot among a total of four unrelated pictures. By pointing to the flower pot, the child is not necessarily showing that she knows what 'floral' represents as a noun modifier, as the knowledge of the noun 'flower' is probably enough to succeed in this instance. In order to turn this test item into an adjective item, a picture of a piece of fabric with a floral pattern, for example, could have been used, alongside a picture of a flower pot, creating a situation in which the child would need to choose between the two related pictures based on her knowledge of the adjective 'floral' as opposed to noun 'flower'.

Figure 1: pictures used on the TOLD's *picture vocabulary* subtest; test item is the adjective 'floral'



The same applies to the adjective 'dental'. The picture chosen to represent it shows a little girl brushing her teeth. Again, the instance is not testing an adjective, as an association with the lexical item 'dentist' is enough to get it right.

Figure 2: pictures used on the TOLD's *picture vocabulary* subtest; test item is the adjective 'dental'

More generally, it does not seem much attention was given to the selection of the distractor pictures. In a few instances, there seems to be some sort of control in order to have distractor pictures representing lexical items that are somehow related to the target item, but, in the majority of cases, distractor pictures have no relation to the target items. For example, the distractors for the target noun 'bulb' are a dumbbell, a clock and a feather.

The subtest presents several other isolated shortcomings in relation to its selection of pictures to represent target lexical items and pictures chosen as distractors. The items 'medical', 'infirm' and 'feeble', for example, have common problems. The target picture for 'medical' is a doctor who could be better represented pictorially. The accompanying distractor pictures are: 1. a girl with a bird resting on her arm; 2. an old man wearing pyjamas and walking with a cane; and 3. a lady drawing on a desk. Here, the target picture is the doctor, but the old man with the cane gives the impression he is ill and could, then, potentially be chosen under the term 'medical' as well, for being in a 'medical condition'. Indeed, the same picture of the old man is used as target picture for the adjective 'infirm'. In addition, the old man is shown again, as a target item for the second time, for the adjective 'feeble'. In sum, the subtest *picture vocabulary* contains several elements that allow us to challenge the validity of its use as a test of basic language skills, as many factors other than primary language abilities seem to be involved. The following paragraphs will show that the same sort of criticism applies to other subtests of the TOLD.

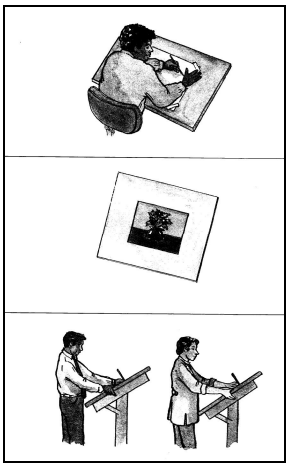
In the second vocabulary task, the subtest entitled *oral vocabulary*, children are asked to give definitions of words and hear the following instructions from the experimenter: "I am going to say some words and I want you to tell me what each word means …". The selection of lexical items on this task is not so problematic as the previous task, as most nouns seem to be relatively frequent in spoken language (and, therefore, not so dependent on formal schooling) and the absence of pictures prevents many of the problems found in the *picture vocabulary* subtest. However, the task itself can be criticised. Giving definitions of words does not seem like an appropriate means to test basic lexical knowledge, as it is quite a subjective task. The authors are aware that the *oral vocabulary* subtest is quite demanding, as the child needs to "tell specifically what a word means, an ability that requires a precise, definitive knowledge of a stimulus word" (page 84 of the test's manual), but they do not question the validity of including it in the test. According to the manual, in order to get the test item 'bird' right, for example, the child needs to provide two of the following: *is an animal, has a beak, has wings, is warm-blooded, has claws, lays eggs, pecks on trees, has feathers, flies, lives in trees, builds nests, something it eats (e.g., seeds, worms, fruit), lives in a cage, sounds it makes (e.g., sings, tweets, peeps, clucks, chirps), a particular species (e.g., peacocks, robin)*. In order to get the item 'kayak' correct, the child has to say one of the following: *a canoe that's enclosed with the exception of a hole in the center to sit in; or two of the following: you paddle it, used by Eskimos, made of canvas or skins, used with a two-headed paddle*. This subtest seems, therefore, to be highly dependent on encyclopaedic knowledge, and not on basic language skills.

2.5.1.2 The TOLD's grammatical subtests

Like the subtests of vocabulary, what Tomblin and colleagues named the 'grammar composite', with 3 subtests from the TOLD, has many weaknesses. The subtest *grammatic understanding* is the one which presents most of the problems, so I will focus my critique on this. *Grammatic understanding* is a picture selection task in which children need to point to the picture that best matches the sentence they heard. The criticism I have here is not about the nature of the task itself, but about the stimulus sentences and the selection of pictures. The authors' aims with this task seem somewhat unclear and this might have affected the way the stimuli and pictures were chosen. At the same time that Newcomer and Hammil state that they placed primary emphasis on the syntactic aspects of the sentence, they say the aim of their task is to "assess the

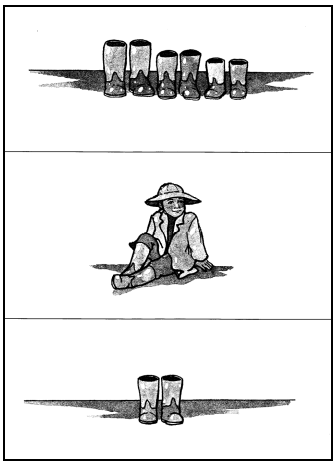
child's ability to comprehend the *meaning* of sentences" (Newcomer & Hammil, 1991: 8, my italics). The lack of precision in the wording when defining the aim of the task is reflected in the stimuli list. In many of the sentence items, the focus is not on the ability to establish syntactic relations or grammatically relevant distinctions: the understanding of the meaning of one or two words in the sentence is enough for choosing the correct picture. This is the case for the test item 'The picture that was drawn by the artist is finished'. As figure 3 below shows, the pictures used for this test item are: 1. upper view of man seated and holding a pencil near a white blank sheet of paper; 2. a finished drawing; and 3. a side view of two men standing by a drawing desk and holding a pencil. The understanding of the words 'picture' and 'finished' is enough to choose the target picture. There is no need to comprehend either the relative clause or the passive sentence embedded.

Figure 3: pictures used on the TOLD's *grammatic understanding* subtest; test item is 'The picture that was drawn by the artist is finished'



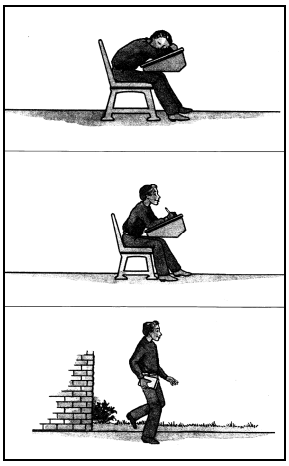
Many other utterances and pictures are questionable or ambiguous. As figure 4 below shows, for the utterance 'The children's boots are here', the pictures available are: 1. three pairs of boots lying on the floor; 2. a child wearing a pair of boots; and 3. one pair of boots lying on the floor. In principle, any of them could be considered correct, even if option 1 is the most suitable. In order to avoid problems, the authors could have chosen a noun which does not refer to an item that comes as a pair.

Figure 4: pictures used on the TOLD's *grammatic understanding* subtest; test item is 'The children's boots are here'



Moreover, a large number of test items share the problem of demanding a high level of inference for successful completion, partly because they are extremely difficult to represent pictorially. The test items 'Because he had already finished his work, he was not kept after school' (reproduced below), 'The boy has been waiting for a long time for his friend to arrive' and 'They could not continue the game until she joined them' illustrate well the amount of inference and picture complexity required in a large portion of the subtest. Indeed, the authors' reasoning behind the selection of sentence structures reveals a very unclear definition of 'complexity': "In constructing items for this subtest, we deliberately selected grammatic forms that were complex and that would challenge older children (…) In short, every type of grammatic relationship that came to mind and that appeared appropriate for pictorial rendering was incorporated into an item" (Newcomer & Hammil, 1991: 67).

Figure 5: pictures used on the TOLD's *grammatic understanding* subtest; test item is 'Because he had already finished his work, he was not kept after school'



The weaknesses of the TOLD are not exhausted by those discussed here. Taken together, the test's weaknesses show a trend that deviates from the original purpose of the test, that of being a test of 'basic language skills'.

2.5.2 Culatta et al's Story retelling task

In addition to the TOLD, Tomblin and colleagues made use of Culatta et al's story retelling task (1983) in an attempt to measure children's narrative skills. As mentioned on page 17, Culatta and colleagues use the terms *language* and *communication* somewhat interchangeably, which reveals that the authors share a view of language which is quite broad in scope. Culatta and colleagues' story retelling task uses a short story about the birthday party of a five year old boy. It contains 11 sentences, 136 different words, and 24 events. A group of 199 children participated in the study, divided into three subgroups: kindergarten (N=66), readiness[5] (N=16), and first-grade (N=117). Children were recruited in elementary schools near Lexington, Kentucky, USA. They were asked to retell the story to the experimenter and were then asked 10 questions about the content of the story. Comparison screening tools were used by Culatta et al with the intention of validating the story retelling task. The tests chosen were the Carrow Screening Test of Auditory Comprehension (STACL) (Carrow, 1973) and the Vane Evaluation of Language Scale (ELS) (Vane, 1975). The STACL is a brief picture selection test with 25 stimulus items varying from isolated lexical items such as the adjectives 'big' and 'red' to utterances with different linguistic

[5] According to the authors, "readiness classrooms are comprised of children of first-grade age but are not considered ready for normal first-grade placement" (page 67)

structures. There are utterances with passive voice sentences, with verbs in the past tense, and negatives. Some of the items are somewhat problematic, such as the one that tests the pronoun 'that', which is difficult to represent pictorially. The pictures for this test item are: 1) a girl reading a book on a chair; 2) a girl pointing to a book on a table; and 3) a girl walking with a book under her arms. Nevertheless, many of the items are generally acceptable, especially in comparison to the scale of problematic items present in the TOLD[6]. Results of Culatta and colleagues' study revealed that more children were identified as needing language services on the basis of story retelling than on the basis of STACL performance. In addition, more readiness and kindergarten children were identified as needing remedial services than first-grade children. Generally speaking, children who did poorly on the STACL also performed poorly on story telling, while many children who performed poorly on story telling performed well on the STACL. The authors conclude that "story retelling is a more stringent measure of communicative performance" (Culatta et al, 1983: 71). However, I will argue that, instead of being a more stringent means of evaluating communicative performance, story retelling (especially if used alone) is way too broad in scope, and difficulties in performing the task can potentially be caused by a number of different factors, not necessarily related to basic language skills. Indeed, the very fact that children who did poorly on the STACL also performed poorly on story telling, while many children who performed poorly on story telling performed well on the STACL suggests that the cognitive resources needed for successfully completing the story telling task are not all the same as those needed for completing the STACL.

Additional evidence of the view that Culatta and colleagues' perspective on language is a broad one comes from the following quotes: "In academic settings, children rarely encounter the need to comprehend or retrieve isolated language rules. Instead, for successful classroom functioning, they are required to use a variety of language rules in order to follow sequentially presented directions and explanations" (p. 66) and "The higher percentage of kindergarten children identified as needing services in the present study suggests the need for using graded story passages for both story retelling and language comprehension tasks. Use of these two measures may more closely evaluate a child's ability to meet the demands for integrative language performance

[6] Unfortunately, I could not gain access to the Vane Evaluation of Language Scale, so I will limit my analysis to the STACL

encountered in academic settings" (p. 72). These passages also serve as a basis for raising questions about the terminology used to refer to the target population of the tasks commonly used in the recruitment of children with SLI. It is clear that Culatta and colleagues have an educational setting in mind. Hence, it is probably safe to say that they consider that 'students' in academic settings are the target population of their narrative task. I will come back to issues about the terminology used to supposedly refer to the SLI population when discussing the CELF test below and we will see that Culatta and colleagues are not the only ones using terminology referring to school contexts.

2.5.3 The Clinical Evaluation of Language Fundamentals (CELF)

The TOLD and Culatta's narrative task are not the only instruments that present potential problems. Many other language assessment tests are subject to strong criticisms (Corrêa, Freitas & Lima, 2003). One of the most widely-used tests, the Clinical Evaluation of Language Fundamentals (CELF), both in its original and revised formats (Semel, Wiig & Secord, 1980, 1987, 1995), attracts similar criticisms to the TOLD. Problems with the CELF start with the very definition of the test's aims. At the beginning of the revised version's technical manual, the authors state that "CELF has proven to be a useful test for identifying students who lack the *basic language* skills which are the *foundations* of mature language use in communication: word meanings (semantics), sentence structure (syntax), and recall and retrieval (memory)" (my italics). As we will shortly see, like the TOLD, the CELF is *not* a test of basic language skills as it claims to be. The construction of the CELF was heavily based on *Language Assessment & Intervention for the Learning Disabled* (Wiig & Semel, 1980), written by two of the three authors of the test. An analysis of the book reveals major conceptual assumptions which are potentially problematic. As the book's title indicates, there is no particular reference to SLI. This, in principle, does not seem to be a problem. Nevertheless, the terminology used by the authors to refer to their target population, along with passages that reveal what the authors effectively used as their 'research object', casts doubts on the adequacy of the CELF for identifying potential cases of SLI.

With respect to the target population of Wiig and Semel's book, it is interesting to note that the authors make use of different expressions throughout the text. In most cases, 'learning disabled children' is used, but other terms are used with high frequency, such as 'language and learning disabled student',

'learning disabled adolescents', 'language disabled student' and 'language and learning disabled high school student'. Moreover, the authors often contrast the performance of 'disabled children and adolescents' with the performance of 'their academically achieving age peers'. It is not clear, therefore, which is the target population of the study presented by Wiig and Semel. In addition, the authors make frequent use of terms such as 'classroom', 'teachers' and 'school curriculum', Can the terms 'children' and 'students' be used interchangeably in a study that serves as basis for a test which is supposedly assessing *basic language skills*? Should the 'language' formally taught in schools be the focus of such studies? I argue that those features greatly diminish the validity of the CELF as a measure of basic language skills, i.e. those mastered by any typically developing child by around the age of five, without any formal instruction[7]. The authors deliberately state that their "book seeks to put the day-to-day management of the learning disabled child with a language disorder squarely within the domain of the classroom teacher" (page vii) and that their focus was on the "language components of the curriculum …" (page vii). Such statements seem inconsistent with the attempt to use Chomsky's generative theory, which, in 1960s' and 1970s' terminology, was concerned with the 'intrinsic competence of the idealized native speaker'. In addition, as raised previously in this chapter, Wiig and Semel make inappropriate use of some of Chomsky's terms, like many studies on language acquisition carried out within the developmental psychology framework of the time. Wiig and Semel start the section they called 'Forming sentences' stating that "in high school, grammar and English are the most difficult subjects for students with learning disabilities" (p. 60). They go on to say that there are many reasons for such difficulties, but "problems of memory and abstraction" would be of primary significance. Immediately after, in what they consider to be an attempt to better understand the difficulties encountered by high school students, they cite Chomsky's transformational grammar (Chomsky, 1957), making reference to *phrase structure rules* and the concepts of *deep* and *surface structure*s. The misinterpretation of the latter terms is particularly striking: "The *surface structure* reflects the syntactic properties of the sentence; the *deep structure* reflects the meaning. Thus, you could understand the *surface structure* of a sentence but not its *deep structure*, if you were unfamiliar with the words used; or you could understand the *deep structure* — know what ideas are being discussed — but not understand the *surface structure* — how the different words

[7] The authors of the TOLD also use 'students' as well as 'children' in their manual, but not so often as the authors of the CELF.

relate to each other" (Wiig & Semel, 1980: 62). To illustrate what they believe surface and deep structures mean, the authors provide the following examples: 'Go to the library and return these books for me', as supposedly the deep structure of 'Take these books back' (p. 62). The following quotes are also illustrative: "When we listen to a story, a lecture, or a discourse, we retain only the meaning or *deep structure* of sentences, paragraphs, and larger units" (p. 299); "when the youngster enters junior and senior high school … he must be accurate, efficient, mature and rapid in processing the *surface structures* of a variety of sentences and retrieving their underlying meaning" (p. 398); "to reduce the syntactic and *transformational* complexity of the language used in instruction, the teacher can use the guidelines we presented above for adapting reading materials" (p. 423) (my italics).

It is, thus, possible to say that Wiig and Semel's (mis)use of Chomsky's generative linguistics is purely as a descriptive tool, and not as a hypothesis about the nature of language acquisition. Indeed, no serious discussion about the nature of language acquisition or the relation between language and other cognitive abilities is undertaken by Wiig and Semel.

An additional problematic element is found in the CELF-R's technical manual. It refers to the description of the reasons why the authors excluded from the revised version the subtests of pragmatics, which were present in the original version of the test. Among other reasons, Semel, Wiig and Secord say that the original version of the CELF was judged by most users to take too long to administer, which influenced their decision to drop the pragmatics subtests in the revision. Just like Tomblin's statement cited above (see page 24), the decision of whether or not to include pragmatics in the test was taken (at least partially) not on the basis of theoretical and conceptual grounds, but on logistical grounds (duration of the test).

2.5.3.1 The CELF's subtests

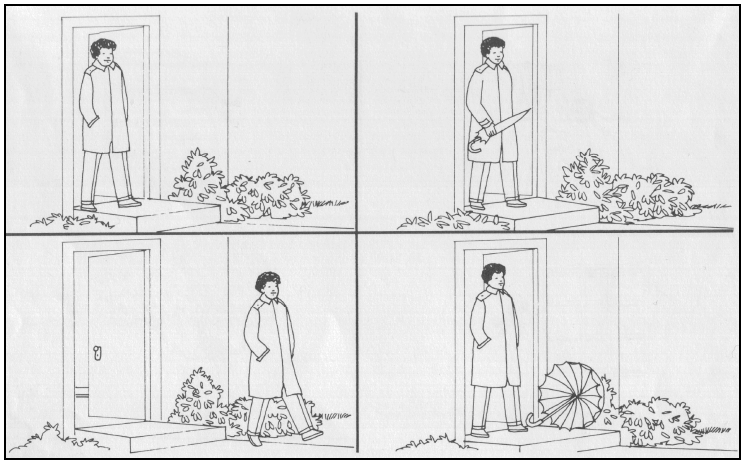
Like the TOLD, many of the CELF's tasks do not tap basic language abilities, but skills which are not specific to language. The subtest *oral directions*, for example, in which the child is supposed to "interpret, recall, and execute oral commands of increasing length and complexity" (CELF-R Technical Manual, p. 8), is poorly defined. The concept of complexity is not explicitly presented, but does not appear to be pertinent to language development. Children are asked to follow

instructions such as ‘Point to the first black triangle and the last small white circle’ or ‘Point to the last small black circle to the left of the big black square’. The reasoning behind the inclusion of such a task in a test which is supposedly assessing language abilities is questionable, as it demands skills which pertain to visual-spatial cognition as well as certain linguistic skills. A similar criticism applies to the subtest *linguistic concepts*, in which children are asked to point to different colour bars after hearing oral instructions. Questionable test items such as ‘After I point to a red line, you point to a blue line after you point to a yellow line’ and ‘If the red line is first, point to the yellow line’ are used with the children. The subtest *sentence structure* is similar to the TOLD’s *grammatic understanding* subtest in its design and format (picture selection task) and it is subject to some of the same criticisms I have proposed for the TOLD’s subtest, with respect to stimulus sentences and the selection of pictures. Many of the test items are difficult to represent pictorially, such as ‘The boy wanted to swim across the pool to sit with his friends’, ‘The woman asked: how much does this apple cost?’, ‘Father asked: shouldn’t you take out the rubbish?’ and ‘Mother asked: shouldn’t you play the piano now?’. Moreover, the two latter examples, which the authors classify as *indirect requests*, are of questionable validity. The interpretation of this sort of structure depends largely upon the context in which it is spoken and, therefore, involve extra-linguistic factors that should not be assessed (or at least should be avoided whenever possible) as part of a test which is supposedly evaluating basic language. Curiously, the book that provided the rationale for the construction of the CELF, *Language Assessment & Intervention for Learning Disabled* (Wiig & Semel, 1980), recognises the pragmatic complexity of *indirect requests* and the contextual and relational rules needed for their correct interpretation: “The pragmatic meanings elude them (language and learning disabled youngsters)” (p. 78). Oddly, thus, the authors provide — most likely unwittingly — arguments against the inclusion of *indirect requests* in a subtest assessing syntactic abilities.

Another problem of the subtest *sentence structure* lies in the design of the items with relative clauses. The four picture options for the test item ‘The man who is carrying his umbrella is walking out of the door’ do not provide the means for adequately testing the child’s knowledge. As figure 6 shows below, the pictures available for the child are: 1) man holding nothing and walking out of the door; 2) man holding umbrella and walking out of the door; 3) man holding nothing and a couple of meters away from the door; and 4) man holding nothing

and walking out of the door (umbrella is open and laying by the door); The correct answer, undoubtedly, is picture n^{o} 2. However, the way the remaining options were designed does not allow the possibility of picking up an error that can potentially be made by the child. If the child misses out the relative pronoun 'who' and instead of registering 'The man who is carrying his umbrella …', registers 'The man is carrying his umbrella …', picture n^{o} 2 is still the most adequate answer. This test item, therefore, is not adequately manipulating the necessary elements to test the abilities involved in the comprehension of relative clauses.

Figure 6: pictures used on the CELF's *sentence structure* subtest; test item is 'The man who is carrying his umbrella is walking out of the door'



The lack of precision in the definition of what is really being tested in the CELF is reflected in other subtests as well. Further in the technical manual, the authors provide additional examples in which they in some way contradict themselves, rather as they did with the *indirect requests* subtest discussed above. Two interesting examples come from the subtest *listening to paragraphs* (in which children have to answer questions based on stories read to them) and the subtest *word associations* (in which children have to provide orally the maximum number of lexical items belonging to a specific category – such as animals or means of transport – within a time limit). At the same time that the authors claim that the CELF is a "useful test for identifying students who lack the *basic language* skills" (p. 1, my italics), they say that *listening to paragraphs* requires children "to recall details they have just heard (…) and to draw inferences based on those details" (p. 24). They go on to say that "Because several sentence boundaries are crossed before the first question must be answered, the subtest probes longer-term aspects of memory". It can be argued that inference-drawing and remembering details such as the colours used by a character to paint a table in a

story (as is demanded in one of the test items) are not basic language skills and do not share the same cognitive mechanisms as basic language skills do. A similar argument can be put across for the *word associations* subtest. Having to recall, under time pressure, the maximum number of nouns referring to animals or means of transport is not a task which assesses basic language skills. As recalling large quantities of semantically-related words under pressure differs greatly from the way words are retrieved in the natural process of sentence formation, the task involves its own cognitive strategies. Once again, the lack of theoretical rigour in implementing what the test set out to do is observed and evidenced by an additional passage of the technical manual: "(…) although it seems a relatively "pure" measure of content, the word associations subtest also provides some insight into strategies employed in the recall of words stored in long-term memory, including planning and grouping strategies" (CELF Technical Manual, p. 24). Therefore, by the authors' own description, the inclusion of such tasks on a test which allegedly assesses basic language skills brings into question its validity and effectiveness for identifying deficits in these basic skills.

2.5.4 The Peabody

The Peabody (Dunn, 1965) is a test of vocabulary assessment with a picture selection format. It is, together with its British equivalent (the British Picture Vocabulary Scale, BPVS), widely used in the English speaking world, and has been translated into many languages. In general, the criticisms that have been made about the TOLD's subtest 'picture vocabulary' are applicable to the Peabody. Corrêa, Freitas and Lima (2003) observed that the Peabody does not offer a systematic list of stimuli, in particular, grammatical category was not taken into account when selecting the items for the task. There are uneven numbers of nouns and verbs, so the test does not allow the therapist or the researcher to evaluate the types of lexical items that the child has acquired. In addition, like the TOLD's 'picture vocabulary', many of the test items in the Peabody are highly dependent on schooling and access to formal, written language or specific topics, for example 'lethargic', 'ornament', 'lobe', 'sepal', 'mendicant', 'edifice', 'quiescence', 'walrus', 'jurisprudence', 'indigent' and many others. Moreover, many of the test items are very difficult to represent pictorially, for example 'convergence', 'astonishment', and 'constrain'. In these cases, the mapping of the lexical meaning with the picture can potentially be the main hurdle to overcome in completing the task. Corrêa, Freitas & Lima (2003) argue — and I endorse their

view — that the Peabody does not provide an informative measure of language development.

2.5.5 The Test for Reception of Grammar (TROG)

The Test for Reception of Grammar (TROG) (Bishop, 2003) is much less subject to the criticisms made for the CELF and the TOLD. The TROG is a standardised test which uses a picture selection task to assess a number of English structures. Unlike the CELF and the TOLD, the structures tested by the TROG are, in general, linguistically relevant. In addition, the TROG's test items and distractors have mainly been carefully selected, avoiding many of the problems presented by the other tests. The quotes below, from the TROG-2's manual, illustrate the author's efforts in desigining the test: "Test pictures are clearly drawn and brightly coloured. A deliberate effort was made to exclude pictures that are hard to discriminate on a visual basis." (p. 30), "Care has been taken to use a restricted simple vocabulary in test sentences, to minimise the likelihood of failure due to the client not knowing the meaning of individual words." (p. 30), "Every attempt has been made to minimise the influence of non-linguistic factors, such as plausibility of pictured events, on performance." (p. 30). Although the TROG is much more linguistically informative than tests such as the CELF and the TOLD, there are some aspects of the test that are problematic. I will address these below, starting with some general points and then focusing on the manner in which it assesses relative clauses (RCs), as the test fails to address issues that have been extensively discussed in the literature.

In general, the appropriateness of the number of pictures the TROG offers to the child in each test item is debatable. The child's task in the TROG is to select one picture in an array of four pictures. An alternative to that is to have the child choose from an array of three pictures. While the TROG's set-up diminishes the likelihood of a child selecting the correct picture by chance, having to scan four pictures instead of three increases the cognitive demand of the task. In the interest of minimising the chances of children making a mistake due to unnecessary processing load, it could be argued that the TROG offers a larger number of potential answers than is desirable.

It could also be said that some of the structures assessed in the TROG rely on logical reasoning more than a test of basic language skills should aim to do. In other words, in order to determine whether a particular sentence matches a

particular picture in some blocks, the child needs to make use of much more complex logical thinking then he or she needs for other blocks. The structures that seem to evaluate logical relations more than basic language skills are the following:

- Block H - Not only X but also Y: 'The pencil is not only long but also red'
- Block O - Neither nor: 'The girl is neither pointing nor running'
- Block P - X but not Y: The cup but not the fork is red'

It could be argued that understanding the logical constructions such as the ones above is very different from understanding basic sentence structure. It seems reasonable, thus, to say that part of the TROG is dealing with higher levels of usage of language which are more likely to be affected by schooling. In fact, in their normative sample, Block O gave rise to systematic errors, with 25% to 40% children below the age of eight years getting all the items wrong.

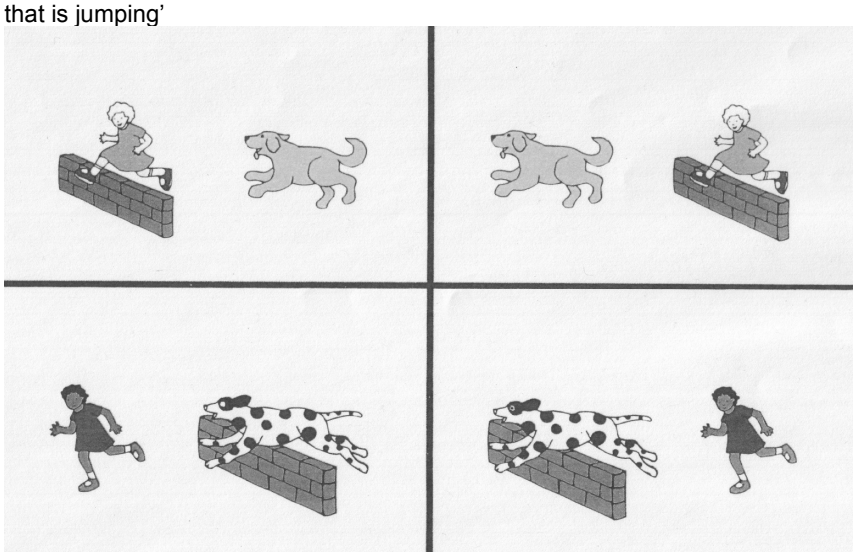
Let us now look at the way the TROG assesses relative clauses (RCs). In order to do so, we need to briefly review studies which have addressed methodological issues pertinent to testing RCs. As Adani (in press) points out, data from spontaneous speech and elicited production experiments (Diessel & Tomasello, 2000; Crain et al., 1990, Guasti & Cardinaletti, 2003, among others, apud Adani, in press) show that (typically developing) children can use RCs from at least age three and four. On the other hand, studies exploring the comprehension of RCs have reported that children continue to perform at chance level until five years of age (Tavakolian, 1981, Slobin, 1971, de Villiers et al, 1979, apud Adani, in press)[8]. This delay in mastering the comprehension of RCs has been claimed by different authors to be a task artefact. In other words, it has been argued that the relatively low performance of five year olds in comprehension studies is due to methodological flaws in the tasks that have been employed.

The studies that originally investigated the comprehension of RCs made use of acting out tasks, in which children have to manipulate toys in response to a stimulus sentence. Lately, picture selection tasks have been widely used instead of acting out tasks but, as Adani (op. cit.) points out, "in a number of

[8] It is interesting to note that these results are paradoxical at a first glance, as comprehension should logically precede production, as is found in many studies.

recent studies (Arnon, 2005; Arosio, Adani & Guasti, 2005; Friedmann & Novrogodsky, 2004) some of the 'old' methodological problems are still at play" (Adani, op. cit.: 4). The block testing RCs in the TROG has similarities to the comprehension tasks which have been criticised in the literature and, thus, it could be argued that the RC block in Bishop's test might trigger more incorrect responses than it should due to methodological issues. Let us now see why this seems to be the case. Figure 7 below is an example from the TROG's block testing RCs. The test item in question is 'the girl chases the dog that is jumping'.

Figure 7: pictures used on the TROG's relative clause block; test item is 'The girl chases the dog that is jumping'



At least two aspects of the TROG's RC block need to be discussed. The first one refers to the lack of the felicity conditions identified by Hamburger and Crain (1982) and the second aspect refers to Grice's (1975) Conversational Maxims.

Let us first look at the issues concerning felicity conditions as discussed in Hamburger and Crain (op. cit.). Note that the pictures above each present two characters: a girl and a dog. The two different dogs never appear together in the same image. According to Hamburger and Crain, this violates felicity conditions which are necessary for the interpretation of RCs. Felicity conditions state "what should be true of the context" (Hamburger and Crain: 258). Using the example in figure 7 above, the felicity conditions of the RC specify that there should be more than one dog available within each picture, since the function of the RC is that of restricting the set of potential referents for the definite expression in question (in Figure 7, 'the dog'). In other words, according to Hamburger and Crain, felicity conditions "stipulate that enough objects must be present in the experimental

setting" (p. 259). In this quote, Hamburger and Crain are referring to objects presented to the child in acting out tasks. The validity of their statement, nevertheless, is extendable to picture selection tasks as the one used in the TROG. In the example above, in order to meet felicity conditions, each picture should show two dogs, instead of just one[9]. Hamburger and Crain showed that, once felicity conditions are met (in the example in figure 7, this could be done by adding an extra dog in each picture), children of three and four years of age performed above chance, contrary to what previous studies had found.

Another semantic/pragmatic issue raised in Hamburger and Crain (op. cit.) refers to Grice's (1975) Conversational Maxims, specifically the Maxim of Manner. According to Hamburger and Crain, the use of the present tense with non progressive aspect, as in figure 7 above, seems to be in violation of Grice's Maxim 'avoid obscurity (of expression)': "this verb form (…) is normally used for definitions and recurrent events, but is unnatural, hence somewhat obscure, in the situation of the experiments" (p. 256).

The violation of Grice's Maxim of Manner is not restricted to the block testing RCs. The TROG's block assessing the comprehension of reversible passives (Block K) also violates Grice's Maxim of Manner, as the example below illustrates:

- Block K – reversible passive: 'The cow is chased by the girl'

Block K makes use of the present tense with non-progressive aspect, which, according to Hamburger and Crain (op. cit.) is unnatural in the experimental contexts.

In sum, therefore, it is possible to say that the TROG avoids many of the problems which are present in tests such as the TOLD and the CELF, making it a test more suitable for SLI investigations. It is undeniable, nevertheless, that the TROG still has some important drawbacks as an assessment of basic language skills, as the above analysis has shown.

[9] It could be argued that the extra dog in the other two pictures satisfies the felicity conditions for the RC to be interpreted. However, as Adani (op. cit.) points out, if more than one picture has to be taken into account at the same time, a different problem arises: in the example in figure 7 above, this would be the infelicitous use of the determiner 'the' to specify 'girl', given that two girls would be present in the context.

2.5.6 Remarks on standardisation

One might argue that the criticisms that were put forward in this chapter, in particular those referring to vocabulary testing, lose their strength when standardisation is taken into consideration. To some extent, standardisation does cover the fact that vocabulary knowledge varies among the population. It could, thus, be argued that items such as 'salmon', 'infirm' or 'abode' are acceptable on a test of vocabulary knowledge because most children would not know them anyway, even typically developing children. However, I would argue that if a test is used to assess basic language skills, it should be free of items that can pose a problem for less advantaged children.

2.6 Summary

So far in this chapter we have seen that, despite the considerable progress of research in SLI over the last couple of decades, the field is still quite unstable and not as settled as some researchers assume. In addition, we saw that the term *language* receives multiple interpretations, leading to different conceptions of *language* being used concomitantly by different researchers who assume that they are addressing the same phenomenon. One group of researchers works with a broad meaning of *language*, often used interchangeably with the term *communication* or, to a certain extent, with language skills relating to educational attainment. Other researchers share a much narrower view, namely that of language as a computational system plus a lexicon which specifies the semantic, the phonological and the formal or lexico-syntactic features of words. The fluctuation in use of the term *language* has had a negative impact on how SLI research has been conducted and implemented, as the different conceptions of *language* have given rise to studies that are only *apparently* dealing with the same phenomena. We have also seen that many of the tests widely used in the field do not effectively assess what their authors claim to be assessing.

In the next section, I develop the framework of two potential scenarios for the field of SLI, in an attempt to clarify some of the issues outlined in the earlier sections.

2.7 SLI: a clinical or a socio-educational matter?

2.7.1 Scenario 1 SLI is NOT a basic language disorder per se, but a non-standard academic profile

The above historical overview of SLI and analysis of the diagnostic procedures commonly used in the field open up the possibility to offer a socio-cultural resolution on the debate on SLI. We have seen that many of the tests available for use with children are not targeting basic language abilities, i.e., those abilities already mastered by any typically developing five year old child, regardless of their level of formal schooling. Instead, tests are assessing skills which are required for successful school performance, which is, in turn, oriented by socio-cultural customs. With this in mind, I will propose a scenario of SLI which can potentially resolve controversies in the field. According to this scenario, what researchers commonly refer to as SLI is not really a disorder specific to language (in its narrow conception), but a failure in achieving the necessary communication skills for educational success. In addition to the analysis offered in the previous section, other evidence support this scenario.

2.7.1.1 Evidence of test bias

One of the most striking indications of the potential plausibility of this scenario is the general acknowledgment that many tests used in SLI studies give rise to poorer performance within most minority populations in the United States. Tomblin et al (1997) report that "Native American and African American children presented the highest rate of SLI, followed by Hispanic children, and then White children" (p. 1255). They also report that the prevalence of SLI in African American children was 11% as opposed to 7% in white children (Tomblin, 1997, apud Hammer, Pennock-Roman, Rzasa & Tomblin, 2002). Hammer et al conducted an analysis of the TOLD-P:2's test items for evidence of differential item functioning (DIF)[10]. Results revealed that 16% percent of all items of the TOLD-P:2 had DIF. In the third version of the TOLD (TOLD-P:3), the authors restructured two of the subtests (*sentence imitation* and *word articulation*), as some of their items were not compatible with nonstandard or regional English. However, as 'non-mainstream' populations scored more poorly across subtests, the restructuring of two subtests of the TOLD-P:3 does not seem enough. Hammer et al (op. cit.) conclude that their findings "should cause speech-language pathologists to use caution when using the TOLD-P:2 with African American children" (p. 282). Performance on the CELF also yielded differences

[10] As Hammer et al point out, the DIF method of evaluation "involves the identification or flagging of items on which two groups of children score differently after taking into account an estimate of their overall skill in the underlying construct (e.g., true vocabulary knowledge, true grammatical understanding). The estimate of the underlying construct is usually derived from the test items themselves and is therefore an internal criterion" (Hammer et al, 2002: 275).

between populations. African Americans scored approximately one-third to one-half of a standard deviation lower than white people (CELF Technical Manual, p. 42).

It is outside the scope of this thesis to discuss, in detail, socioeconomic issues pertaining to different social groups within the United States. However, it is essential to consider a socioeconomic angle in a potential explanation for the differences stated above. As it is widely known, African Americans, Native Americans and Hispanic Americans have historically been disadvantaged in North America. Needless to say, the influence of socioeconomic status (SES) on child development is a highly complex issue. Nevertheless, there seems to be compelling evidence that parental SES accounts for a substantial portion of ethnic gaps in school achievement and attainment (Kao & Thompson, 2003; Sirin, 2005). Indeed, the mathematics test scores of African Americans lag behind those of white students (in addition to scores of vocabulary and reading tests) (Kao & Thompson, op. cit.). A very similar pattern is found when analysing the results of the SAT Reasoning Test, a standardised test for university admission in the United States: white students perform on average better than African Americans, Hispanic Americans and Native Americans (Kao & Thompson, op. cit.). The same seems to be true for variation in grades, which are highly correlated with parental SES, school dropout, high school completion, college transition and completion. In the case of the SAT, Dorans and Kulick (1986, apud Hammer et al, 2002) have demonstrated that between 12.9% and 16.5% of all verbal items on the SAT display DIF.

Kao and Thompson (op. cit.) report an interesting phenomenon linked to the Mexican community in the United States, who make up more than 75% of all Hispanic population in the country. The average educational attainment of many Mexicans who migrate is very low, but, once factors such as generation, language and social capital are controlled for, the school dropout rates improve significantly (White & Kaufman, 1997, apud Kao & Thompson, 2003), suggesting that differences in school performance (including SAT and other forms of testing) between Mexican immigrants and mainstream North Americans are at least partly due to SES factors.

The case of Asian immigrants is also enlightening. Tomblin et al (1997) report that none of the 70 Asian children they examined were found to present

SLI. Interestingly, the performance of Asians on the SAT Reasoning Test and the data available on school dropouts, grades and school/university completion reveal that Asian immigrants are usually more successful than the group formed by white students (Kao & Thompson, op. cit.). Contrary to immigrant Mexican students, immigrant Asian students are generally extremely advantaged in terms of parental education levels, which can explain at least part of their relatively high performance and attainment. Data on Japanese descendants in Brazil might also shed some light into explaining the differences found between Asian immigrants and other groups in the United States. São Paulo state, in the southeast region of Brazil, is home to 1.3 million Japanese descendants, who comprise around 3.25% of the state's total inhabitants[11]. Interestingly, however, the proportion of Japanese descendant students at the University of São Paulo is around 10%[12]. Moreover, some comparison data between what is called HIP Asia (highly performing Asia — Japan, South Korea, Singapore, Hong Kong and Taiwan) and Latin American countries is also informative (Moura Castro & Verdisco, 2002). According to their Inter-American Development Bank publication, the public sectors of HIP Asia and Latin America countries spend similar amounts on education. One of the main differences between the two regions, however, is that HIP Asian countries allocate more of the total expenditure to primary education, while Latin American countries spend disproportionately on higher education. Further, the authors discuss regional differences in commitment of families to educate their children. In HIP Asia, family expenditure on education is astonishingly high and parents invest a considerable amount of their own time helping their children with their homework and studying for their tests. According to Moura Casto and Verdisco, similar behaviour has been observed in Asian immigrants living in the United States. The high parental education levels of Asian descendants in the United States, combined with the importance the culture gives to schooling can arguably explain why SLI was not found among the Asian children in Tomblin et al's sample.

2.7.1.2 School age disorders in the context of law and school policies

In addition to indications of test bias on the TOLD, it is important to note another socio-cultural-educational factor present in the context of North American schools (and, possibly, at least to a certain extent, in British schools). According to

[11] Figures takes from the website of the Government of São Paulo on 2 April 2008.

[12] Figures obtained on the 2007 annual review of University of São Paulo. Downloaded on 2 April 2008.

Newcomer and Hammil, one of the uses of the TOLD is to "provide a means for evaluating children's progress in prescribed remedial programs. The monitoring of children's progress is an important component of special instruction and *often required by law or by school policy*" (p. 13, my italics). With this in mind, it is plausible to think that at least part of the 'SLI labeling' of children that occurs in the United States is due to pressures imposed by the school system. Something similar is likely to be taking place with the diagnosis of ADHD (Attention Deficit/Hyperactivity Disorder). There have been increasing concerns about the overdiagnosis of ADHD in school age children, especially in the U.S., where most of the world's supply of psychostimulants is consumed (Marshall, 2000; Mackey & Kipras, 2001, apud LeFever et al, 2003). Such pressure would not apply to countries such as Brazil, due to the characteristics of its educational system. The educational system in Brazil offers roughly two types of schools: (1) *privately-owned schools*, which are either religious or commercially-driven; in general, these are attended by those at the high end of the social class divide (a relatively small percentage of the population); and (2) *public schools* (public in the sense of no tuition charged/free access), which are funded by the Government at 3 levels: municipal, state or federal; with very few exceptions, these schools are attended by children coming from low income families. Comparatively, the school systems of Brazil (and possibly other countries in Latin America or other regions in the world) are not as homogeneous as the schools systems in the United States or the United Kingdom. While the private school sector in Brazil has to follow the guidelines of the Ministry of Education, schools are run relatively independently. The public school sector, on the other hand, is controlled by the three government levels mentioned above and schools are, therefore, relatively more centralised. The public sector, nevertheless, lacks a substantial amount of funding and many schools do not have enough teachers for basic subjects. Special educational needs are, therefore, not a major concern for many Brazilian public sector schools, as it is for North American and British schools. It could be argued, thus, that pressure imposed by the school systems in some countries (e.g. access to government funding for special instruction) might actually inappropriately increase the number of children diagnosed with SLI.

2.7.1.3 Validity of the category 'Specific Language Impairment'

Under Scenario 1 outlined above, I argue that Specific Language Impairment would be an invalid and inappropriate category. If a case is made for this scenario, the use of the terms 'specific' and 'language' should be revised, as they

do not seem suitable to refer to a set of problems greatly related to difficulties in reaching the expected school performance.

2.7.2 Scenario 2 SLI in its narrow conception is a valid category

In this section, I outline an alternative scenario, under which SLI could be a valid category. Here, I argue for the adoption of a working hypothesis which assumes the existence of an impairment in basic language skills, i.e., those skills independent of formal instruction and naturally acquired under normal circumstances. The reasoning behind Scenario 2 is strongly based on the logical possibility of the existence of selective disorders in the human mind. Here I argue in favour of the view that, if selective disorders are a logical possibility, they must be explored, at least as a starting point of research. The approach advocated here makes use of a narrow conception of *language*, grounded in generative linguistics and, consequently, in a (functionally) modular approach to the human mind. Nevertheless, as will become clear, this approach differs fundamentally from the current accounts of SLI which are commonly grouped under 'linguistic accounts'. While I argue for the adoption of a narrow conception of the term *language*, I also argue for an integrative approach to the disorder, combining linguistics and psycholinguistics in a way the disciplines may complement each other.

What has been put forward under Scenario 1 would still be partially valid under Scenario 2, as the prevalence of a truly language deficit would be expected to be much lower than the 7% that is estimated by the current literature, and an explanation based on impairment in the communication skills that are necessary for school achievement would still be needed to account for many of the children who are currently diagnosed with SLI.

2.7.2.1 Need to differentiate between distinct underlying problems

Needless to say, basic linguistic abilities are essential in most learning tasks at school and, therefore, a basic language system which is not functioning normally will certainly impact on school performance. Nevertheless, many other abilities are required for formal learning, and a careful investigation of the cognitive mechanisms involved in learning tasks is imperative. Here, I endorse Barrett and Kurzban's (2006) view that there is little doubt that different kinds of information are handled by different systems in our minds and that no computational mechanism can simply process any kind of information in any way. Presumably,

then, there is no reason to think that the natural and, to a certain extent, involuntary acquisition of a native language is handled by the same mechanisms responsible for learning content through formal instruction (such as the formal, infrequent words used in the tests we analysed or, arguably, the inferential knowledge needed for completing parts of the standardised tests). Research on SLI must be able to distinguish problems which originate in the process of spontaneous acquisition of a language from problems which emerge in the school context. This assertion does not mean that both types of problems cannot co-exist. The downside of current research and clinical practice is that tests do not generally provide the means to differentiate between a child with a genuine language impairment from a child with intact basic language skills who demonstrates difficulties in the school environment. Two crucial problems, of different sorts, arise from this. First, from a theoretical viewpoint, current research is not as productive as it could be, as it does not inform us much about the functional architecture of the human mind, how different types of information are accessed and how they are processed. Second, from a clinical perspective, such a broad approach to SLI cannot adequately direct the therapy offered to children who are referred to clinical services.

2.7.2.2 Shortcomings of current 'Linguistic' approaches to SLI

The call for an approach to SLI under a narrow view of language naturally points to a set of current hypotheses about the disorder commonly grouped under 'linguistic hypotheses'. Much has been discussed about the main hypotheses that have been put forward to explain the nature of SLI, but a note on the methodology used by those who argue for a 'linguistic' explanation is required here. Despite important differences, the work of Clahsen, Wexler and van der Lely share some common assumptions in their attempt to formulate hypotheses for SLI by making explicit use of Chomsky's generative linguistics theory. Importantly, however, a look at their research, with special attention to their methodology, reveals some inconsistencies between the arguments they promote and the way they diagnose their subjects.

Rice and Wexler (1996), for example, make use of the Peabody and the TOLD to recruit children for their SLI study, while van der Lely and colleagues make occasional use of the CELF (Marshall & van der Lely, 2008, Ebbels & van der Lely, 2001); and frequent use of the BPVS (British equivalent to the Peabody) (van der Lely, Rosen & Adlard, 2004; Marshall & van der Lely, 2006; van der

Lely, 1997; van der Lely & Stollwerck, 1997). In the case of Clahsen's work, a hypothesis that became well-known in the field was put forward on the basis of two types of data: 1) English data originally collected by van der Lely and 2) German data of children independently diagnosed with SLI by speech and language therapists. Clahsen (1989) states that "Independently from our studies, the children have been assessed as dysphasics[13] by speech and language therapists …". Further on in the same paper, Clahsen writes the following passage, in which he criticises the classification of dysphasia used in the German clinical context: "… little is known about dysphasia in German-speaking children, either about the characteristic linguistic features or about the possible causes. In the clinical context, Liebmann's (1901!) classification of dysphasia is still in use … This system is just an unsystematic collection of some salient features of a child's linguistic behavior which is completely out of date, given the state of the art in first-language-acquisition research".

It could be argued that such approaches to diagnosis weaken the claims made by these authors: their proposed accounts of SLI are conceptually incompatible with the tests being used to recruit children for their studies. With this in mind, it is plausible to question their hypotheses on very basic grounds: are they putting forward hypotheses about the nature of the disorder or providing descriptions of the patterns that emerge in the data? I will argue for the latter and propose that they in some way invalidate their own argumentation. Wexler, Clahsen and van der Lely claim to be taking a narrow, linguistic perspective, but, in effect, they are using a relatively broad approach in the recruitment of subjects, one which, according to what has been proposed earlier in this chapter, is not suitable for testing basic language skills. Therefore, a very basic problem arises, which can be formulated as follows: highly specific and selective claims about a developmental language disorder are being made on the basis of population samples recruited on broad criteria.

Although Wexler's and his colleagues work, for example, focuses on relevant linguistic aspects and seeks to identify linguistic markers, by using tests like the TOLD and the Peabody as recruitment tools, they disregard what causes their subjects to perform poorly on the subtests which have questionable validity when it comes to identifying cases of SLI under a narrow interpretation. A

[13] See section 2.2 of this chapter for a review of different terms used to refer to children with SLI.

question that needs to be addressed in this situation is why the children they identify as SLI make a large number of errors on the Peabody and on the TOLD and the children they use as part of the control group do not. Is it really the case that the problems presented by the children who are identified as SLI are caused by a deficit specific to language in its narrow conception?

Here, the issue of *primary* and *secondary* language deficits becomes crucial. Many factors can cause difficulties in the performance of tasks that require verbal language. Some deficits, for example, even if manifested in language performance, may result from problems in cognitive domains not probed by the non-verbal tests used to identify children with SLI. Even if two populations perform similarly in a certain behavioral task, it does not necessarily mean that they have the same underlying problems. It has been reported, for example, that individuals with Down Syndrome present difficulties with passive sentences (Bridges & Smith, 1984) and with reflexive pronouns but not with non-reflexives (Perovic, 2003), a pattern which is somewhat similar to the one reported for children with SLI (Fukuda & Fukuda, 1994; Jakubowicz et al., 1998; Silveira, 2002). In principle, then, taking a broad approach when recruiting children with SLI and then conducting experiments manipulating relevant linguistic aspects does not guarantee a reliable group of subjects. Ideally, theoretically-motivated and well-grounded *diagnostic* procedures should be adopted from the very start of any investigation. The administration of language tests based on a narrow conception of the term, however, is not enough, since, as mentioned before, many factors can affect performance. More comprehensive non-linguistic tests are also needed, including, perhaps, tests assessing skills related to language in its broad conception, under the assumption that not all verbal behavior is controlled by the same cogntive processes. That way, SLI research would be able to differentiate between cases of *primary* language deficit and *secondary* language deficit.

In practice, however, researchers have to deal with several limitations and can only work with the tests which are available. Having said that, while it is likely that researchers working with a narrow view of SLI have to rely on broad language tests to recruit participants due to the lack of more appropriate alternatives at this stage, it seems that this is done without much discussion regarding the limitation of current diagnostic procedures. This lack of questioning regarding the appropriateness of current diagnostic tools in the field of SLI is, in

my view, a major issue that should be given much more thought than it is given at this time. Ultimately, differentiation of the different types of problems is important for identification of children's needs and, hence, appropriate intervention.

Before moving on to the next section, van der Lely's proposal needs to be looked at in more detail. Van der Lely and colleagues, besides making use of generative linguistics as an attempt to provide explanations for their data, make very specific claims about an alleged subgroup of SLI, which they have named Grammatical SLI (G-SLI). According to these researchers, individuals with G-SLI "suffer from a relatively pure developmental deficit in the grammatical aspects of language (syntax, morphology and phonology) that are core to the human language faculty" (van der Lely, 2004: 122). G-SLI supposedly affects around 10-20% of children within the population of SLI. Selection of G-SLI participants conforms to the following steps:

> "Selecting G-SLI participants is a two-stage process. In the first stage, children between the ages of 8 and 16 who have received a diagnosis of SLI are recruited from residential language schools or from language units within day schools. This recruitment is done with the help of speech and language therapists, who are asked to select only children with normal hearing and articulation, with English as a first language, and without a diagnosis of autistic spectrum disorder. Non-verbal intelligence tests are administered (e.g. British Ability Scales, BAS, Elliot, 1996; Raven's Progressive Matrices, RPM, Raven, 1998) to ensure that we only select children with non-verbal IQ scores of greater than one standard deviation below the mean (i.e. a standard score greater than 85). Scores from standardised language tests, including the Test for Reception of Grammar (TROG; Bishop, 1983), British Picture Vocabulary Scales (BPVS; Dunn, Dunn, Whetton & Burley, 1997) Test of Word-Finding (TWF; German, 2000) and Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig & Secord, 1995) are obtained, often from the child's speech and language therapist, in order to build up a profile of the child's general language abilities. Children who have been recruited in this way, and who show a pattern of a more severe impairment in grammar than in vocabulary, as based on comparison of standardised scores in language tests, then pass through to the second stage.
>
> In the second stage, children are administered a series of tests devised by van der Lely to assess the specific grammatical abilities that characterize G-SLI (van der Lely, 1996b, 1997c, 2000). Although standard tests assess a wide variety of skills within the area of syntax or vocabulary, van der Lely's tests target specific areas of grammar that children with G-

> SLI find particularly difficult – verb agreement and tense, reversible passives and pronominal reference." (Marshall, 2004: 41)

At least two aspects of the G-SLI selection procedure need to be discussed. Both are related to the administration of the BPVS to test the children's vocabulary. First, as we have seen earlier, a poor performance on the Peabody or the BPVS cannot be taken as a reliable measure of basic language skills. With this in mind, it can be argued that van der Lely and colleagues' selection of subjects starts in a misleading manner, even if in a later stage more relevant tests are administered. Second, even if they claim that the vocabulary of children with G-SLI, although impaired, is not as severely impaired as their 'core grammar', their own data show that this does not seem to be the case: in Marshall (2004), for example, the standard scores of 13 out of 24 children with G-SLI (and of the group as a whole) on the BPVS are lower than their standard scores on the TROG. Among the reasons van der Lely and colleagues use to explain this paradox is the idea that their problems with the lexicon are caused by syntactic difficulties, since "vocabulary acquisition also relies on syntactic bootstrapping (Bloom, 2000), and syntax is impaired in G-SLI" (Marshall, op. cit., p. 45).

In order to reverse the present picture, the language tests that are used with potential cases of SLI must improve. Only theoretically motivated standardised tests can lead to reliable subjects.

2.7.2.3 Relevance of research with younger children

An approach to SLI under Scenario 2 would also call for a change in the age range of children recruited in the studies. Curiously, despite the fact that researchers agree in defining SLI as a 'developmental language disorder' (as opposed to an 'acquired language disorder), most studies focus on children who are already at a relatively late stage in the process of language acquisition, when lexical, morphological and syntactic units have been segmented and have been produced for a long time, even in the case of children with delayed acquisition. For example, Clahsen et al (1997), Archibald and Gathercole (2006) and Brinton et al (2007) presented data on children with SLI with mean age of 6;4, 9;8 and 9;1, respectively. As is widely known, and supported by the following passage from Slobin (1985), typically developing children acquire the largest chunk of the grammar of the language surrounding them much earlier than that: "It is safe to say that except for constructions that are rare, predominantly used in written

language, or mentally taxing even to an adult (like 'The horse that the elephant tickled kissed the pig'), all parts of all languages are acquired before the child turns four (Slobin, 1985)". Needless to say, a child with SLI will behave in a different way and her/his language acquisition process will not achieve stability (if achieved at all) by 4 or 5 years old. Nevertheless, the relatively short time taken by a typically developing child to master his/her native language, along with the compelling evidence offered by studies with newborns, supporting the hypothesis that the process of language acquisition starts well before the child begins to speak (cf. Mehler and Dupoux, 1990), allow one to argue for the need for SLI studies targeting younger children.

2.8 Final observations about the inconsistencies in the field of SLI

I have argued throughout this chapter that the field of SLI is marked by a range of inconsistencies of various kinds. The lack of coherence among some researchers and studies gives rise to the use of inconsistent definitions and imprecise statements. We saw earlier, for instance, that Clahsen selects children who have been independently diagnosed as SLI by speech and language therapists although he criticises the way children are diagnosed in the German clinical context. I suggested earlier in this chapter that researchers are inaccurately discussing data and studies as if those referred to the same phenomenon. At this stage, it is possible to bring together further evidence to support this proposal. A very clear example of the lack of coherence in the field is the comparison set by Clahsen and Almazan (1998) where German data collected by the authors is directly contrasted with previous English data collected by van der Lely and colleagues. Recall that van der Lely and colleagues claim to have identified a subgroup of SLI, who they refer to as G(rammatical) SLI. Based on this paper, one could assume Clahsen and colleagues also focus on the alleged subgroup of children with SLI. However, Clahsen and Almazan use van der Lely's data without making any reference to the subdivision, and consequently, it is possible that they are not comparing like with like.

In his 1998 book, which has become widely cited in the field, Leonard states, as we saw on page 19 of this dissertation, that "the prevalence of SLI is about 7%". Later, when discussing the hypotheses about the nature of SLI, he makes use of the two common groups of hypotheses in the literature and puts together, under linguistic approaches to SLI, Clahsen's *Missing Agreement Hypothesis*, Wexler's *Extended Optional Infinitive Account*, Gopnik's *Feature*

Blindness Hypothesis and van der Lely's *Representational Deficit for Dependent Relationships Account*. Are all these accounts trying to handle the supposed 7% of children who have SLI? Are they alternative accounts to the hypotheses Leonard puts together under processing capacity accounts? Or are they referring to the alleged subgroup of children with SLI van der Lely claims have G-SLI?

The lack of coherence in the field can also be exemplified by the age range of subjects recruited to take part in studies. Conti-Ramsden and Durkin (2008), for example, focused on adolescents with SLI and stated, quoting Nippold (1998), that the focus on such a group is relevant as "language continues to develop through adolescence" (Conti-Ramsden and Durkin, 2008: 70). Can the term *language* in Conti-Ramsden and Durkin's statement be equated with the term *language* in, for example, the study by Rice, Wexler and Cleave (1995) or by Anderson and Souto (2005), who investigated the abilities of children with SLI of 5 years of age or even younger? It is undeniable that our communication abilities can expand through adulthood. However, the type of *language* mentioned by Conti-Ramsden and Durkin as continuing to develop through adolescence is unlikely to be the same type referred to in the other two studies.

It seems, therefore, that evidence put forward by authors who share one view of SLI is often reported and used by authors who seem to take a different view of the disorder, and differences are not always acknowledged. It may be concluded that the term 'SLI' is likely being used without enough discussion about what the acronym really stands for.

2.9 Discussion and conclusion

Part I of this thesis presented a thorough analysis of the field of SLI, in light of the controversies that surround the investigations of the language impairment and the disciplines that share an interest in the disorder. I sought to present an analysis of the field from a wide perspective, revisiting many aspects of SLI research, culminating in two possible scenarios. Next, in Part II, I present an experimental investigation carried out with six Brazilian children with language impairment. The experimental study was designed on the basis of Scenario 2 proposed in Part I. As previously discussed, according to Scenario 2, SLI is potentially a valid category if a narrow approach is undertaken. The experimental study reported in Part II of this thesis is an attempt to put such an approach into practice. It differs, nevertheless, from the other approaches which assume a

narrow conception of language. Contrary to Clahsen or Wexler, for example, who work with a narrow conception of language but use standardised tests based on a broad definition of the concept for recruiting participants for their studies, the current study attempted to take a narrow approach from its start. Participant recruitment made use of MABILIN, a test being developed by the Psycholinguistic and Language Acquisition Laboratory (LAPAL) at Pontifícia Universidade Católica at Rio de Janeiro (PUC-Rio). MABILIN uses a picture selection task and has been specifically constructed for the assessment of Brazilian Portuguese-speaking children. Differently from many other tests, MABILIN has been "constructed on the basis of an integrated theory of linguistic competence, that is, a theoretical approach for the cognitive competence of language which integrates a model of linguistic knowledge (…) with the psycholinguistic study of processing abilities" (Corrêa, 2005a). Although no standardised test is perfect and the MABILIN is still under development, I believe that its use for participant recruitment provided the opportunity for avoiding at least some of the pitfalls associated with other tests and discussed extensively in Part I. MABILIN avoids these pitfalls by attempting to make a clear distinction between linguistic skills and skills dependent on other cognitive domains, such as encyclopaedic knowledge and inferencing. In addition, the linguistic and psycholinguistic variables used in MABILIN are well controlled, providing a balanced way to test children's skills. Moreover, being specifically designed for Brazilian-Portuguese children, MABILIN provides a much more reliable measure for testing children in the Brazilian context than translated versions of foreign tests. In chapter 6, more details of MABILIN are presented, particularly with respect to how problems present in the TROG (cf. section 2.5.5) were avoided.

In addition to using an appropriately targeted test to recruit participants, the experimental investigation that follows focused on a very specific linguistic phenomenon, namely grammatical gender. As we will see in Part II, gender is adequately defined on grammatical grounds and its determining criterion is *agreement*. I aimed for tasks that probe linguistic knowledge instead of encyclopedic or inferential knowledge of the sort required by the TOLD or the CELF. I anticipate, however, some difficulty in testing input processing abilities related to gender agreement, as experienced in Silveira (2002) (cf. 7.2.1 for a discussion about the obstacles in testing the comprehension of gender processing).

PART II
PSYCHOLINGUISTIC INVESTIGATION

Chapter 3
INTRODUCTION AND THEORETICAL CONSIDERATIONS

3.1 Introduction

In Part II of this thesis, I report a study which investigated aspects of gender agreement within the Determiner Phrase (DP) in Brazilian Portuguese. A series of six experiments exploring gender retrieval, agreement between determiner and noun and agreement between noun and adjective was administered to six children with language impairment and 60 typically developing children.

As mentioned at the end of Part I, the experimental study was undertaken on the basis of a narrow conception of the term *language* (understood in terms of grammatical knowledge), according to which SLI could be a valid category for research if it focuses on basic language skills, i.e., those abilities acquired spontaneously by any typically developing child by roughly the age of five. In addition to adopting a narrow conception of *language*, the experimental study that follows was carried out under the assumption that our understanding of SLI can only increase if an integrated and conciliatory approach to the disorder is adopted. This approach is based on the idea, as anticipated in the main Introduction, that the study of SLI should be done under a framework in which linguistic and processing accounts are not mutually exclusive, but complementary. Research on SLI is strongly marked by contrasting approaches. On the one hand, linguistically-oriented hypotheses interpret the problems presented by children with SLI strictly on the basis of formal linguistic models, without any reference to the mental processes involved in language production and comprehension. The so called 'linguistic hypotheses' (Clahsen, 1989; Clahsen et al., 1997; Rice, Wexler & Cleave 1995, Rice & Wexler 1996, Rice, Wexler & Hershberger, 1998; Gopnik, 1990; van der Lely, 1998; van der Lely, 2003; van der Lely, 2005) generally assume that the problems encountered by children with SLI result from incomplete knowledge of some particular grammatical aspect, attributed to a deficit in the child's grammar. Thus, linguistic theory is often used as a tool for describing the patterns of errors characteristic of children with SLI, and descriptions are then given explanatory status.

On the other hand, the 'limited processing' accounts of SLI tend to be restricted to issues related to speech perception and information processing, without reference to linguistic models that provide a characterization of what needs to be acquired by the child or processed in production and comprehension. In general, hypotheses belonging to this group (Tallal & Piercy, 1973; Tallal et al, 1996; Gathercole & Baddeley, 1990) share the assumption that the underlying deficits of SLI are caused by a limited processing capacity, albeit differing in the scope of the processing deficit.

It is often the case that one type of approach is adopted in sharp contrast to the other type of approach and treated as an alternative to the opposing approach. In other words, SLI research is frequently carried out as if 'processing' and 'linguistic' accounts provided distinct answers to the same problem. Such polarisation, nevertheless, seems misleading. It is my view that an 'explanation' for SLI data on the basis of formal linguistic models and no reference to psycholinguistic models of production and comprehension does not grasp all relevant aspects of the phenomenon. The opposite is also arguably applicable. Explanations for SLI data that make no reference to linguistic models leave out a crucial part of the puzzle.

In light of the above, it seems that an approach to SLI which incorporates a procedural dimension to the disorder, as well as formal models of language, is desirable. In 3.2, below, I discuss a proposal that deals with the relationship between linguistics and psycholinguistcs in an interesting way.

3.2 Metatheoretical considerations

3.2.1 A potential dialogue between linguistics and psycholinguistics

The recent publication of papers about the relationship between linguistics and psycholinguistics (Neeleman & van de Koot, 2009; Phillips & Wagers, 2007; Corrêa and Augusto, 2006), which discuss a range of issues such as misconceptions about notions used by both fields and matters regarding the research object(s) of the two disciplines, suggests that it is not clear yet how research should be carried out in order to establish an articulated dialogue between linguistic theory and theories of language processing. Establishing a dialogue between the two disciplines is certainly not an easy task. Moreover, despite the recent publications, it is not a topic that is discussed as often and as thoroughly as it should be. Yet, attempting to articulate the two disciplines is

arguably crucial for a better understanding of language in its cognitive dimension. The current thesis has as a background the assumption just outlined, namely that an attempt to put forward an articulated approach to the study of language cognition and, consequently, to Specific Language Impairment, is necessary. Implementing such an approach is demanding and, to my knowledge, not many proposals for articulation have been advanced.

Among those who have addressed the issues above, Neeleman and van de Koot (2009) and Corrêa and Augusto (2006) refer to the work of Marr (1982) on vision, arguably extendable to any complex information-processing systems. For Marr (op. cit.), these systems must be approached at three levels: computational, representational/algorithmic and implementational[14,15]. The computational level is the most abstract level of description, in which the device is characterized in terms of the task to be performed. In the representational/algorithmic level, the device is described with reference to an algorithm that explicitly sets out the steps that must be followed for a task to be carried out. Finally, in the implementational level, the system is described in terms of its physical implementation. Marr's theory is arguably pertinent and applicable to the description of any complex information-processing system. In the case of language, Marr's theory would apply as follows: the computational level is the level where generative grammar belongs; the representational/algorithmic level is arguably the work of psycholinguistics, in an attempt to characterize, from a functional[16] perspective, procedural models of language processing. The implementational level would refer to the neurology associated with linguistic activity, i.e., the operational characteristics of the physical device responsible for the task to be carried out.

The adoption of Marr's proposal for the study of language cognition is not arbitrary. Neeleman and van de Koot (op. cit.) present a series of arguments in favour of approaching language cognition on the basis of Marr's viewpoint. The authors show that a potential alternative to Marr's approach, according to which the grammar would be a separate module consulted in some way by

[14] Neeleman and van de Koot changed the terminology orginally used by Marr. The terminology used here is the same as Marr's.

[15] For a different approach to the study of language cognition from the one advanced in Neeleman and van de Koot (2009) and Corrêa and Augusto (2006), see Phillips (1990).

[16] Throughout this chapter, the term 'functional' is used to refer to the mental processes involved in language cognition. The use is unrelated to what is known as 'functional grammar'.

performance systems, arguably positioning linguistics and psycholinguistics at the same level of description, does not seem to work (see Neeleman and van de Koot (op. cit.) for a detailed discussion of these issues).

If Marr's proposal about complex information-processing systems is right and it is indeed applicable to language cognition under normal circumstances (i.e., it is applicable to describe the language capacities of typically-developing individuals), it can be argued that it also applies to studying the language capacities of those individuals with language disorders. In other words, a language deficit also needs to be described at three levels.

The adoption of the tripartite model in question involves a series of implications regarding linguistic and psycholinguistc modelling that need to be addressed. It is crucial to emphasise that different criteria guide the formulation of the two types of modelling. As Corrêa and Augusto (op. cit.) point out, this becomes clear when looking, for instance, at models concerning the lexicon:

> "With respect to the lexicon, for example, both the lexicon within a formal model of language and the Mental Lexicon within psycholinguistic theories are characterised as constructs of representational nature. Only the latter, however, is conceived in order to explain phenomena concerning lexical access, such as phonological and semantic interferences that take place in linguistic performance, as well as instances of anomia and agnosia, present in the performance of individuals with aphasia. The lexicon as presented in a formal model of language, even if of representational nature, only contains information necessary for characterising its elements in terms of phonological, semantic and formal properties, in order to explain the various possibilities of syntactic combinations, to make explicit the contribution of lexical semantics to interpreting a linguistic expression and account for the phonological processes resulting from the combination between these elements in a linguistic expression." (Corrêa and Augusto, 2006: 4; my translation)

Arguing for distinct levels of description for linguistics and psycholinguistics does not entail that the two levels do not bear any relation to each other. As mentioned above, distinct criteria guide the formulation of the two types of modelling, but the properties of grammar described in the computational level need to be somehow incorporated into a description of the representational/algorithmic level, i.e., a description of the mental operations involved in the production and comprehension of linguistic utterances.

Under this framework, it seems beyond doubt that linguistic models should not be taken as a characterisation of the mental processes that take place on-line, since these models do not refer to the steps that need to be followed for language processing to occur. Having said that, the direct pursuit of empirical validation for linguistic models in performance data does not seem to be a legitimate endeavour. In other words, without the support (mediation) of psycholinguistic models, which are less abstract and which attempt to take into consideration factors concerning mental processes from a functional viewpoint, the use of performance data to authenticate linguistic models does not seem appropriate.

One common misconception related to the discussion above lies in the term *theory* in *linguistic theory*. It is often the case that researchers view psycholinguistics studies as ways to 'test' theories formulated within generative linguistics, i.e., as means to validate proposals in linguistics, as if the latter was a 'theoretical' field lacking any sort of empirical evidence of its own and psycholinguistics dealt with no theory of its own. The following quote, from Phillips and Wagers (2007), illustrates my point:

> "The term 'theoretical' in Theoretical Linguistics is all too often taken to imply that the field is somehow less concerned with empirical facts. This is unwarranted. The term merely reflects the fact that the empirical side of the field is sufficiently easy that most time is spent worrying about what the facts all mean. Similarly, psycholinguists take questions of theory seriously, although such questions take up less time on a day-to-day basis" (Phillips and Wagers, 2007: 6).

With respect to the study of SLI and the deficits that children with the disorder manifest, the discussion thus far raises the question of whether behavioural data from the performance of children with SLI can shed light on theories formulated within the discipline of linguistics. One possible answer to this question is yes, as long as psycholinguistic models are also part of the puzzle. It could be argued that behavioural data can be compatible or not with psycholinguistic models of language processing, which, in turn, can be compatible or not with linguistic models. I will return to these issues in the final dicussion, after reporting the experimental study.

The analogy between Marr's theory of vision processing and language processing is adopted throughout this thesis. In Part II, in which I present a behavioural study of grammatical gender abilities in Brazilian children with SLI, the tripartite description of the human language faculty is assumed.

Before moving on to the literature review about gender agreement, a note about the notion of development is needed. In addition to proposing that SLI should be approached via an articulated perspective that brings together linguistics and psycholinguistics, it could be argued that a developmental dimension also needs to be incorporated into the study of the disorder. Next, I make the case for the inclusion of a developmental dimension to the study of SLI

3.2.2 The case for a developmental dimension

As Corrêa and Augusto (2005) point out in a paper concerning the potential loci of SLI, "problems in the representation of functional features/categories may stem from difficulties in the way children extract linguistically relevant information from the speech signal" (p. 1). It can, thus, be claimed that a developmental dimension also needs to be incorporated into the study of SLI. Taking a developmental perspective is not equivalent to carrying out detailed longitudinal studies of language development in children with SLI. The point is, rather, that the study of SLI should pay more attention to young infants and toddlers, considering the possibility that at least part of the difficulties which are characteristic of SLI could be a direct result of problems during the course of acquisition of the language the child is surrounded by. In other words, SLI research should explore the logical possibility that some step(s) required for the acquisition of a certain linguistic structure might not follow its normal course in the case of children with the disorder.

The developmental dimension defended here is partly in line with the proposal put forward by Karmiloff-Smith and colleagues (Karmiloff-Smith, 1998; Thomas & Karmiloff-Smith, 2005), who make the case for the importance of looking at developmental trajectories: ”For developmental disorders, a central feature of explanations of the behavioral profile will be the way that language structures are acquired over time …” (Thomas & Karmiloff-Smith, 2005: 76). Their approach, however, is largely confined to a psychology perspective, not taking into account formal models of linguistic theory, argued to be crucial above.

Bearing the discussion above in mind, the next two chapters present an extended literature review of gender agreement. Chapter 4 consists of a review of linguistic aspects of gender and chapter 5 includes sections on SLI studies, acquisition studies and adult processing studies.

Chapter 4
LITERATURE REVIEW – The Linguistics of Gender

4.1 What is gender?

Gender is considered the most puzzling of the grammatical categories (Corbett, 1991). The word 'gender' derives etymologically from Latin 'genus', and originally meant 'kind' or 'sort'. Grammatical gender is essentially a system of noun classification. Importantly, however, there are various ways in which nouns can be classified across human languages and it is crucial to identify the type of classification that counts as a gender system. Hockett suggests that "genders are classes of nouns reflected in the behavior of associated words" (Hockett, 1958, apud Corbett, op. cit.). Similarly, Mathews defines gender as "a system in which the class to which a noun is assigned is reflected in the forms that are taken by other elements syntactically related to it" (Mathews, 1997, apud Corbett, op. cit.). It is, therefore, appropriate to define gender on grammatical grounds. According to Corbett (op. cit.), the determining criterion of gender is *agreement*, i.e., in order to count as gender, the noun-classifying marking should be reflected beyond the noun itself. In other words, the grouping of nouns into different classifications should determine other forms beyond the noun. In a language like Russian, for example, adjectives have to change in form according to whether a noun is feminine, masculine or neuter, which demonstrates the existence of a gender system in this language. Other ways in which nouns can be classified, such as denoting nonflesh food — which is a valid criterion for gender agreement in the Australian Aboriginal language Dyirbal — are not genders in Russian because they do not determine other forms beyond the noun.

Gender systems across human languages vary considerably in terms of the syntactic categories which are involved in gender agreement relations. These may involve determiners, adjectives, verbs, and sometimes even adverbs and conjunctions. In addition, gender systems vary in the number of possible values that can be taken: while Romance languages generally have a two-value gender system, the Northeastern Caucasian language Tsez has a four-value gender system and the Bantu languages generally have between ten and twenty different genders reflected in a complex agreement system. Languages also vary in terms of how gender is represented in lexical items. In Romance languages, the gender morpheme is generally adjoined to the right of the item. Isangu, a

Bantu language, on the other hand, overtly marks gender exclusively by means of prefixes.

From a diachronic viewpoint, gender classes might have originated on the basis of semantic motivation, as a means to distinguish words denoting classes of animate elements from words denoting classes of inanimate elements. Other semantic criteria would have followed the [± animate] distinction, such as masculine vs feminine vs neutral or rational vs irrational (Name, 2002). Although semantic motivation is a plausible account for the origins of gender systems, languages in which the gender of nouns is defined solely on the basis of semantic criteria are not particularly common (Corbett, op. cit.). A synchronic analysis of different languages reveals that morphological and phonological factors can also play an important role in defining the functioning of a gender system. Moreover, although some languages may have a preference for a particular system, factors can overlap, such as in German, whose gender system consists of a complex interplay of overlapping semantic, morphological and phonological factors (Corbett, op. cit.).

4.2 Gender in Portuguese

Like other Romance languages, Portuguese has a two-value gender system: nouns are either masculine or feminine. Also like other Romance languages, Portuguese allows the possibility of the gender feature to be either intrinsic or optional, as exemplified below.

(a) Intrinsic: all inanimate nouns such as ‘carro’ (car_{masc}) and ‘casa’ ($house_{fem}$) and a few animate nouns such as ‘criança’ ($child_{fem}$) – which refer to both male and female children;

(b) Optional: there is correlation with sex and there is variation according to the referent of the DP, as in ‘menin**o**’ (boy_{masc}) and ‘menin**a**’ ($girl_{fem}$).

When intrinsic, the value of the gender feature would be specified in the lexicon entry and, when optional, the value would vary, and the expression of such optionality would be morphological. The following table presents the classification of nouns in Portuguese with respect to the nature of the gender feature (adapted from Name, 2002).

Table 1: gender feature in nouns in Portuguese

Animacy	Optionality	
	Intrinsic	**Optional**
[-animate]	*mesa* (table$_{fem}$) *livro* (book $_{masc}$)	-
[+animate]	*girafa* (giraffe$_{fem}$) *criança* (child$_{fem}$)	*amigo/a* (friend $_{masc}$/$_{fem}$) *coelho/a* (rabbit $_{masc}$/$_{fem}$)
	dentista (dentist $_{masc}$/$_{fem}$) *colega* (colleague $_{masc}$/$_{fem}$)[17]	

Although the intrinsic gender of [-animate] nouns and some [+animate] nouns is arbitrary, there seems to be some phonological regularity in many languages, which allows generalizations to be made. This is also the case in Portuguese. There is a co-relational pattern between the phonological form of the noun and its gender (Corrêa & Name, 2003). Nouns ending in an unstressed 'o' ([u]) are usually masculine and nouns ending in 'a' are usually feminine. However, this is not always the case, and other noun ending patterns are very frequent, such as masculine words ending in 'a' and feminine ones ending in 'o': 'planeta' (planet$_{masc}$), 'problema' (problem$_{masc}$), 'tribo' (tribe$_{fem}$), 'foto' (photo$_{fem}$). Nouns ending in unstressed 'e' ([i]) or in consonants can be either masculine or feminine: 'dente' (tooth$_{masc}$), 'ponte' (bridge$_{fem}$), 'tambor' (drum$_{masc}$), 'flor' (flower$_{fem}$).

4.3 Gender in other elements of the Determiner Phrase

From a syntactic point of view, the gender of the noun controls agreement with syntactically related constituents. In Portuguese, agreement can be observed in the morphology of determiners, adjectives and participial forms. Gender is always manifest, both in singular and plural forms, in determiners and adjectives (those ending in – *o*). Adjectives ending in – *e* are invariant.

Table 2: Morphological pattern of the Portuguese gender classes in the Determiners (adapted from Name, 2002)

	Masculine Determiners	**Feminine Determiners**
Definite articles	**o** (s) *	**a** (s)
Indefinite articles	um (ns)	um**a** (s)
Demonstratives	este/esse/aquele (s)	est**a**/ess**a**/aquel**a** (s)

* (s) stands for the number morpheme marking plural

[17] Nouns like 'dentista' and 'colega' do not go through the morphological process of inflection, but they do require different gender marking in determiners and adjectives depending on the referent. A male dentist is, thus, referred to as 'o dentista' (the$_{masc}$ dentist), while a female dentist is referred to as 'a dentista' (the$_{fem}$ dentist). It is not clear whether these nouns should be treated as two lexical entries with one intrinsic gender feature each or only one lexical entry with two intrinsic gender features.

Within the category D, the feminine subset presents phonological regularity in the ending (-a). This regularity is similar to the endings of many nouns and inflected adjectives. On the other hand, not all forms of masculine determiners present phonological similarity with endings on nouns and adjectives.

The adoption of the notions optional and intrinsic, discussed in 4.2, is compatible with the idea that masculine forms are default forms in languages with a two-value gender system. In other words, masculine forms are often considered unmarked forms, i.e, absent of features, while feminine forms are marked. This is consistent with the use of masculine plural forms to refer to mixed groups. While a masculine plural noun can be used in reference to, for example, a group formed both by female and male individuals, given the absence of a gender feature in masculine nouns, the use of feminine nouns as reference to mixed groups is not allowed.

Next I look at how linguistic theory has been investigating phenomena pertaining to determiner phrases (DPs) and adjective phrases (APs). The discussion of both structures is relevant for the experiments that are reported in chapter 7. Following the discussion of the configurations of DPs and APs, I present an overview of how linguistic theory deals with the phenomenon of agreement.

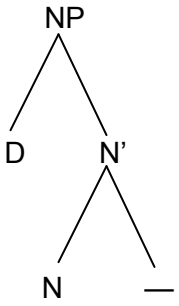
4.4 Determiner Phrases and Adjective Phrases

4.4.1 Determiner Phrase

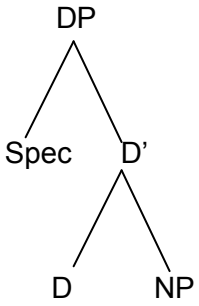
Traditionally, noun phrases (NPs) were considered maximal projections of a lexical head N, as the diagram in (1) illustrates. In early versions of generative theory, the NP had a unique head, N, and the other components of the NP occupied either the complement position(s) or the specifier position(s). Accordingly, the article was represented as being an integral part of the projection of N (Jackendoff, 1977). Later, the work of Abney (1987) and others drew attention to potential structural similarities between verbal clauses and noun phrases. They proposed that noun phrases, like clauses, are embedded under a higher functional projection and headed by a functional element labeled D(eterminer). NPs are thus treated as a complement of the functional categories

of Determiners, setting up the projection known as Determiner Phrase (DP). In other words, the DP hypothesis postulates that, in the same way that the projection of the verb is dominated by functional material, the projection of the noun is part of a larger functional complex, the DP. The Determiner, thus, gains a central role in the nominal system (for a discussion of the motivations for the DP hypothesis, see Radford, 1997; Carney, 2002; and Alexiadou, Haegeman & Stavrou, 2007). The DP hypothesis has achieved a broad consensus and will be adopted throughout this thesis. (2) below illustrates the Determiner Phrase:

(1)[18]



(2)



4.4.2 Adjective phrase

4.4.2.1 The status of adjectives

There is a long-standing debate in the literature regarding the nature of adjectival positions, but adjectives (and adverbs) are considerably less studied than other major lexical categories such as nouns and verbs. Different proposals have been put forward in an attempt to account for the status of adjectives across human languages. These studies have discussed, for example, the ordering of adjectival strings, cross-linguistic asymmetries concerning the order of adjectives in relation to nouns, and semantic interpretation of various adjective types to explain their syntactically different structural positions.

Three alternative proposals for the syntactic status of adjectives will be discussed here: (a) the adjunction analysis of adjectives, (b) the head analysis of adjectives and (c) the specifier analysis of adjectives.

Before introducing the different proposals for the status of adjectives, I present a description of how adjectives behave in Portuguese[19], and in Romance

[18] Abney (1987) suggests that in the traditional account D would have to be a maximal projection, given X-bar theory. This issue is not relevant to the current work.

languages in general, in contrast to the behaviour of adjectives in English, and a brief note on how adjectives have been semantically classified in the literature.

4.4.2.2 The behaviour of adjectives in Portuguese

As the examples in (3) illustrate, some adjectives in Portuguese can appear both in pre-nominal and post-nominal positions:

(3) a. as flores bonitas
the flowers beautiful
'the beautiful flowers'
b. as bonitas flores
the beautiful flowers
'the beautiful flowers'

There is, however, a strong tendency for adjectives of the type described in (3) to appear to the right of the noun. In addition, there are some adjective types that can only occur in post-nominal position:

(4) a. as flores campestres
the flowers country
'the country flowers'
b. *as campestres flores
the country flowers
'the country flowers'

And other adjective types which can only occur in pre-nominal position, as shown in (5):

(5) a. o suposto criminoso
the supposed criminal
'the supposed criminal'
b. *o criminoso suposto
the criminal supposed
'the supposed criminal'

As is widely known, the adjectives illustrated in (3), (4) and (5) appear pre-nominally in English.

[19] The description presented here is based on Gonzaga (2004) and Schmitt (1996).

In Portuguese DPs with more than one adjective, their order is not free. For example, adjectives like 'errado' (wrong) must be in the rightmost position, while colour adjectives like 'branco' (white) usually appear after the noun (Schmitt, 1996), as (6) illustrates:

(6) a. o livro branco errado
the book white wrong
'the wrong white book'
b. *o livro errado branco
the book wrong white
'the wrong white book'

According to Ticio (2003), adjectives in pre-nominal position in Spanish do not accept complements. The same is true for Portuguese, as the examples in (7), adapted from Ticio (op. cit.), show.

(7) a. *el orgulloso de su hija padre (Spanish)
*o orgulhoso de sua filha pai (Portuguese)
'the proud of his daughter father'
b. el padre orgulloso de su hija (Spanish)
o pai orgulhoso de sua filha (Portuguese)
'the father proud of his daughter'

As regards English, pre-nominal adjectives do not accept complements either, as the examples above illustrate.

4.4.2.3 Semantic-based classifications of adjectives

The different patterns in which adjectives are positioned in the nominal domain have been associated with semantic patterns. The task of capturing semantic generalisations among adjectives is, however, not trivial. While it is not always possible to have a clear cut classification, it is evident that some patterns are very strong.

An adjectives can be said to be 'predicative' when it is used in post-copular position, in contrast with an attributive use, when it is employed as a noun modifier located within the boundaries of a nominal phrase. Alexiadou et al (2007), however, offer an arguably more adequate definition of the term 'predicative'. The authors claim that 'predicative' is better defined in terms of the possibility of an adjective to be paraphrased with a copular construction: "If an

adjective modifier does not allow the paraphrase with a copular construction, it is termed 'attributive'; if it does allow it, it is called 'predicative'" (Alexiadou et al, 2007: 291). Based on this proposal, the adjectives in (8) are predicative, as they can be paraphrased with a post-copular construction, while the adjectives in (9) are non-predicative, as they can only be used within the boundaries on the nominal phrase.

(8) a. the interesting problem (Alexiadou et al, 2007)
The problem is interesting.
b. the proud student
The student is proud.

(9) a. the former policeman
*The policeman is former.
b. a mere detail
*The detail is mere

While both attributive and predicative uses of adjectives in English mostly occur pre-nominally, the Romance languages present a different, somewhat clearer, pattern. In this language group, DP-internal pre-nominal adjectives are typically attributive and post-nominal adjectives are typically predicative. In other words, post-nominal adjectives in the Romance languages generally allow for paraphrase with the copular construction while pre-nominal adjectives usually do not. I reproduce the constructions in (3) and (5) to illustrate this point. (3a) allows for paraphrase, while (5a) does not.

(3a) as flores bonitas
As flores são bonitas.
'The flowers are beautiful.'
(5a) o suposto criminoso
*O criminoso é suposto.
'The criminal is supposed.'

Additional classifications of adjectives have been proposed. For example, predicative adjectives, in the sense of Alexiadou et al, are often called intersective (when they are part of the class for which the resulting projections designate a subset of entities that belong to two sets at the same time) in contrast with non-intersective adjectives (their combination with nouns denotes only one property instead of two) (see Knittel (2005) and Laenzlinger (2005) for

details). Further, Demonte (1999) establishes a division of adjectives into three major groups: relational adjectives, qualitative adjectives and adverbial adjectives.

It is outside the scope of the current thesis to provide a thorough review of semantic-based classifications of adjectives, so I will not discuss the proposals above any further. It is important, however, to state that most of the discussion on syntactic analyses of adjectival placement to follow will focus on what is known as predicative adjectives, as these are the adjectives the experiments to be reported in chapter 7 explore.

4.4.2.4 Syntactic analyses of adjectives: an overview

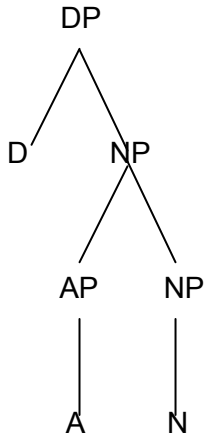
This section introduces some of the main approaches to the analysis of DP-internal adjectives. As mentioned earlier, I will present three different groups of hypotheses about the placement of adjectives in the DP: (a) *Adjectives as adjuncts* (Bernstein, 1991; Ticio, 2003; Picallo, 1991; among others); (b) *Adjectives as heads* (Abney, 1987; Delsing, 1993); and (c) *Adjectives as specifiers* (Cinque, 1994; 2004; 2005; Basic, 2004; Giusti, 2002; and others).

(a) *Adjectives as adjuncts*

The analysis of adjectives as adjuncts dates back to at least the principles and parameters framework (cf. Chomsky, 1981) and can be considered the 'traditional' approach for Adjective Phrases within generative grammar. According to Svenonius (1994, apud Pysz, 2006), one of the arguments usually put forward to support adjunction analyses of adjectives refers to the properties of *optionality* and *iteration*: adjunction, which is optional by assumption, matches the fact that the presence of adjectives in nominal phrases is not obligatory and the possibility of iterating adjectival modifiers in a nominal phrase fits with the idea that syntactic adjuncts can be freely added to a given structure.

Originally, it was proposed that the AP was adjoined to N'. Over the years, different authors have proposed other adjunction sites for adjectives (see Pysz (2006) for an overview of adjunct proposals and the different functional projections adjectives can be adjoined to). I will restrict the discussion in this section to the proposal that has adjectives adjoined to the NP. The diagram in (10) illustrates a structure from English, on which the adjunction proposal was originally based and in which the adjective is in pre-nominal position.

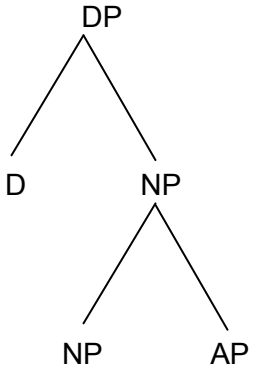
(10)



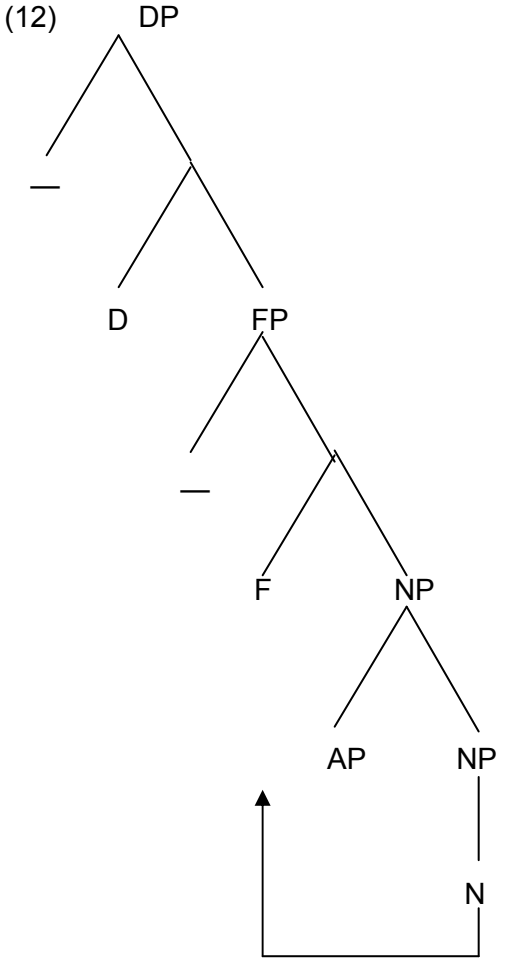
When we turn to Romance languages, in which most predicative adjectives appear in post-nominal position, two possibilities arise:

1. The AP is base generated to the right, as illustrated in (11)
2. The AP is base generated to the left and there is movement, as shown in (12)

(11)



(12)



The two diagrams above have been used as attempts to explain the adjective ordering typically found in Romance languages, i.e. (predicative) adjectives following the N. In (11), the adjective has been adjoined to the right of the NP and there is no need to move the N leftwards to get the typical order in Romance. In (12), on the other hand, the adjective has been adjoined to the left of the N. The proponents of (12) resort to the N moving leftwards to get the surface order of Romance[20]. A crucial question that needs to be asked when N-movement is proposed refers to what actually triggers the N to move in some languages but not others. It has been proposed that the absence vs. presence of noun raising lies in morphological properties of nominal systems: if the nouns in a particular language show inflection for gender and number then N-movement to the respective functional projection takes place. Bernstein (1993), for example, contrasts French, a language in which nouns inflect for number and supposedly has noun movement, with Walloon, a language in which plural is never realized in its nouns and, therefore, shows no noun movement.

The proposal illustrated in (12), nevertheless, has been criticised by Lamarche (1991) and others authors. They have claimed that the N-movement approach does not really apply to the Romance languages, as it cannot, in fact, account for the linear ordering of structures with more than one adjective in this language group (Lamarche, 1991; Ticio, 2003; Alexiadou et al, 2007). The line of reasoning is roughly as follows. Sproat and Shih (1988) have proposed that universal ordering restrictions within a sequence of certain classes of adjectives apply. In other words, they noted that there is a relative ordering of the different classes of adjectives which is by and large the same across languages. For instance, if colour adjectives appear closer to the noun than size adjectives in a particular language, Sproat and Shih's analysis predicts that this will be true in other languages as well. The examples in (13), below, adapted from Ticio (2003) and Lamarche (1991) show the typical ordering of a DP with multiple adjectives in Romance and English.

[20] The proposal of Adjectives adjoined to the left is compatible with Kayne's (1994) antisymmetry theory, a theory of syntactic linearization according to which right adjunction is not allowed. In addition, Kayne proposes that there is a universal order of constituents, namely Specifier-Head-Complement, and that constructions with non Specifier-Head-Complement order are derived through movement.

(13) a. un coche blanco oxidado (Spanish)

b. um carro branco enferrujado (Portuguese)

a car white rusty

c. a rusty white car

As can be seen above, ‘blanco’, ‘branco’ and ‘white’ all appear closer to the noun in each language, leaving ‘oxidado’, ‘enferrujado’ and ‘rusty’ further from it. The Noun-movement approach illustrated in (12) claims the only difference between the order of adjectives relative to one another in Romance and English is the position of the nominal head and their different surface position is attributed to the raising of the N in Romance (but not in Germanic). The predicted orders for the two language groups are abstractly presented below in (14), showing that the Noun-movement approach, nevertheless, is not able to account for the data, as it derives ungrammatical structures.

(14) a. Adj_2 Adj_1 N (English)

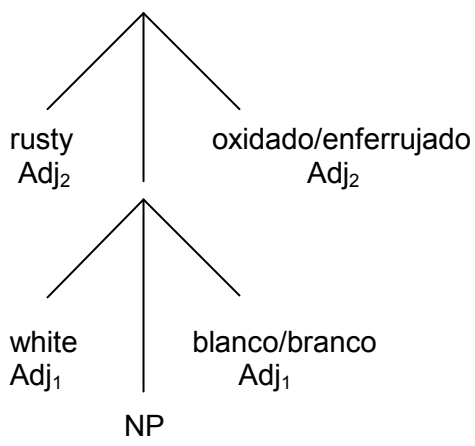
b. * Adj_2 Adj_1 N (ungrammatical in Romance)

*un coche oxidado blanco (Spanish)

*um carro enferrujado branco (Portuguese)

According to Lamarche (op. cit.), in contrast to an analysis involving Noun-movement, a proposal that argues that post-nominal adjectives in Romance are generated to the right of the noun predicts a sequence for this language group that mirrors the sequence found in languages like English and, therefore, provides a more appropriate account for cross-linguistic adjectival ordering. The diagram in (15) illustrates what Lamarche calls the *mirror image phenomenon*.

(15)



In addition to the empirical problems raised above, the N-movement approach to adjectival placement encounters difficulties related to its motivation. As mentioned earlier, it has been proposed that the motivation for noun raising lies in morphological properties of nominal systems and the absence vs. presence of noun raising in a particular language depends on it having a 'rich' or 'non-rich' inflectional system. However, it has been shown that languages can present a 'rich' morphology and lack noun movement. Alexiadou (2001) has shown that this is the case of Greek, whose nouns manifest a 'rich' nominal inflection system (all nouns inflect for gender, number and case) and there is no N-movement (the head noun always follows the adjectives that modify it). Similar observations can be made for West Flemish: in spite of its robust overt morphological markings, the language has rigidly pre-nominal adjectives (Alexiadou et al, 2007).

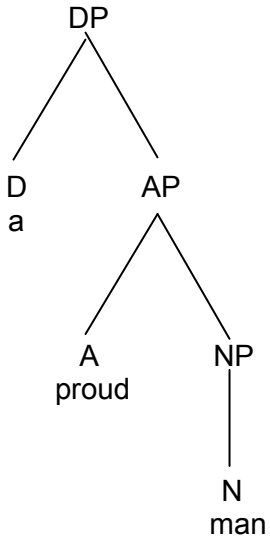
(b) *Adjectives as heads*

Abney (1987) proposes an alternative analysis for adjectives. His proposal focuses on pre-nominal adjectives in English. Under his 'adjective as head' perspective, (pre-nominal) adjectives are heads in the extended projection of the noun phrase. The functional head[21] D selects AP as a complement, under the phenomenon termed as functional selection (f-selection)[22]. The structural diagram corresponding to Abney's analysis for English is provided in (16).

[21] Lexical heads are nouns, verbs and adjectives and, for some researchers, prepositions. The extended projection of a lexical head consists of the projection of the lexical head and all the projections dominating it, up to the point that a new lexical head is merged. Accordingly, a functional head is any head in the extended projection of a lexical head which is not the lexical head itself. Typically, functional categories comprise words which have no descriptive content, as opposed to content words, which have descriptive content. Here, the debate on the status of adjectives refers to whether adjectives should always be categorised as belonging to the list of lexical heads or whether they can sometimes be functional heads.

[22] Functional selection is defined as 'a syntactic relation between functional elements (attributive adjectives in this case) and their (f-selected) complements. Complements which are f-selected are not arguments and do not require case (Abney, 1987, apud Pysz, 2006).

(16)



Abney argues that his approach is able to account for the fact that, in languages like English, adjectives in the pre-nominal domain are not allowed to take complements, as (17) shows.

(17) *a proud [of his children] man

However, while the prediction that pre-nominal adjectives will not take complements is true for English, the same is not valid for a number of other languages, including Dutch and German, which belong to the same language family as English. The Dutch example in (18) illustrates the possibility of pre-nominal adjectives taking complements.

(18) Een op zijn kinderen trotse vader
a on his children proud-INFL father
A father proud of his children

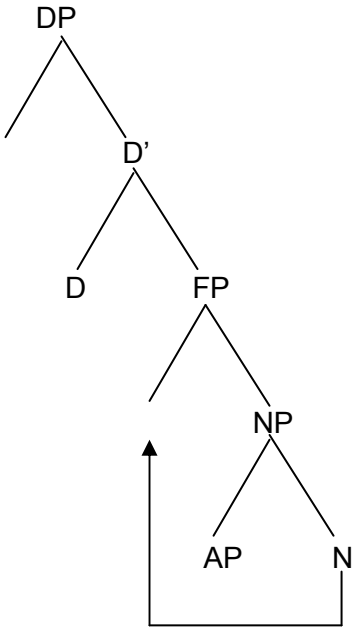
Other languages which allow adjectival complements in the pre-nominal domain are Russian and Polish (Pysz, 2006). In addition to predicting restrictions that do not hold in many languages, Abney's analysis presents another problem that has a direct bearing on this thesis: the head analysis does not handle post-nominal adjectives within the DP, the most common pattern found in Romance languages and, therefore, does not accommodate the adjectival system of this language group.

(c) *Adjectives as specifiers*

Cinque is the main representative of the 'adjectives as specifiers' type of analysis, which is widely used in attempts to account for the adjectival patterns found in Romance languages. Here I will review Cinque's proposals in different publications over the last decade or so as some important changes have occurred. Cinque (1990) proposes that (i) APs are generated in specifier positions of distinct functional projections, (ii) nouns can undergo partial N-movement. Following a similar argument to the 'adjunct to the left' approach outlined above, Cinque (1990) proposed that N-movement is responsible for adjectives appearing in pre- or post-nominal position. In Cinque's proposal, however, the N in the Romance languages moves to the head of a functional projection to the left of the NP, a phenomenon that would not take place in English. Cinque assumes a universal hierarchical order for adjective positions, which is claimed to correspond to world knowledge notions such as size, color, nationality. He claims that this hierarchy cannot be accommodated within an adjunction analysis as adjunctions are normally intended to be free. According to the specifier analysis, the universal hierarchy is determined by selectional properties of the functional categories. The specifiers of functional projections in which adjectives are generated are situated between the highest functional projection (i.e. DP) and the lowest lexical projection (i.e. NP) of the nominal phrase. Each and every adjective present in the phrase is said to be uniquely associated with a specific functional projection[23] (Pysz, op. cit.). The specifier analysis is also claimed to account for the fact that there is a maximum number of Adjectives per DP. According to Cinque, the adjunction analysis cannot explain this restriction, since the number of adjuncts (contrary to the number of functional categories) allowed in a given structure is unlimited. The diagram in (19) illustrates structure according to Cinque's analysis.

[23] Under this proposal, Functional Projections are said to be specialised for different types/groups of adjectives: there would be a FP for colour adjectives and one for size adjectives, for example.

(19)



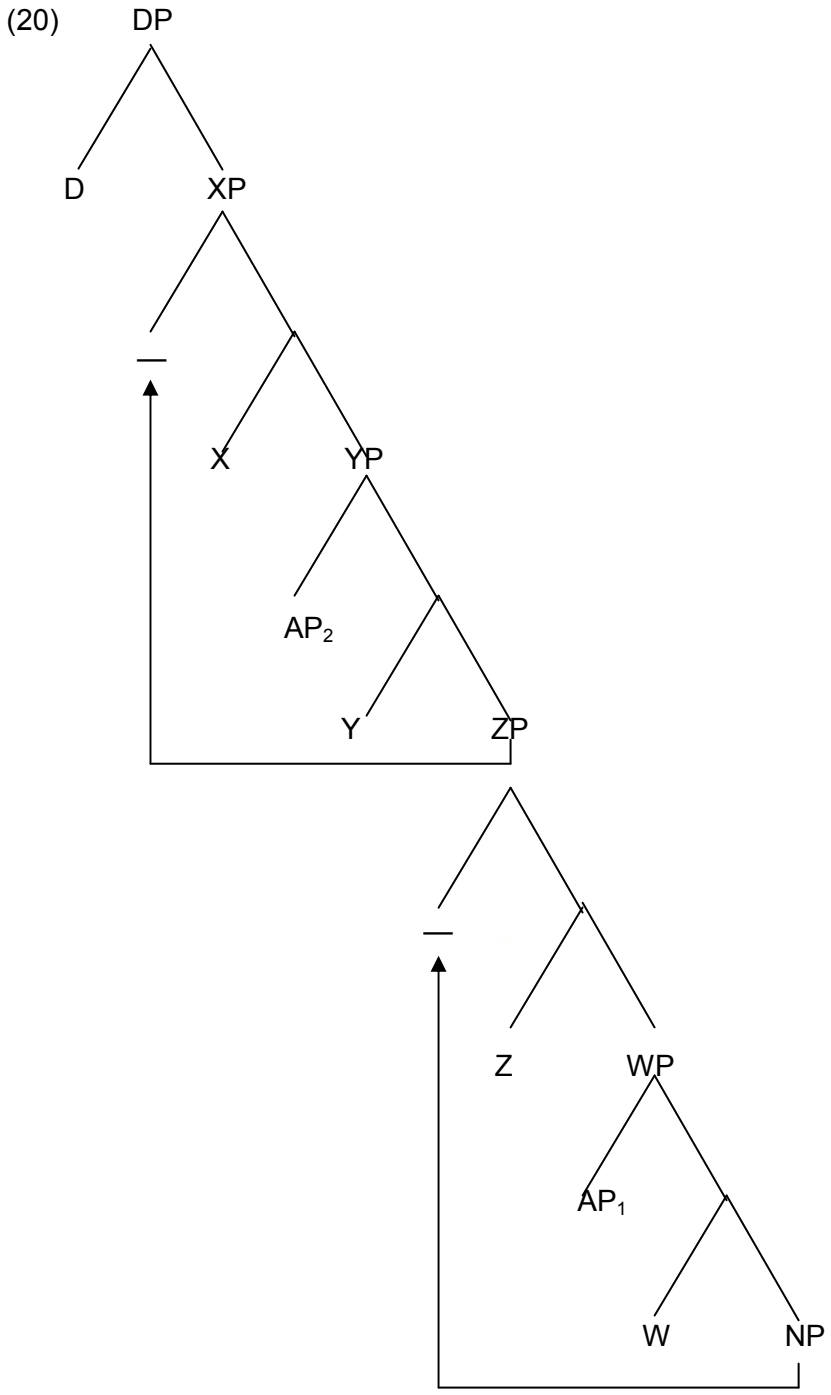
However, when it comes to the surface differences between Romance and English, Cinque's 1990 version of the specifier analysis is subject to the same criticisms I have put forward when discussing the adjective as adjunct analysis: it cannot account for the phenomenon Lamarche calls *mirror image* (cf. page 77).

In an attempt to deal with the problems just raised, in later versions of his hypothesis, Cinque (2003; 2005) reformulates the notion of 'movement' which is responsible for the derivation of different surface adjectival positions in Romance and Germanic. He introduces the notion of 'snowballing' or 'roll-up' movement, an operation in which a constituent moves to the specifier of a higher projection and then pied-pipes[24] the containing projection in the next cycle of movement. Snowballing movement triggers movement of the NP and not of the N as proposed in the earlier versions of Cinque's specifier analysis. Cinque suggests that, on its way up the derivation, the NP collects the lowest adjective and moves to the next specifier; when the NP moves to the highest specifier it has pied-piped all the adjectives it encountered. The diagram in (20) illustrates the roll-up movement[25].

[24] Pied-piping is "a process by which a moved constituent (or set of features) drags one or more other constituents (or sets of features) along with it when it moves. For example, if we compare a sentence like *Who were you talking to?* with *To whom were you talking?*, we might say that in both cases the pronoun *who(m)* is moved to the front of the sentence, but that in the second sentence the preposition *to* is pied-piped along with whom" (Radford, 1997).

[25] Cinque's (2005) discussion is centered around what he claims to be a single, universal, order of Merge, namely Dem > Num > Adj > N. I have somewhat simplified his proposal and left out the details that were not pertinent to this thesis.

(20)



At the end of the derivation, the order will be

(21) D NP AP_1 AP_2

as opposed to (22), which would be the order obtained with the NP moving without any pied-piping of other material

(22) D NP AP_2 AP_1

While the introduction of the notion of snowballing or roll-up movement does allow the specifier analysis of adjectives to account for adjectival ordering in Romance, it can still be criticised on conceptual grounds.

Abels and Neeleman (2009) have identified problems of two different sorts in Cinque's analysis. They propose an alternative account for linear asymmetry found in language to the one proposed in Kayne's (1994) Linear Correspondence Axiom (LCA), which is the basis of Cinque's theories. According to the LCA, specifiers universally precede heads and heads universally precede their complements and movement (only allowed to the left) is applied in order to derive cross-linguistic surface orders. Abels and Neeleman argue, however, that "a weaker theory, one that embraces only the restriction to leftward movement and jettisons the idea that base-generation is universally ordered, is to be preferred" (p. 60). First, Abels and Neeleman claim that, after applying what they have named 'shrinking'[26], a mechanical procedure that preserves gross constituency, to a tree in Cinque's system, it becomes very similar to the representation stemming from the proposal which adjoins the adjective to the right or the left of the noun (cf. page 75). According to Abels and Neeleman, the application of the 'shrinking' algorithm does not change the c-command relations between the functional nodes in the Cinquean tree. In other words, the application of 'shrinking' shows that Cinque's tree does not offer any new insights about how adjectives behave, as it advances the same 'grouping' of elements as the traditional theory of adjunction does. A second line of criticism that Abels and Neeleman advance refers to the validity of the movement operations required by Cinque's theory. In spite of the observation that, after the application of 'shrinking', the trees in Cinque's proposal and in the traditional adjunction analysis are essentially equivalent, the two theories are not identical. Abels and Neeleman argue that Cinque's theory, and LCA-based work more generally, require a type of movement which, instead of contributing to strengthening the theory, weakens it: "the movements required to reconcile the LCA with the attested word-order patterns stand in the way of arriving at a restrictive theory of movement" (p. 73). Abels and Neeleman's criticism is broad in scope and, arguably, has implications for many other syntactic domains besides the DP. As

[26] 'Shrinking' is defined as follows by Abels and Neeleman: "Prune the Cinquean tree by deleting the functional heads (W, Agr*W*, X, Agr*X*, Y, and Agr*Y*) and their intermediate projections maintaining dominance relations" (p. 68).

this discussion is not directly relevant for the current thesis, I will not develop Abels and Neeleman's arguments in any further detail.

Before ending this section, it is important to highlight that, despite the problems we have identified in Cinque's analysis, it offers relevant insight into some restrictions on the hierarchy of adjectives. While we have seen that his proposal that APs are generated in specifier positions of distinct functional projections and that nouns undergo partial N-moment in Romance cannot be maintained, some basic aspects of his claims about the hierarchy of adjectives are valid and need to be accommodated in alternative accounts of adjectives. Cinque (1994) proposes a structure in which a strict and highly elaborated series of functional heads governs the allocation of adjectives according to the class of property that they denote. Cinque's hierarchy is presented in the serialization in (23).

(23) poss > cardinal > ordinal > quality > size > shape > colour > nationality

Nevertheless, the idea that classes of adjectives across languages follow a sequence that is not arbitrary and that allows no recursion does not need to be translated into different functional projections and their selectional properties, as Cinque proposes. Instead, this hierarchy of adjectives is arguably better defined in terms of a 'linguistic principle' that governs adjective adjunction[27].

4.4.3 Summary and concluding remarks

In this chapter we have seen a characterisation of the DP hypothesis and how it has changed the internal structure of the nominal domain. We then looked at the behaviour of adjectives cross-linguistically, with a strong focus on how they occur in Portuguese, and we briefly looked at semantically-based classifications of adjectives. Crucially, we saw that there is a strong tendency for predicative adjectives to appear in post-nominal position in Portuguese (and in Romance more generally). We have also seen that, when more than one adjective is present, there are clear restrictions as to which adjective precedes the other. In the second half of the chapter we looked at three of the main proposals for the syntactic status of adjectives: (a) the adjunction analysis of adjectives, (b) the head analysis of adjectives and (c) the specifier analysis of adjectives. It has

[27] For a related argument involving adverbials in the verbal extended projection, see Nilsen 2003.

been shown that one version of the adjunction analysis (the adjunction to the left analysis), according to which adjectives are adjoined to the left of the noun and N-movement takes place in an attempt to account for cross-linguistic ordering differences, presents a number of problems and, therefore, needs to be rejected. It has also been shown that the head analysis of adjectives presents difficulties: while it apparently accounts for some observed restrictions in English, it does not for languages that also belong to the Germanic group, such as Dutch and German, nor does it deal with post-nominal adjectives, the most common position for adjectives to occur in Romance languages. Therefore, the head analysis is also rejected. Next, the specifier analysis of adjectives was considered and we saw that both versions of the analysis are subject to strong criticisms: the first version is not able to account for constructions with multiple adjectives and, in the second version, when the mechanism 'shrinking' is applied, the resulting tree is essentially the same as the tree which stems from the traditional, 'free' adjunction analysis (to the left or to the right, depending on the language). In addition, the second version of Cinque's theory makes use of movement types which are problematic and heavily depends on theory-internal restrictions that do not seem to be motivated.

Given the issues raised above, which have allowed us to reject the head, adjunction to the left and specifier analyses of adjectives, it seems plausible to consider that the adjunction to the right analysis is the most appropriate in capturing the behaviour of predicative adjectives in Portuguese and in the Romance languages.

4.5 Agreement

4.5.1 Introduction

The goal of this section is to give a (selective) overview of how linguistic theory deals with the phenomenon of agreement. I will discuss ways in which the Theory has handled agreement between a determiner and a noun, as well as agreement between a noun and an adjective.

Agreement can be defined, in general terms, as a syntactic process in which syntactic relations between different items are established and features are shared. Corbett (2006) defines agreement as a phenomenon occurring "when grammatical information appears on a word which is not the source of that

information". Also according to Corbett (2006), "despite extensive research, agreement remains deeply puzzling". In comparison with agreement elsewhere, less studies have been carried out on the nominal domain. Nevertheless, in spite of the less favourable scene, a number of relevant issues have been raised and looked at.

4.5.2 Agreement in the Minimalist Program

In the framework of the minimalist program (Chomsky, 1995), *agreement* is conceived as a feature checking operation. A given language, in minimalist terms, is composed by a computational system (arguably common to human language) and a lexicon, which is constituted by a series of lexical features, grouped according to three different criteria: (i) semantic features (interpreted in the semantic interface); (ii) phonetic features (interpreted in the phonetic interface); and (iii) features that are not interpreted in either of the two interfaces and, therefore, must be checked during the course of derivation. Furthermore, features can be also be divided into: Formal Features (FFs) that are subject to syntactic operations (such as categorial, ϕ features – gender, number and person – and case features) and other features that are not relevant for syntax. Together with features of person and number, gender is part of the group of phi-features (ϕ features). According to the proposal, ϕ-features can be either intrinsic or optional. Chomsky defines gender as an intrinsic feature in nouns and an optional feature in determiners, adjectives, for example. Under the minimalist framework, ϕ features are taken to be interpretable or non-interpretable, i.e., semantically read at Logical Form. Gender feature is considered to be [+interpretable] in nouns and [–interpretable] in determiners and adjectives. Under the minimalist perspective, *agreement* occurs because there are non-interpretable formal features that need to be checked, given the *Full Interpretation Principle*.[28].

Before moving on to the next section, it is important to point out that Chomsky's proposal that gender feature is [+interpretable] in nouns and [–interpretable] in determiners and adjectives faces minor difficulties with languages like Portuguese. As seen in 4.2, Portuguese (and other Romance languages) have two different types of nouns: [+animate] and [-animate]. We also

[28] According to the *Principle of Full Interpretation*, PF and LF must contain only elements that are readable at these levels, requiring that non-interpretable features are checked. Therefore, non-interpretable ϕ features (gender, number and person) of functional categories are attracted by their interpretable counterparts in lexical categories such as nouns and verbs and, thus, checked (deleted)

saw that Portuguese allows the possibility of the gender feature to be either intrinsic or optional. All [-animate] nouns, such as 'mesa' ($table_{fem}$), and some [+animate] nouns, such as 'criança' ($child_{fem}$), have intrinsic gender, while many [+animate] nouns (such as 'menino/menina' (boy/girl) would qualify for optional gender. It is, thus, not possible to say that all nouns in Portuguese have a gender feature which is interpretable, as Chomsky says. The proposal introduced by Pesetsky and Torrego (2004) provides an interesting alternative to that of Chomsky's. The two proposals share a number of similarities, but Pesetsky and Torrego's work deals more adequately with the interpretability issues raised above. The authors combine the conception of 'feature sharing' with a proposal that valuation and interpretability of features are actually independent concepts. They do not discuss the interpretability of the gender feature of DP items and, thus, their position regarding this type of feature is not clear. Nevertheless, it is possible to extend their line of reasoning and propose a way in which the interpretability of gender feature in languages like Portuguese would be better characterised. Their proposal would result essentially in the following: nouns which allow optional gender would be interpretable and come from the lexicon with a value for gender, while nouns with intrinsic gender would also come from the lexicon with a value for gender, but would be uninterpretable. Both determiners and adjectives, on the other hand, would have uninterpretable gender features, which would come from the lexicon unvalued.

Having looked at how the minimalist framework handles agreement in general terms, I next discuss how the field developed in relation to the nominal domain.

4.5.2.1 Agreement in the nominal domain

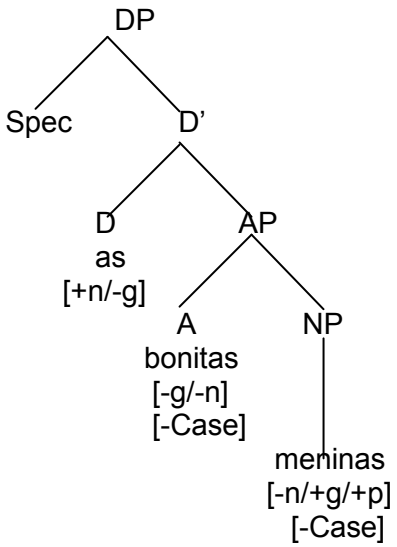
Regarding the nominal domain specifically, in the 1990s, after the introduction of the DP hypothesis (discussed in 4.4.1), agreement within this domain started to be investigated. Researchers, nevertheless, were initially concerned with identifying the nominal middle field, i.e., the functional phrases arguably found in between the D and the N. According to Roehrs (2006), the debate on those intermediate functional phrases has undergone substantial revisions and there is little consent on the kind, number or sequence of them. Ritter (1993), for example, proposed the existence of a number phrase (NumP), while gender would be a feature, with no functional category status. Picallo (1991), on the other hand, proposed a functional category for gender, which would be located in

between the NumP and the NP. Further, Di Domenico (1997) advances the idea of the gender feature of Nouns being projected in the node of the NumP, based on Greenberg's Universal 36, according to which languages only have gender marking if they also have number marking. Because these proposals are not directly dealing with agreement per se, and, therefore, not directly relevant to the discussion in this chapter, I will not expand them any further (for a detailed discussion of the proposals, see Name (2002)). For ease of exposition, I will make use of a neutral representation of the nominal middle field. Next I look at a more recent proposal, which actually discuss nominal agreement under the framework of the minimalist program.

4.5.3 Probe and goal – Magalhães (2004)

Magalhães (2004) discusses the minimalist agreement configuration making use of data from Brazilian Portuguese (BP). She claims that D has [+interpretable] number features and [-interpretable] gender features, while N has [-interpretable] number features and [+interpretable] gender and person features. Using the concepts of *probe* and *goal*[29], Magalhães advances a proposal of feature valuation on the basis of the configuration illustrated in (24), below[30]:

(24)



Under the above configuration, Magalhães argues that the first agreement relation is established between the adjective and the noun. The [+interpretable]

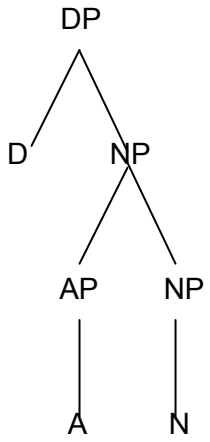
[29] The notions of *probe* and *goal* were introduced by Chomsky in an attempt to account for the way features are valued in the course of a derivation. An [-interpretable] feature is said to be a *probe* which has to search for an [+interpretable] feature of similar nature, named *goal*.

[30] The '+' and '-' signs preceding the symbols for 'gender', 'number', 'person' and 'case' in (1) refer to the interpretability (or lack of) of the features in question.

gender feature of the noun meets the [-interpretable] gender feature of the adjective, resulting in its valuation. Further, the [-interpretable] gender feature of the D triggers another agreement relation. According to Magalhães, the adjective is found by the D and has its [-interpretable] number feature valued by the [+interpretable] number feature of the D. However, the gender feature of the adjective is [-interpretable] and, thus, cannot value the equivalent feature in the D, which is also [-interpretable]. Consequently, the D needs to continue its search for an [+interpretable] gender feature and finds it in the noun.

Magalhães claims her proposal accounts for feature valuation within the DP. And, in fact, there does not seem to be any problems in the way she deals with determiner and noun agreement. However, at the same time that she includes an adjective in the configuration she discusses, she explicitly avoids entering the debate over the potentially different positions in which an adjective can be located within the DP. In order to attempt to handle adjective agreement within the minimalist framework, Magalhães must analyse the adjective as a head, which has to look into its c-command domain to try and find matching features. If we tried to incorporate the proposal discussed in 4.4.2.4, according to which predicative adjectives are best analysed as adjuncts to NP (either to the right or the left), to Magalhães's proposal, it becomes clear that it is not able to fully account for feature valuation as she claims. Recall that, as shown in (10), reproduced below, if AP is treated as an adjunct, then its head A does not c-command the noun and can therefore not act as a probe.

(10)



In sum, while Magalhães's proposal might be considered an adequate account of agreement between determiner and noun, agreement between noun and adjective remains unresolved under her analysis.

The discussion of Magalhães's analysis of agreement under the minimalist framework has allowed us to conclude that the probe and goal configuration is not able to account for agreement between a noun and an adjective. There seems to be two distinct configurations regulating, on the one hand, agreement between a determiner and a noun and, on the other hand, agreement between a noun and an adjective. Although Magalhães's analysis is able to account for determiner and noun agreement, it is possible to say that the probe and goal approach is very specific to one version of the minimalist theory. In light of this, next I discuss an alternative proposal for dealing with determiner and noun agreement, one which is compatible with other versions of generative theory. Subsequently, I discuss a potential way to account for noun and adjective agreement.

4.5.4 Extended Projection – Grimshaw (1991)

In this section, I discuss Grimshaw's extended projection theory as a way to provide the necessary configuration that enables agreement between a determiner and a noun to take place. In this theory, the idea of identity of category between a lexical head (such as verbs and nouns) and the functional projections which occur above it is explored. Grimshaw proposes that projections of lexical heads form larger projections of some kind with the functional heads above them and these projections are formed on the basis of identity of category, revealing a similitude among all the elements within the same projection. Based on the idea that, from a semantic viewpoint, the verb selects the thematic information of a sentence and that functional projections above VP are, thus, projections of V, Grimshaw proposes something along the same lines for the nominal domain: namely, the DP would be an extended projection of the noun. As with the verbal domain, the same category features would be assigned to N and the functional head occurring above it. Under extended projection, there would be identity of category between D, N and, more controversially, P. Grimshaw proposes that, once these category features are abstracted away from the lexical/functional distinction, the heads can be considered part of the same syntactic category.

With respect to agreement, extended projection can arguably provide the necessary configuration which would allow gender agreement between a determiner and a noun to take place. If, as claimed, information projects from all of the heads of an extended projection, there should be consistency within a

projection for all projected features, including gender features. Under extended projection, it can be argued that agreement between a determiner and a noun is a full syntactic phenomenon. Both determiners and nouns would be fully specified after agreement takes place. This sort of agreement would be a consequence of how extended projection works.

In the case of agreement between a noun and an adjective, the theory of extended projection does not resolve all the issues. We saw that predicative adjectives are best characterised as adjuncts to NP, either to the left or to the right, depending on the language. Under this configuration, the AP is not part of the same extended projection of the noun, contrary to what happens with the D. A different syntactic dependency would hold between these two elements and, thus, a different agreement configuration needs to be considered. It is discussed next.

4.5.5 Theta Identification – Higginbotham (1985)

To my knowledge, no one has provided a thorough satisfactory account of the agreement configuration underlying the relationship between nouns and adjectives. As we saw earlier in this chapter, Magalhães's proposal for agreement within the DP based on the current notions of minimalist grammar *probe* and *goal* does not seem to account for adjective agreement. In order for the *probe* and *goal* proposal to be advanced, one needs to assume that adjectives are heads and that nouns move to the left in Romance languages so that the correct linear order is achieved. It was argued earlier that such proposal is not able to fully account for feature valuation as Magalhães claims. Grimshaw's extended projection theory does not account for noun and adjective agreement either. Here, I attempt to provide a satisfactory account for the agreement relation between a noun and an adjective, based on Higginbotham (1985). In particular, I develop the idea of agreement between noun and adjective configured as constituents whose theta-roles are identified. Higginbotham's work addresses questions related to theta theory[31]. Importantly, it is assumed that all words, including nouns, have a theta grid. As with other proposals within theta theory, Higginbotham assumes that theta roles have to be associated with their arguments. However, he proposes that there are three distinct ways for theta roles to be discharged: *theta marking*, *theta binding*, and *theta identification*. In

[31] For an overview of general issues involved in theta theory, see Cullicover (1997) and Carnie (2002).

(25), below, I reproduce the description of each type of thematic discharge given by Schmitt (1996):

(25)

a. theta-marking: this is the case in which a predicate V, for example, theta marks its internal argument (the nominal phrase under V');
e.g. 'eat the cake'

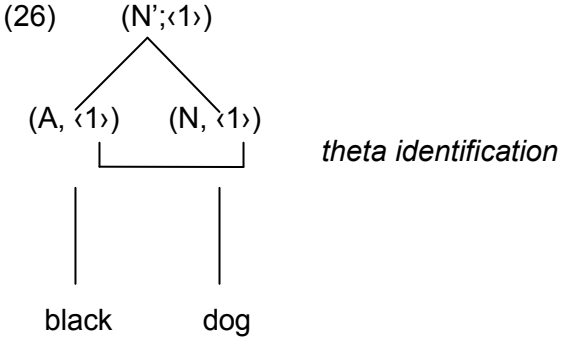
b. theta binding: this is the case in which a determiner or quantifier theta binds an open position in a nominal;
e.g. 'the cake'

c. theta identification: this is the case in which one open position from the adjective and one from the noun merge into a single open position;
e.g. 'red flower'

I will focus on *theta identification*, as it is directly relevant to the type of adjectival modification this thesis is concerned with. For Higginbotham, *theta identification* captures the idea that a phrase modifies the head of a another phrase. Specifically, he develops the concept of *theta identification* on the basis of the idea that modification of one predicative expression by another can sometimes express conjunction. In his words, "a white wall is a thing that is white (on the outside) and a wall" (Higginbotham, 1985: 562). This would not be true for cases such as 'bad violinist', as "it is not a thing that is, on the one hand, bad, and, on the other, a violinist". Following this line of reasoning, Higginbotham makes the case for the treatment of structures of the type in 'white wall' as *theta identification*.

Consider the diagram in (26), below:

(26)



Higginbotham argues that the noun 'dog' has a single open position in the theta grid and the adjective 'black' must have open positions because it occurs as a predicate. The semantic interpretation of the phrase would be something along

the lines of ‘any member which is both of the set of black entities and of the set of dogs’. This would be achieved by considering that the theta-role of the adjective ‘black’ is identified with the theta-role of the noun ‘dog’. So, in other words, semantically, *theta identification* corresponds to the intersection of the set denoted by the noun and the set denoted by the adjective. Next, I explore two possibilities according to which *theta identification* can render the configuration which is needed for agreement between a noun and an adnominal adjective to occur. The discussion that follows is based on an assumption shared by a number of linguistic proposals within distributed morphology, namely that of late vocabulary insertion, i.e., the idea that phonological information is inserted into syntactic structure only after syntax (Beard, 1995; Sproat, 1985). In addition, I will also assume that masculine forms are unmarked forms with respect to gender in Portuguese, thus, the default form. In other words, I will consider that feminine is a feature while masculine forms are characterised by the asbsence of that feature.

A note about the notion of default is important at this stage. It is possible to say that default is an individual notion, in the same sense that grammar is an individual notion, i.e., there is a grammar for speakers but there is no grammar of a language, not in any real sense. On the other hand, however, there are very strong tendencies regarding the notion of default and one of those tendencies is for masculine to be the default gender in languages which have a two gender system. Assuming the notion of default in the above terms possibly means that the vast majority of speakers of a language (perhaps all typical speakers) have masculine as a default.

The notion of elsewhere condition is also relevant here, as it allows us to capture the concept of default. Kiparsky (1973, apud Carstairs-McCarthy (1992)) introduced the elsewhere condition as a principle governing the application of rules. It explores the idea of disjunctive ordering: either one rule applies, or the other, but not both. On the basis of the elsewhere condition, it is possible to account for all the least general cases first and then simply state the most general case. Kiparsky’s (1982, apud Carstairs-McCarthy, op. cit.) description of the elsewhere condition is given in (27) below.

(27) Rules A, B in the same component apply disjunctively to a form Φ if and only if

(i) The structural description of A (the special rule) properly includes the structural description of B (the general rule).

(ii) The result of applying A to Φ is distinct from the result of applying B to Φ.

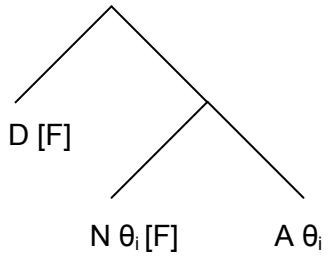
In that case, A is applied first, and, if it takes effect, then B is not applied.

Let us now consider the two alternatives that explore *theta identification* as a potential way to account for noun and adnominal adjective agreement:

1. As a result of *theta identification*, adjectives are NOT specified for gender

Under this alternative, a relationship between noun and adjective via *theta identification* would hold, but no gender feature copying conditioned by *theta identification* would take place. What the *theta identification* configuration would do is provide the context for a spell-out rule according to which, at the level of Vocabulary Insertion, the adjective needs to get its form by looking at the noun which stands in this relation of *theta identification*. The diagram in (28) below illustrates this configuration.

(28) A casa branca (The white house)



where [F] stands for 'feminine' gender feature, the marked gender feature in Portuguese.

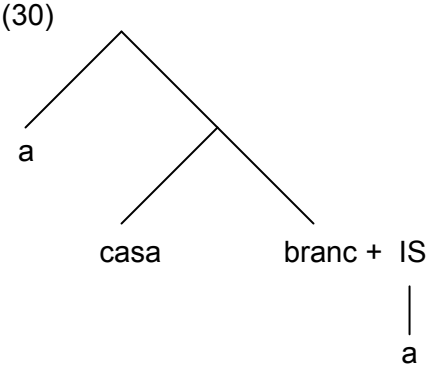
What we see in (28) is that both the determiner and the noun, at the end of the syntactic derivation, are fully specified for gender: the noun has its gender feature as an intrinsic feature and the determiner gets its gender feature from the noun on the basis of either extended projection or the *probe* and *goal* configurations discussed previously. According to this alternative, no feature copying would hold at the morphological level. At vocabulary insertion, a spell-out rule would apply, whereby an adjective would get its form by looking back at the noun. In (29), I illustrate the spell-out rules that would apply at the level of vocabulary insertion under this alternative.

(29) D → o
D[F] → a
N[F] → casa
A → branc + IS
IS → o
IS → a iff N θ_i [F] A θ_i

where 'IS' stands for inflectional slot and 'iff' for 'if and only if'. The notion of inflectional slot used here is intended to make the proposal currently under discussion and the proposal discussed subsequently as easy to compare as possible. One can see it a theory-neutral expression of the regularity that the a/o alternation on adjectives reflects gender. In the proposal under discussion, the gender feature expressed is only present on the noun, so the inflectional slot marks a position for what Emonds (2000) calls 'alternative realization'. An inflectional slot would only be assigned the 'a' form if the adjective accompanies a feminine noun.

The diagram in (30) below illustrates the configuration in question after vocabulary insertion.

(30)



2. As a result of *theta identification*, adjectives are specified for gender

The idea to be explored under this alternative is based on Bobaljik (2008). *Theta identification* itself would hold in a similar way to the alternative sketched above but, in contrast with the previous alternative, a morphological rule of feature copying conditioned by *theta identification* would apply. The same diagram illustrated in (29) above would suit the syntactic level of current alternative but, before vocabulary insertion, a morphological rule, illustrated in (31) would apply.

(31) N θ_i [F] A θ_i → N θ_i [F] A θ_i [F]

Using the same example discussed in the previous alternative, the spell-out rules at vocabulary insertion applicable under this alternative would be the ones

represented in (32) below. In this proposal, the inflectional slot can be seen as marking a position for what is called fission in Distributed Morphology (Halle & Marantz, 1993).

(32) D[F] → a
N[F] → casa
A → branc + IS
IS → o
IS → a iff A[F]

In this case, since a morphological rule of feature copying is assumed, spell-out rules for the adjective at vocabulary insertion are less complex, as it is already specified for gender.

Bringing together the configurations under *probe* and *goal* and extended projection theory discussed above and the current alternative for noun and adjective agreement, it is possible to say that all elements in the DP would be specified for gender before Vocabulary Insertion.

In the next chapter, I will return to the alternatives above and sketch potential ways to incorporate these configurations in a discussion about gender processing in psycholinguistic models.

Chapter 5
LITERATURE REVIEW – The Psycholinguistics of Gender

5.1 What do we know about gender agreement within the DP in SLI?

Not many studies of gender agreement abilities in children with SLI have been carried out. In this section, I will review the work that has been done in Portuguese, Spanish and French.

5.1.1 Portuguese

Silveira (2002) presented a preliminary investigation of gender agreement abilities in Brazilian children with SLI, as part of the development of the language test MABILIN (see chapter 6). MABILIN covers a wide range of structures and its module 2 focuses on morphosyntactic abilities. Four children with the SLI profile were tested in Silveira (op. cit.). For the comprehension tasks, 150 typically developing children ranging from three to seven years old were tested (15 of each age group), while the production task involved 20 typically developing children of five and seven years old (10 children per group). Since tasks were devised to be part of a language abilities test covering a wide range of structures, they are not full experiments, i.e. not many items were included in each condition. Results can, therefore, show only tendencies and their interpretation needs to be cautious. Nevertheless, Silveira's preliminary investigation suggests that gender agreement is an area where children with SLI encounter difficulty. On a picture selection task, exploring if gender redundancy provided by the noun ending and the adjective plays a role in input processing, all children with SLI performed more poorly than the typically developing children. Conditions varied with respect to the presence of a gender morpheme in [+animate] nouns (gender inflected -- e.g. o gato [the_{masc} cat_{masc}] and a gata [the_{fem} cat_{fem}] -- and non-gender inflected -- e.g. o tenista [the_{masc} tennis $player_{masc}$] and a tenista [the_{fem} tennis $player_{fem}$]). The presence/absence of an adjective and presence/absence of a determiner were also varied. No specific pattern with respect to gender redundancy (presence of determiner, adjective or both) was identified, but children with SLI tended to perform poorly when there was no gender morpheme on the noun.

A second picture selection task dealt with the assignment of gender to recently-learned novel nouns. The conditions on this task were the same as those used in the previous task, with the only difference being the type of noun (novel nouns, instead of known nouns). Imaginary characters were introduced on

each trial and children were asked to select the picture that matched the utterance produced by the experimenter. Although there was no specific pattern of performance among the children with SLI, two of the children had considerable difficulties with this task.

A third task dealt with the production of gender agreement in the DP. Children's production was elicited through pictures. The variables manipulated in this task were: animacy of the noun (animate or inanimate); presence of gender morpheme on the animate nouns or typical endings on inanimate nouns (present or absent); and elements participating in the agreement relation (determiner and noun; determiner, noun and adjective). All four children with SLI performed more poorly than the typically developing children. A few errors of non-inflection of the determiner or the adjective were found (in particular in the condition where there was no gender cue conveyed by the noun).

Additional evidence that gender agreement is an area which potentially causes problems for children with SLI is provided by Haeusler (2005). In her study, Haeusler made use of a revised version of MABILIN to identify possible cases of children with SLI for an investigation of argument omission in this population. The data reported here refer to two children who later participated in the current study, WM and FR. A production task on the basis of the assignment of gender to novel nouns was used by Haeusler. A page with three pictures was shown to the child, depicting one imaginary character (for animate nouns) or one imaginary object (for inanimate nouns) and two known images. All three pictures were introduced by the experimenter. A second page was then shown to the child, from which one of the pictures previously present was missing. The child was then asked which picture is missing. Conditions varied with respect to presence/absence of typical ending of the noun. Since all the pictures were introduced by the experimenter with a DP formed by an indefinite determiner and a noun, the expected response was always a DP formed by a definite determiner and a noun, as both the experimenter and the child were familiar with the referents. On this task, both WM and FR produced three mistakes each and all of them consisted of masculine utterances instead of the targeted feminine ones (e.g. 'dabo' when the target was 'a daba').

In addition to the elicited production data above, it is worth noting that Haeusler reports some data from spontaneous speech during therapy sessions.

Like the MABILIN data, WM's and FR's mistakes in spontaneous speech mainly involved producing a masculine item (either a determiner or an adjective) instead of feminine target item (e.g. 'uma porca gordo' [a_{fem} pig_{fem} fat_{masc}], 'um banana' [a_{masc} $banana_{fem}$] and 'bolsa pesado' [bag_{fem} $heavy_{masc}$]. Most of their mistakes consisted in producing an adjective whose gender feature mismatched that of the noun it accompanied.

Although very preliminary and not extensive by any means, the evidence in Silveira (2002) and Haeusler (2005) indicate that gender agreement is an area where children with SLI encounter problems. Further investigations are, thus, warranted.

5.1.2 Spanish

The Spanish gender system is very similar to the system in Portuguese. Nouns are either feminine or masculine and other items in the DP must agree in gender with the noun (with the exception of some adjectives like 'triste' [sad], which are invariant). Like Portuguese, nouns that end in *–a* tend to be feminine and nouns that end in *–o* tend to be masculine. Also as in Portuguese, many nouns do not follow this tendency, and several different noun endings are possible in Spanish.

Bedore and Leonard (2001) were likely the first authors to explore gender agreement in Spanish SLI. In a broad study that investigated a range of what they named grammatical morphology deficits, Spanish-speaking children were recruited in the San Diego metropolitan area. As this was a broad study, which investigated both verbal and nominal morphology, gender errors were not analysed in detail. Nevertheless, the authors report that gender errors were found, such as feminine plural adjectives being substituted by masculine plural forms. Bedore and Leonard recruited children from a questionable setting[32] and do not provide a thorough investigation of gender agreement. Other studies have been carried out since then.

Anderson and Souto (2005) sought to evaluate the pattern of article use by a group of Puerto Rican Spanish speaking-children with SLI. Their ages

[32] The authors report that several steps were taken in an attempt to select children whose difficulties were genuinely characteristic of SLI, such as only recruiting children with proven record of very limited knowledge of English. However, it is undeniable that those children acquiring Spanish in Mexican communities in the United States have, to say the least, a different linguistic experience from those acquiring Spanish in a Spanish-speaking Latin American country or in Spain.

ranged from 4;3 to 5;4, and their performance was compared with age-matched controls. Only their results relating to gender marking will be discussed here. In speech samples obtained through picture description, narrative story telling and play interaction, gender errors accounted for 9.5% of the SLI group's non-target responses, which also included article omissions and number errors. The percentage of gender errors made by the group of age-matched typically developing children was reported by the authors as minimal, but the exact numbers were not provided. With respect to the children with SLI, the authors performed further analysis in order to try and identify patterns of gender errors according to noun ending (typical versus non-typical) and to what was referred to as the noun's semantic transparency (animate versus inanimate). 42% of the gender errors were due to the use of a feminine article when the noun was masculine, whereas 58% of the errors resulted from the use of a masculine article with a feminine noun. 40% of the non-target responses occurred with animate nouns and 60% of the errors referred to inanimate nouns. With respect to noun endings, 70% of the errors were with typical endings, while 30% occurred with nouns with non-typical endings.

In addition, Anderson and Souto (2005) report results obtained in an experimental task in which the production of DPs was elicited. Both the experimenter and the child had a set of cards showing the same pictures. Each card of the child's set contained four different pictures (two different objects, each differing in attributes, for example, color and size). The experimenter's set, on the other hand, contained cards with single pictures. The child's task was to describe the pictures on his/her card, following the order of appearance, so that the experimenter could organize her pictures in the same order. An error analysis showed that 84% (21 out of 25) of the errors produced by the typically developing children consisted of gender errors, while this type of error was present in 21% (28 out of 180) of the non-target forms produced by the children with SLI. At a first glance, these results seem highly surprising, as the task should not have presented any considerable difficulty for the typically developing children. However, a brief analysis of the list of target nouns used in the experimental task reveals major problems and confusion with respect to properties of nouns supposedly controlled for by the authors. The paper contains several instances of misunderstanding of theoretical notions such as feature (in particular the notion of formal feature), agreement, semantics and even gender. For example, the authors make use of the expression 'semantic transparency' in a highly

misleading way, claiming that nouns that are semantically transparent have what they call 'inherent gender'. By doing so, the authors are confounding properties of nouns with properties of the referent denoted by a noun. For instance, in the category *feminine/atypical word ending/transparent*, the authors group together DPs such as 'la mujer' (the_{fem} $woman_{fem}$), 'la bebé' (the_{fem} $baby_{fem\ or\ masc}$) and 'la piloto' (the_{fem} $pilot_{fem\ or\ masc}$). The word 'mujer' refers solely to female referents (the word for 'man' ('hombre') is completely different), whereas the words 'bebé' and 'piloto' can have both female and male referents, without inflecting for gender. Therefore, only the word 'mujer' can be argued to have an intrinsic gender. The words 'bebé' and 'piloto' have no intrinsic gender, as they can be used to denote both female and male referents. In addition, in the category *masculine/typical word ending/transparent*, the authors group together nouns such as 'niño' (boy) and 'toro' (bull), which behave differently as regards their morphological properties. The noun 'niño' can inflect for gender with the addition of a feminine gender marker, becoming 'niña' (girl), while the noun 'toro' does not go through the same morphological process, as the word for 'cow' ('vaca') is completely unrelated. Most likely, such an unbalanced list of nouns had a distorting impact on the results obtained. Because of these problems, the results of Anderson and Souto's study are, unfortunately, not very informative.

Sanchez and Grinstead (2004) also investigated gender agreement in Spanish SLI, but their study was carried out in a different setting, namely, Mexico City. In total, 10 children with SLI (mean age 58 months) participated in the study and their performance was contrasted with that of an age-matched control group and a language-matched control group. Sanchez and Grinstead employed an elicited production task which explored the use of DPs formed by a determiner, noun and adjective. Children were shown pages containing two pictures. The first page contained two items with one characteristic and the second page contained two of the same item with a different characteristic. The experimenter modelled the task on the basis of the first page (e.g. 'Aqui ha una flor roja' ['Here there is a red flower']) and the child's task was to produce an utterance based on the second page of pictures (for the example in question, the page contained a picture of two yellow flowers). The child was, therefore, expected to produce an utterance containing a DP marked for gender and number. The authors report that there was a significant difference between the performances of the children with SLI and the two control groups in this task. However, there is no information detailing the types of mistakes produced by the children. It is, therefore, not

possible to know whether children with SLI produced any errors of gender and, in case they did, their percentage in relation to number errors.

5.1.3 French

Jakubowicz and Roulet carried out a comprehensive study of gender agreement with French-speaking children with SLI (Roulet et al, 2004; Jakubowicz & Roulet, 2007; Roulet-Amiot & Jakubowicz, 2006). Their initial work focused on the agreement between determiner and noun and they later expanded the investigation to agreement between noun and adjective as well.

The initial work on agreement between determiner and noun tested 18 French-speaking children with SLI (mean age 8;9, SD 1.4) and a group of 18 typically developing children (mean age 6;6, SD 0.1). It included an elicited production task and a semantic categorisation/gender perception task. According to the authors, they sought to find out whether or not inconsistent use of grammatical morphemes by children with SLI was the result of a syntactic deficit. Specifically, they wanted to know whether gender errors made by children with SLI in ordinary conversation and in experimental context result from deficits in feature recognition, as would be predicted by the 'feature blindness' account proposed by Gopnik (1990) and the 'missing feature' hypothesis proposed by Clahsen (1989; 1997).

For the production task, 72 pictures were presented one by one to the child, who was asked to answer the question 'What do you see in this picture?'. The expected response in this context is a DP headed by a singular feminine or masculine indefinite article depending on the gender of the target noun. On the perception experiment, Jakubowicz and Roulet used a semantic categorisation task in which children had to decide if what they heard does or does not belong to the semantic category indicated by the experimenter. Two conditions were used: a gender matching condition and a gender mismatching condition. For example, children had to say, by pressing a button as quickly as possible, if 'pantalon' [pair of trousers] and 'cravate' [tie] are items of clothing. Nouns were preceded either by an article matching it in gender or by a mismatching article. The main reasoning underlying this task was the idea that mismatching DPs, if perceived, would cause longer response times (RTs). If incorrect agreement or omission of determiners characteristic of children with SLI were due to a selective impairment in establishing agreement relations, these children would behave

differently from the typically developing ones not only in the production task but also in the perception task. An agreement effect in the perception task would only be present for the typically developing children.

As regards the production task, no single error was produced by the typically developing children. The children with SLI, on the other hand, omitted, on average, 17.5% of determiners and produced, also on average, 6.9% of gender agreement errors. Considering the ease of the task and the fact that the nouns used are usually acquired very early, the number of incorrect responses in the SLI group is quite high. An error analysis showed that agreement errors were relatively more frequent for feminine nouns (child produces 'un' [a_{masc}] instead of 'une' [a_{fem}]) than for masculine nouns (child uses 'une' [a_{fem}] instead of 'un' [a_{mas}]).

Results of the semantic categorisation/gender perception task show that decision latencies were faster in the agreeing condition than in the disagreeing condition, for both the children with SLI and the typically developing children. In addition, the children with SLI were faster than the typically developing children in both conditions. These results suggest that both groups of children were sensitive to gender agreement. An analysis of the categorisation errors that occurred shows that their distribution was not random: the mean number of errors was higher in the incorrect agreement condition than in the correct one for both groups of children and no between group differences were observed. A post-hoc analysis was carried out to determine whether categorisation errors varied in relation to the predictive value of the noun ending. It was observed that, although this factor had no effect, it interacted with the agreement factor. For both groups of children, DPs with nouns whose endings are of low predictive value gave rise to more categorization errors in the incorrect agreement condition than in the correct one. In addition, an individual analysis of the children with SLI did not reveal any specific pattern or relation between what they did in the production task and their sensitivity to mismatching in the semantic categorization task. The authors argue that, with respect to the phenomena considered in the study, children's performance on production does not constitute a reliable reflection of the state of their grammatical competence. They conclude that children with SLI do not suffer from feature-blindness or a break-down of the operation Agree, as Gopnik's and Clahsen's hypotheses claim. The results show, instead, that processing of agreement seems to be irrepressible and automatically calculated even though it is not required by the semantic categorization task. Given these

remarks, the authors suggest that the difference between the two groups of children "does not seem to lie in the properties and modes of functioning of the syntactic component of the language faculty. Rather, the difference seems to reside in the modes of functioning of systems that access the structural representations made available by the syntactic component but are external to this engine, the production system in particular" (Jakubowicz & Roulet, 2007: 26).

Further work from Jakubowicz and Roulet expanded the investigation of gender agreement within the DP to constructions containing a (pre- or post-nominal) adjective (Roulet-Amiot & Jakubowicz, 2006). They wanted to understand whether what they named a 'heavier' DP (noun + det + adj) would increase the number of agreement errors in production and whether a 'heavier' DP would prevent sensitivity to agreement violation from appearing in perception. The authors had as background a hypothesis according to which the difficulties children with SLI have reflect syntactic complexity, calculated on the basis of the number of Merge operations involved in the derivation: DPs containing an adjective have one more Merge than DPs containing a determiner and a noun only. In addition, the authors, citing Giorgi and Longobardi (1991), adopt the view according to which, in a derivation with a post-nominal adjective, the noun has moved over the adjective (cf. page 75). A structure of a DP with a post-nominal adjective would, thus, involve one more Merge operations than one with a pre-nominal adjective, adding complexity to the derivation.

Different groups participated in the study: children with SLI (14 children in total, aged 6;10 to 12;6 years), 4 and 6 year-old typically developing children and a group of adults. Like their previous work on determiner and noun agreement, this study comprised an elicited production task and an input processing task. In the elicited production task, participants were shown, on a computer screen, drawings depicting an animal wearing, holding, standing/sitting on or playing with an object. The experimenter started by providing the participant with a sentential context describing the agent and the event: 'Ici, l'éléphant porte …' ('Here, the elephant is wearing …'). The participant was then required to complete the sentence with a determiner phrase containing an article, a noun and a pre- or post-nominal adjective to finish describing the drawing. Like the previous study, the children with SLI produced significantly more agreement errors than controls in production. Results for the SLI group were the following: errors on determiner only and on both determiner and adjective within the same DP consisted of 5.2%

of total responses while agreement errors on adjective only consisted of 25.7%. It is very clear, therefore, that agreement with the adjective caused many more problems for the children with SLI than agreement with the determiner. With respect to the position of the adjective, very few pre-nominal adjectives were produced, but the authors report that DPs containing a post-nominal adjective triggered relatively more agreement errors than phrases containing a pre-nominal adjective (this was the pattern of all groups).

The second task carried out by Roulet and Jakubowicz was a semantic categorisation task aiming at testing whether participants would be sensitive to violations involving DPs formed by three elements (determiner, noun and adjective), in contrast with two-element DPs tested in the earlier study. Participants were presented with auditory stimuli consisting of concordant and discordant conditions. On the concordant condition, both the determiner and the adjective agreed in gender (e.g. 'une grande cuillère' ['a_{fem} $large_{fem}$ $spoon_{fem}$']; on the discordant condition, either the determiner or the adjective did not agree in gender with the noun (e.g. '*un grande cuillère' ['a_{masc} $large_{fem}$ $spoon_{fem}$'] or 'une *grand cuillère' ['a_{fem} $large_{masc}$ $spoon_{fem}$']. The stimuli also varied with respect to the position of the adjective (pre- or post-nominal). The experimenter provided the participant with a semantic category, some background information, and then asked him/her to press a yes/no 'smiley button' on a touch screen as soon as possible after the presentation of the DP. An example is illustrated below:

> 'Maintenant, le singe va dans un magasin ou l'on vend des vêtements, des choses que l'on met pour s'habiller. 'A ton avis, est-ce que dans ce magasin de vêtements, il va pouvoir acheter …' ('*Now, the monkey is going to a shop where clothes are sold, things we put on to get dressed. According to you, in this clothing shop, will he be able to buy. . .*')

Then, the experimenter presented determiner phrases one after the other for each category: 'une chemise verte' ('a_{fem} $green_{fem}$ $shirt_{fem}$'), 'un *nouvelle balai' ('a_{masc} new_{fem} $broom_{masc}$').

Like the experiment carried out in their previous study, which investigated effects of violation in DPs containing only two elements, overall results of the current experiment show that participants (all groups) were slower and less accurate to categorise DPs in the discordant condition than in the concordant condition. The performance of the children with SLI, however, show different

patterns when compared to the control groups. When the disagreeing element was the determiner, all groups were slower in the discordant condition than they were in the concordant condition. However, when the disagreeing element was the adjective, the children with SLI behaved differently from the other groups and did not react slower in the discordant condition. Such effect, however, was not equal for both types of adjectives: a comparison between pre- and post-nominal adjectives shows that DPs containing a post-nominal adjective did not yield a violation effect. In other words, pre-nominal adjectives generated slower responses in the discordant condition than in the concordant condition but post-nominal adjectives did not. With respect to errors of categorisation, the authors report that only disagreeing determiners yielded higher rates of categorisation errors in the discordant condition than in the concordant one. Disagreeing adjectives, whatever the adjective position in the DP was (pre- or post-nominal) did not create a context for many categorisation errors.

Roulet and Jakubowicz claim that this study provides additional evidence to the findings of their earlier study: agreement is automatic and irrepressible independent of sample population and, despite agreement errors in the production of the children with SLI, those skills needed for input processing seem to function well. With respect to the dissociation observed between determiners and adjectives, the authors raise several possible explanations, such as frequency (determiners, since they are obligatory in French, are more frequent than adjectives) and phonological regularity of determiners in comparison to adjectives. Since errors were proportionally more frequent for adjectives in post-nominal position, Roulet and Jakubowicz argue that the results provide evidence in favor of their hypothesis: adopting the linguistic view according to which DPs containing a post-nominal adjective have the noun move over the adjective for correct ordering, DPs whose adjective follows the noun would be better characterised as ‘more complex DPs’. The number of pre-nominal adjectives produced were, nevertheless, very small compared to the number of post-nominal adjectives, so caution should be used when making generalisations about the two types of adjectives.

5.2 What do we know about gender acquisition in typically developing children?

In this section, I will discuss studies that have investigated how typically developing children acquire the gender system of their native language. Although

research on SLI tends to focus on children who have passed the early stages of language acquisition, it is crucial to look at how young typically developing infants and toddlers in order to have a better understanding about SLI.

The acquisition of gender systems tends to evolve smoothly for typically developing children. It has been reported to occur without problems for children acquiring French (Karmiloff-Smith, 1979), German (MacWhinney, 1978, apud Mills, 1985), Spanish (Pérez-Pereira, 1991), Czech (Henzl, 1975; Polišenská, 2006) and Portuguese (Name, 2002). However, while there seems to be broad consensus regarding the relatively early onset of gender and its smooth acquisition, authors disagree with respect to the mechanisms supposedly responsible for the acquisition of gender in a given language. Most of the studies to date generally assume that grammatical gender is somewhat idiosyncratic and that its acquisition depends on general learning processes, sensitive to frequency, phonological cues and semantic properties. A few recent studies, on the other hand, present robust data suggesting a different perspective, namely one which is based on a syntactic mechanism. In what follows I present an overview of both types of studies investigating gender agreement abilities in infants and children. We will see that the approach according to which gender acquisition depends on general learning mechanisms is subject to well-founded criticism and that a proposal which views gender acquisition as a process dependent on syntactic mechanisms seems to capture the phenomenon in a more refined and more thorough manner.

5.2.1 Karmiloff-Smith's approach

The first type of approach will be illustrated with the work of Karmiloff-Smith (henceforth KS). KS studied gender agreement abilities in monolingual French-speaking children between the ages of 3;2 and 11;10, using a series of elicited production tasks, targeting both known and novel nouns. For KS, language is a 'problem space' approached by children with different strategies. The aim of her study was to identify the cues assumed to participate in the learning of gender, namely cues in the determiner, the noun ending or its semantic properties. The child would see a picture and hear a comment from the experimenter ('Voici l'image de ....' ['Here is the picture of …']. When introducing a second picture, the experimenter would ask the child 'Et ça?' ['And this?']. Variables were controlled to create congruent and incongruent conditions. Children were expected to provide answers containing a definite article and a noun.

General results suggest that children used information on the determiner correctly 82.4% of the time and used information on the noun's ending correctly 86.2% of the time. When cues were present both in the determiner and in the noun ending, the child's task was apparently made easier, with mean score of correct responses reaching 95.9%. The rate of correct responses of children younger than 6 when there was incongruence between determiner and noun ending indicates that no strategy was used predominantly over the other: 47% of the children gave priority to information conveyed by the determiner, while 53% of them preferred the cue provided by the noun ending. KS suggests that these results could indicate that some children prefer one strategy over another, or that the same child adopts more than one strategy.

KS's work has been highly influential and it has inspired further work on the acquisition of gender, such as the study Pérez-Pereira (1991) carried out with children acquiring Spanish. Name (2002), however, offers a different interpretation of the results provided by studies such as KS's and Pérez-Pereira's. According to Name, KS assumes that the strategies children employed in completion of the tasks are part of the acquisition of a given language from the early stages. Name (op. cit.), however, suggests that children's behavior in KS's and similar studies does not reflect the natural way language is acquired. Instead, KS's results reflect the use of procedures mediated by a general cognitive system, employed in order to account for the demands of the experiment. Different strategies would compete when conflicting information is available, which can explain the results obtained in the incongruent condition. The argument against KS's assumption is strengthened by the fact that the children in her study are well advanced in the language acquisition process. Name & Corrêa (2001) argue that phonological and semantic properties are taken into account by children only after nouns are ascribed to gender classes on a syntactic basis. Awareness of these patterns then gives rise to the occasional gender errors that have been reported in the literature on children acquiring Portuguese, such as the data in Figueira (2000, apud Name & Corrêa (2001), who report errors in the speech of children even at the age of six.

5.2.2 Corrêa and Name's approach

An alternative account to the one advanced by KS has been proposed. Corrêa and Name (Corrêa, 2000b; Name & Corrêa, 2001; Name, 2002; Corrêa & Name, 2003) argue that the acquisition of gender, instead of being dependent on

general learning mechanisms, is fundamentally dependent upon syntactic computational operations. Corrêa and Name's work is based on the acquisition patterns of Brazilian Portuguese but it is arguably extendable to, at least, other Romance languages. In contrast with KS's work, Corrêa and Name focus on very young children and explore input processing abilities as well as production abilities. Taking into account that, within the Determiner Phrase, the category of Determiners is the most consistent in terms of phonological regularity, Corrêa and Name's working hypothesis is that children acquiring Portuguese identify morpho-phonological variation related to gender within items in the closed class Determiners. The parsing and the delimitation of morphologically marked gender classes would then "bootstrap" the syntactic operation of the linguistic system as far as agreement with the DP is concerned, enabling the gender of the determiner to be assigned to the noun. This hypothesis assumes early discriminatory abilities and the availability of the functional category Determiner at an early age[33]. Corrêa and Name have carried out a series of experiments and obtained strong evidence for their proposal. Next, I look at their main findings.

Name and Corrêa have investigated early sensitivity to gender agreement between determiner and noun in sentence processing (Name, 2002; Name and Corrêa, 2003). 32 children (mean age 23.2 months) acquiring Portuguese participated in their study. Children were asked to identify a picture in an array of four, after hearing a sentence. Due to the young age range of participating children, which makes it difficult to use a picture selection task in its normal settings, a puppet with synthesized speech provided the target sentence. A target determiner appeared in five different conditions: 1) Gender congruent determiner (GC); 2) Gender incongruent determiner (GI); 3) Inadequate functional item (COMP); 4) Pseudo-functional item (PS); and 5) Random lexical arrangement

[33] Infants' abilities to segment functional categories have been investigated in several studies (Shady, 1996; Shafer et al., 1998). Name (2002) reports data which provide evidence particularly for discriminatory abilities and the availability of the functional category Det at an early age in typically developing children acquiring Brazilian Portuguese. She carried out an experimental task with the preferential head-turn paradigm in an attempt to detect children's sensibility to phonological alterations in the members of the Determiner class. Two versions of four short stories were presented auditorily to children (mean age 15 months): one version contained a story in a condition referred to as 'normal' and the other version contained a story in a condition referred to as 'modified', in which Determiners were systematically replaced by phonologically legal pseudo-Determiners. Results revealed that the listening time on the normal condition was significantly longer than listening time on the modified condition, suggesting that young children are sensitive to Determiners as a class by the beginning of their second year of life.

.

control (RAN). The two conditions manipulating gender agreement are exemplified below (Gender Congruent – CG and Gender Incongruent – GI):

1) GC – Mostre a / aquela / essa bola pro Dedé
   'Show the$_{fem}$ / that$_{fem}$ / this$_{fem}$ ball$_{fem}$ to Dedé
2) GI – Mostre o / aquele / esse bola pro Dedé
   'Show the$_{mas}$ / that$_{mas}$ / this$_{mas}$ ball$_{mas}$ to Dedé

Results showed that, by the age of two, children are able to detect morpho-phonological alterations concerning gender. A higher proportion of correct responses was given for sentences that respected gender agreement between Determiner and Noun than for sentences that violated agreement between those elements. According to the authors, these results provide evidence that is "compatible with the view that young children take into account the information provided by the determiner in the identification of the gender system" (Name & Corrêa, 2001: 6)[34].

Further tests of the hypothesis were carried out in an investigation of gender assignment to inanimate pseudo-nouns by young children. An elicited production task was used by Name (2002), which was later extended in Corrêa and Name (2003), to verify which type of information young children take into account when assigning gender to novel nouns. Thirty young children acquiring Portuguese participated in the study. Children were equally distributed in two age groups: in the younger age group, the age range was 2;2 to 2;10 (mean age 2;7), while in the older age group, the age range was 3;0 to 5;4 (mean age 4;6).

[34] Additional evidence that young children are sensitive to gender agreement between determiner and noun in sentence processing is provided by Lee-Williams and Fernald (2007). Using an eye-tracking procedure, they investigated whether typically developing children acquiring Spanish can use gender-marked articles as an informative cue in interpreting noun phrases. Children were shown a pair of objects as they listened to speech naming one of the objects. On same-gender trial, the nouns depicted by the pictures were either both masculine or both feminine. On different-gender trials, the object names differed in grammatical gender. Their aim was to set up a design which allowed them to evaluate whether or not the gender of the article would be useful in predicting the referent of the subsequent noun. Lee-Williams and Fernald claimed that, if children do use the gender of the article, they would orient to the correct referent more quickly on different-gender trials than on same-gender trials. Twenty six children (mean age 37;7) from Mexican families which had recently immigrated to California took part in the study. Results show that participants identified the referent of a noun more rapidly in the different-gender condition than in the same-gender condition, which indicates that young children are sensitive to gender agreement between determiner and noun in the early stages of language acquisition, even before language production is fully accomplished.

Children were introduced to short stories with imaginary objects on a PowerPoint presentation. Each imaginary object appeared twice, each time in a different colour, previously selected on the basis of the potential of the adjective to inflect for gender[35]. On the first slide of each story block, an object was introduced to the child. Then the same object in a different colour was introduced on the second slide,. On the third slide, both objects appear on a background scene and, on the fourth slide, one of the objects takes part in an event (e.g. falls on the floor or in the water). The child's task was to say what had happened to the object on the fourth slide, following the final question put by the experimenter, which aimed to elicit a referential expression.

The experiment consisted of three conditions, created as a function of a phonology-gender co-relation[36]:

1) *Positively co-related* ⇨ the final vowel of the Noun is the same as the one in the Determiner (-o for masculine and –a for feminine Nouns), like in o mabo and a depa;
2) *Negatively co-related* ⇨ the final vowel of the Noun is opposite to the one in the Determiner (-a for masculine and –o for feminine), such as o bida and a puco;
3) *Neutral* ⇨ nouns with the final vowel –e, which cannot be co-related with gender, such as o mipe and a tobe.

The authors predicted three potential outcomes for the task:

1) Children would make exclusive use of gender information present in determiners when assigning gender to novel nouns;
2) Children would make exclusive use of gender information expressed by the noun ending;
3) Children would make use of both the information conveyed by the determiner and the noun ending.

[35] Not all adjectives referring to colour in Portuguese inflect for gender. 'Vermelho/a' (red) and 'amarelo/a' (yellow), for example, do inflect, while 'azul' (blue) and 'verde' (green) do not.

[36] As we saw in section 3.2, on the characteristics of the Portuguese gender system, the three conditions explored here are valid in Portuguese, i.e., all the combinations can naturally occur in the language.

A main effect of the phonology-gender co-relation was found, indicating that children are sensitive to the phonological pattern Det-N from an early age: both groups of children (younger than three and older than three) scored higher in the positively co-related condition. In addition, a significant interaction between phonology-gender co-relation and age was found. The direction of the means indicates that the sensitivity to the phonological pattern Det-N increases with age: the older group obtained a smaller percentage of correct responses in the negatively co-related responses.

Corrêa and Name (2003) argue that the developmental trend identified in their study supports the hypothesis that "it is necessary for the gender of Nouns to be identified by means of the processing of agreement within the DP in order for a co-relational pattern to be established between the phonological form of the Noun and gender" (p. 20). According to the authors, the 'gender errors' found in children's speech production (such as those consisting of 'regularizations' of forms that are not phonologically positively co-related) are explained as follows: the co-relation between the phonological pattern of the Determiner and the final vowel of the Noun starts to interfere in the processing of gender agreement in production as the children's vocabulary expands. In fact, this type of errors, as well as self-repairs leading to errors, has been identified in natural longitudinal data of two children acquiring Portuguese (Figueira, 1996; 2001). These data show that these errors are occasional and seem to start to occur after the age of 2;3. The data in Figueira reveals different types of behaviour. Figueira reports data of children producing new words, such as 'fado', referring to a male fairy, when the word 'fada' (fairy) is the only one existing in the language, showing that they start exploring the morphological marking of gender in nouns, establishing a relation with the sex of the referent denoted by the noun. In addition, Figueira reports examples of when children perform a phonological harmonization between determiner and noun ending, such as 'um tapo', when the correct form for 'a slap' is 'um tapa'. Figueira emphasises that these types of 'error' do not seem to occur in the early stages of language production, so she interprets the data as evidence for a reorganization of the children's linguistic system at a later stage, i.e. a rearrangement on the basis of the phonological regularities experienced by comparatively older children.

Following Corrêa and Name (2003), which deals with gender assignment to novel inanimate nouns, a new study investigating gender assignment to novel

animate nouns was carried out (Corrêa, 2005b). The main aim of the new study was to verify if the same conclusion held for the study with inanimate nouns. In other words, Corrêa wanted to find out whether children rely on the processing of agreement in the acquisition of the gender of animate novel nouns, as they seem to do with novel inanimate nouns. In the case of animate novel nouns, potentially, the noun's theme vowel can be associated with a semantically interpretable gender inflection and, therefore, impact on the children's performance. A similar task to the one reported in Name (2002) and Corrêa and Name (2003) was designed, with the use of animate imaginary characters instead. As in the experiment with inanimate nouns, two groups of children were tested: a group of children younger than three years old and a group of children older than three years old. Similar conditions were used and a similar procedure was followed. The results revealed that the determiner's gender was maintained by the children in the great majority of test items, suggesting that it is the determiner's gender, rather than the noun final vowel, that provides the most relevant information in the assignment of gender to a novel animate noun. In addition, the results reveal that some responses were characterized by an alteration of the noun ending vowel. This suggests some interference of a correspondence between gender and theme vowel in the production process. These alterations, nevertheless, were significantly more frequent for feminine nouns in the incongruent condition, suggesting that feminine nouns are more vulnerable to congruence effects. Furthermore, group comparisons showed that older children were more vulnerable to congruence effects. Corrêa suggests the effect of correlation between gender and theme vowel is post-syntactic, originating during the morphophonological encoding of the new noun. Older children, given that they are more aware of metalinguistic factors, are more subject to the effects of a correspondence between gender and theme vowel.

In sum, the collective work of Corrêa and Name offers an integrated account of the acquisition of gender in Portuguese, bringing together what is known in different subfields of the Cognitive Sciences to provide a procedural model of the phenomena. They assume that language processing by the child and by the adult functions in a similar way once language and memory capacities are comparable. In addition, Corrêa and Name assume that agreement is a post-lexical process, in which features are 'checked' for compatibility. Their model can be summarised as follows:

> "A phonological phrase is perceived in which a Determiner and a Noun can be segmented on the basis of their phonological and distributional properties (Christophe, 2002; Gout & Christophe, in press). A parsing operation takes place merging D and N in a D-Complement configuration. Given this configuration, the value of the gender feature of the Determiner (say, 0 or 1, corresponding to unmarked and marked forms) and the value of the gender feature of the Noun are matched. If the Noun is not represented in the lexicon, the DP configuration requires that the value of its gender feature be the same as the value of the gender feature of the Determiner." (Corrêa & Name, 2003: 7)

5.2.3 Summary

In this section we looked at the way research in language acquisition has approached grammatical gender in the last decades. Two types of proposals were reviewed: Karmiloff-Smith (1979) studied a group of French children of relatively old age and proposed that gender acquisition depends on general learning processes, sensitive to frequency, phonological cues and semantic properties. Corrêa and Name (Corrêa, 2000b; Name & Corrêa, 2001; Name, 2002; Corrêa & Name, 2003) claim that, given the age of participants represented in KS's work, her results best reflect a strategy used by the children to complete the task, rather than gender acquisition per se. Corrêa and Name put forward an alternative account of gender acquisition, namely one which assumes infants' early discriminatory abilities to segment functional categories and is dependent upon syntactic computational operations.

I share the view of Corrêa and Name with respect to their claim that language processing by (typically developing) children is akin to language processing by adults and that only an integrated approach to language acquisition can capture the phenomenon thoroughly. In an ideal setting, young children who are potential cases of SLI (either because of familial history or because of concerns regarding late speech or any other applicable reason) should participate in studies such as those carried out by Corrêa and Name.

5.3 What do we know about gender processing in adults?

In recent years, gender agreement has attracted the interest of a growing number of researchers in psycholinguistics and neuropsychology. Different aspects of the production and comprehension of gender agreement have been investigated (see Schriefers & Jescheniak, 1999 and Friederici & Jacobsen, 1999 for a

review), ranging from studies with typical adults, using behavioural experimental, electrophysiological (ERP) and spontaneous data, to studies investigating the way in which grammatical gender is represented, used and lost in patients with aphasia. In addition, different models depicting the functional architecture of gender processing have been proposed (Dell, 1986; Levelt, 1989). Studies investigating the psycholinguistics of gender have focused on several aspects involved in its processing, namely accessibility of gender features and nouns in the mental lexicon (Badecker et al, 1995), agreement between different items within the DP (Hagoort & Brown, 1999, Faussart, 2000), agreement between a noun and an adjective in post-copular position (Vigliocco & Franck, 1999). The quote below, from van Berkum (1997), captures well the dynamics of gender processing in real time and sheds some light on the aspects of the phenomena that should be taken into account in experimental investigations:

> "The frequency with which gender must be retrieved from the mental lexicon clearly imposes a considerable real-time demand on a speaker: not only must he or she frequently recover a noun's gender, but this must be done in time, often before the noun itself has been used, and early enough to have the appropriate word forms read for uninterrupted, fluent speaking. Non-native speakers of gender languages will readily appreciate this demand. But most native speakers will hardly be aware that it exists at all; to them, gender agreement usually comes for free. How do these speakers retrieve grammatical gender information from memory as they speak, such that their fluency is preserved?" (van Berkum, 1997: 117)

In this section, I look at how gender agreement processing within the Determiner Phrase has been investigated in the psycholinguistic literature. I focus mainly on evidence based on studies with neurologically healthy adult populations, but I also touch on studies with adults with brain damage.

Models of both speech production and comprehension assume multiple levels of representation and processing for grammatical gender. Therefore, if a child with SLI demonstrates difficulties in, for example, the morphological expression of gender agreement, such difficulties could potentially occur at different stages of speech production. In other words, what is manifested as a mismatch of gender features in the outcome of production can, in principle, be the result of a disruption at any level of processing that is involved in the

phenomenon. Studies of gender agreement processing (and agreement processing in general) are arguably crucial to a better understanding of SLI, as they can help us pinpoint where agreement errors may occur.

5.3.1 The production of grammatical gender

In general, psycholinguists assume that different types of mental processes need to be considered in their accounts of speech production (cf. Levelt, 1989; Dell, 1986; Marx, 1999; Franck et al, 2008 and references therein). Naturally, conceptualisation would take place first, specifying which concepts are to be expressed verbally. The nodes in the conceptual stratum provide the input for the first stage of lexical access. In this stage, the selection of so-called lemmas occurs. Lemmas are defined as nodes containing morphosyntactic properties of words, such as their syntactic category and other morphosyntactic features, such as gender in the case of nouns (Levelt et al, 1999). Structure building then takes place, and constituents are structured hierarchically to express relevant syntactic dependencies. Morphosyntactic representations are then converted into phonological representations that specify prosodic structure and include lexemes, that is, the lexical representations of the phonological form of words. These phonological representations are then converted into phonetic ones which are spelled-out in preparation for articulatory planning and execution.

A number of specific factors concerning grammatical gender needs to be addressed. In most cases, the production of a DP containing a determiner, a noun[37] and an adjective would require the following steps:

1. The retrieval of the gender feature of the noun via lemma retrieval;
2. The gender feature needs to be shared with the determiner and adjective via syntactic processes;
3. The correct phonological forms of the determiner and the adjective need to be selected;
4. The phonetic representation of the whole DP needs to be generated and sent to articulatory planning and execution.

[37] These steps are presumably required in the production of inanimate nouns, which have an intrinsic gender feature, such that gender cannot be determined at conceptual level, prior to lemma selection. This is the type of noun used extensively in the experiments of this thesis, to be reported in chapter 7.

Let us now discuss these steps in more detail, with particular reference to Portuguese. According to the model put forward by Levelt and his colleagues (Levelt, Roelofs & Meyer, 1999), one of the most influential models of language production, information on a noun's gender should become available during lemma access, before the retrieval of the noun's phonological form. Indeed, various studies provide convincing evidence for the relative independence between the representation of gender features and the activation of phonological information about a lexical item. For instance, speakers experiencing 'tip-of-the-tongue' states can describe the gender of a noun even though they cannot retrieve its full phonological form, and this has been attested in both Italian and French (Caramazza & Miozzo, 1997; Ferrand, 2001). Studies of gender retrieval by people with aphasia provide additional evidence that gender features are not stored with the phonological form of a noun. Badecker et al (1995), for example, have reported the case of an Italian anomic patient who, in various naming tasks, showed intact ability to identify the grammatical gender of nouns for which he was unable to provide any indication whatsoever about the phonological or orthographic form.

After the first stages of production, when the message the speaker wants to convey is conceptualised and lemmas are selected, presumably the gender feature on the noun is shared with other elements in the DP. In the case of Portuguese, this sharing occurs with determiners and adjectives (although only variant adjectives overtly mark gender; those ending in –e are invariant (see 4.2)).

Let us now consider how the discussion carried out in 4.5, about the linguistic properties of agreement between the different elements in a DP, could be incorporated into the current discussion on the stages and levels required in a psycholinguistic model of gender production. We saw that there are potentially different proposals that could be used for laying out the necessary configurations for agreement between determiner and noun. The same can be said about noun and adjective. In the case of agreement between determiner and noun, we looked at Magalhães' proposal for DPs in Portuguese, adopting the notions of *probe* and *goal* from recent developments in the Minimalist Program (Magalhães, 2004). We also looked at the possibility of characterising determiner and noun agreement according to Grimshaw's Extended Projection Theory (Grimshaw, 1991). In terms of a model of gender processing, it seems that the crucial factor

to point out is that both these proposals would imply that determiners are fully specified for gender at the end of syntactic encoding. In other words, once syntactic encoding is completed, the phonological form of determiners can be selected without the need for any further syntactic or morphosyntactic process.

A debate in the literature concerns when exactly the selection of the correct form of determiners occurs, since, in some languages, the phonological form of determiners depend on the phonological form of the word following it. Before discussing the aspects of this debate, however, we first need to look at the implications for a model of gender processing of the accounts for noun and adjective agreement proposed in chapter 4. Recall that both the accounts proposed in chapter 4 were based on the idea of *theta identification* between nouns and adjectives, i.e. the theta-role of an adjective (e.g. 'black) is identified with the theta-role of a noun (e.g. 'dog'). We took the idea of *theta identification* into two different directions: one in which *theta identification* would imply full specification of gender for adjectives and another in which *theta identification* would solely imply that the two elements are identified, in preparation for a further agreement process. These two alternatives have different implications for a model of gender production. In the first alternative, the 'weight' of adjective gender production is shared by feature copying processes and the expression of inflectional slots. In the second alternative, there would be no feature copying processes, so the 'weight' of adjective gender agreement is placed on later processes.

As anticipated above, a debate in the literature concerns when the selection of the correct form of determiners occurs. Languages vary with respect to the type of information that is necessary for determiner selection to take place. In Dutch, the form of the article is determined by the syntactic properties of the noun, since, in order to select a determiner such as *het*, it is enough to know that a noun is singular and neuter. In Italian, on the other hand, determiner selection depends on the phonological characteristics of the word that follows it (for example, selection of one of the two singular masculine definite articles *il* and *lo* depends on the phonology of the subsequent word), which means that the phonological form of articles can only be selected once the onset of the first syllable of the subsequent word is available (for a detailed discussion of this debate, see Miozzo & Caramazza, 1999). Portuguese is similar to Dutch in that

determiner selection depends on the retrieval of the noun's gender and number, but not on the phonological form of the noun.

To my knowledge, the debate over determiner production has mainly focused on *when* determiners are selected. However, less attention has been paid to *how* exactly determiners are best represented in the mental lexicon. Different logical possibilities can be considered for Portuguese. I will discuss two of these possibilities: 1. determiners are lexicalised and stored as full forms for both masculine and feminine genders; 2. determiners are stored as a root plus an inflectional slot. These two alternatives are illustrated below with reference to indefinite articles.

(1) INDEF → um
INDEF [FEM] → uma

(2) INDEF → um + IS
IS → ø iff x = INDEF
IS → a iff x = INDEF, FEM

In the case of option 1, it is possible to assume that masculine and feminine versions of determiners are accessed directly, without the need of inflectional processes. It can also be argued that, although multiple forms would be involved, the masculine forms of determiners would act as default forms, given that these are more frequent in Portuguese (and in many other languages). 'Default' could be given different psycholinguistic interpretations, one of which could be in terms of greater activation or lower selection threshold for the masculine forms in comparison with the feminine ones.

Given that the number of determiners in Portuguese and other languages is relatively small and that their frequency of occurrence is very high, the possibility of determiners being lexicalised and stored as full forms seems reasonable. Nevertheless, it is also possible that determiners are not stored as full forms but as described under option 2 above. If this is the case, determiners would need to go through an inflectional process depending on the accompanying noun. At the moment, there is little empirical evidence to decide between these two options.

Unlike determiners, adjectives are open class words. Nevertheless, their final form in many languages is crucially dependent upon gender information supplied by the noun. As with determiners, there are different logical possibilities for the storage of adjectives in the mental lexicon. Unlike determiners, however, there is empirical evidence suggesting that adjectives in Portuguese (and possibly at least in other Romance languages) are not stored as full forms. Corrêa, Almeida and Porto (2004) investigated the representation of Portuguese gender-inflected animate nouns and adjectives in the mental lexicon. As they point out, the inflectional process which animate nouns go through, on the one hand, is strictly lexical, with a feminine affix adding semantic information (e.g. 'menino' – 'boy' → 'menin-*a*' – 'girl' or 'professor' – 'teacher$_{masc}$' → 'professor-a' – 'teacher$_{fem}$'). The inflectional process that adjectives go through, on the other hand, is essentially the morphological expression of agreement (e.g. 'pequeno' or 'pequena' – 'small' depending on which noun it accompanies). Corrêa et al hypothesise that this difference might affect the way nouns and adjectives are represented and accessed. They carried out a series of lexical decision tasks which manipulated *grammatical category* (noun vs adjective), *gender* (feminine vs masculine) and *frequency dominance*[38]. Taken together, results suggest that nouns and adjectives are represented and accessed in different ways. According to the authors, adjectives are not represented as full forms but feminine nouns from feminine dominant (FD) pairs (where the feminine surface form is dominant) are likely to be.

Following the stages of production sketched above, phonetic representation would be generated and sent to articulatory planning for overt speech. The different alternatives proposed in the previous paragraphs will be looked at again in chapter 7 when the experiments conducted in this thesis will be reported and discussed.

[38] "Frequency dominance refers to the relative frequency of the surface forms of an inflected pair. For instance, in a language with singular and plural forms, a pair of number-inflected words is singular-dominant if the singular form is more frequently used than the plural one, and it is plural-dominant, if it is the plural form that occurs more frequently. Frequency of use is a reliable predictor of the speed of the recognition of monomorphemic words. For complex words, surface frequency effects can be taken as evidence for the recognition of complex words as full forms, as predicted by the Full Listing Model of word representation (Butterworth, 1983)" (Corrêa, Almeida and Porto, 2004: 64).

5.3.2 Determiner & Noun agreement vs Noun & Adjective agreement

The discussion presented thus far seems to point towards the idea that processes involved in agreement processing between determiner and noun and between noun and adjective might be distinct. As discussed in chapter 4, different linguistic configurations seem to be involved in agreement between determiner and noun and agreement between noun and adjective. These configurations were incorporated in the discussion of the factors that need to considered when sketching a model of gender production. Experimental evidence from production and comprehension studies seems to support the idea that gender agreement between determiner and noun and between noun and adjective are indeed different phenomena.

Some evidence for a dissociation between gender agreement with determiners and agreement with adjectives comes from a single case study of a Spanish-speaking person with agrammatism (Centeno and Obler,1994, apud Antón-Mendez et al, 2002). The patient and a matched control were asked to describe pictures using a determiner, a noun and an adjective. While the performance of the aphasic patient with respect to number did not vary across the items produced (i.e. a number morpheme was always produced for determiners, nouns and adjectives), her performance with respect to gender differed: she performed better on adjectives than on determiners. These results seem to indicate some independence between what is involved in the processing of gender agreement between determiner-noun and between noun-adjective.

Barber and Carreiras (2005) investigated gender agreement relationships using electrophysiological data (ERPs)[39]. Spanish participants read two types of word pairs: (1) word pairs formed by a determiner and a noun (e.g. el piano [the piano]) and (2) word pairs formed by a noun and an adjective (e.g. faro alto [lighthouse high]). Masculine and feminine nouns were selected as part of the experimental list and gender agreement relationships were manipulated in order to present syntactically congruent and incongruent constructions. Results show that disagreement in word pairs of type (2) produced an N400-type effect[40], while

[39] With the ERP technique, electrophysiological activity is recorded via electrodes placed on the scalp of participants. It is a passive technique, in the sense that presentation of stimuli (visual or auditory) to participants is enough for the electrophysiological activity to be recorded. In other words, while behavioral responses are sometimes part of an ERP testing, they are not essential.

[40] The N400 effect is a component peaking approximately 400ms after the presentation of the stimulus

word pairs of type (1) showed an additional left anterior negativity effect (LAN)[41,42]. However, the effects for condition (2) need to be interpreted with caution: the gender agreement relationship present in a construction formed by a noun and an adjective is arguably different from the gender agreement relationship found in constructions made up of a determiner and a noun, but it is also true that word pairs formed by a noun and an adjective are not grammatical constructions in Spanish, a language in which determiners are obligatory. Therefore, it is not clear whether the ERP effects obtained in condition (2) result from the ungrammaticality residing in the lack of determiners or the ungrammaticality residing in the gender feature violation.

Faussart (2000) looked at gender agreement in the input processing of spoken French by neurologically healthy adults. She tested the effects of gender agreement on a lexical decision task by presenting utterances containing grammatically congruent and incongruent Noun Phrases of different types (determiner + noun and determiner + adjective + noun). The rationale behind her task rests on the assumption that, if there is an agreement relation between a prime and a target item, it is automatically computed by the syntactic processor; when a syntactic violation is detected, it interferes with the lexical decision in progress. The study showed that syntactic violation effects are greater when the prime is a determiner than when it is an adjective. In other words, subjects showed slower lexical decision times after hearing utterances such as (a), below, than after hearing utterances such as (b).

(a) *la studieux collégien
'the$_{fem}$ studious$_{masc}$ schoolboy'
(b) *le studieuse collégien
'the$_{masc}$ studious$_{fem}$ schoolboy'

---

[41] LAN stands for Left Anterior Negativity and is another ERP effect.

[42] Additional data concerning gender agreement violation between article and noun are offered by Hagoort and Brown (1999). These authors investigated the effect of grammatical gender violation in Dutch using the ERP technique. The Dutch gender system has two values: nouns have either common gender or neuter gender. Dutch subjects read sentences in which a definite article and a noun had the same gender and sentences in which gender agreement was violated. The authors report that a very clear cut P600 effect was found for the condition with agreement violation. The P600 is known as a syntax-related ERP which is a positive polarity shift that starts at about 500ms. Thus, a mismatch of gender between an article and a noun in Dutch noun phrases produced an effect which has become associated with syntactic phenomena.

5.3.3 Summary

In this section, we looked at several issues concerning the production of grammatical gender agreement in Determiner Phrases. We discussed the necessary stages that need to be incorporated into a model of gender production and reviewed evidence suggesting that agreement between determiner and noun is a different phenomenon compared to agreement between noun and adjective. This discussion is crucial for a better understanding of the difficulties encountered by children with language impairment, as a mismatch in the outcome of agreement production can, in principle, be triggered by a breakdown in different production processes. In the next chapter, we look at the methodological issues of the behavioural study carried out for this thesis.

Chapter 6

METHODOLOGY

6.1 Introduction

Specific Language Impairment is not a category which is often used by Speech and Language Therapists (SLTs) in Brazil. SLI diagnosis is almost inexistent and the assessment measures currently used in the clinical setting are limited. A survey conducted by Corrêa (2000a) with speech and language therapists in the city of Rio de Janeiro, which included an evaluation of the most common tests used in the clinical setting, demonstrated that most of the tests used by therapists are old translations of tests originally devised in English. Thus, in addition to the problems most language assessment tests in English present, which I discussed in chapter 2, their use by Brazilian SLTs encounters an extra number of striking problems, outlined below:

1) If tests have English as a basis, they do not take into account linguistic phenomena that are absent in English but present in Portuguese, such as those related to nominal inflection;
2) When tests are translated, it is harder to control for factors related to lexical items, such as frequency, age of acquisition, phonological complexity;
3) Cultural differences are usually not considered.

Having acknowledged the picture sketched above, the Psycholinguistic and Language Acquisition Laboratory (LAPAL) at Pontifícia Universidade Católica at Rio de Janeiro initiated a project that aimed to devise a language test entirely conceived for Brazilian children, and grounded in the most recent developments of Linguistics and Psycholinguistics. This test – referred to as MABILIN (Módulos de Avaliação de Habilidades Psicolingüísticas) is in the process of being standardized and comprises a series of different modules. Given the lack of formal diagnosis of SLI in the Brazilian clinical setting, the children who participated in this study had to be recruited on the basis of informal reports by the SLT in charge of their therapy and the administration of module 1 of MABILIN. More than 300 SLTs were contacted in the greater area of Rio de Janeiro.

Module 1 of MABILIN uses a picture selection task and tests processing abilities dependent on syntactic operations. It includes structures such as simple

sentences, passive sentences, relative clauses, the equivalent of wh-questions, sentences with reflexive and full pronouns (see the appendix on page 229 for a list of all test items and examples of pictures of module 1 of MABILIN; for a detailed description and presentation of the test, see Silveira (2002), Haeusler (2005) and Corrêa (2005a)).

6.2 Methodological considerations

Careful thought was put into the issue of identifying children with SLI for the current study. Apart from the WISC and the Ravens, which were used to test the children's non-verbal cognitive abilities, module 1 of MABILIN was used to test their linguistic skills. Module 1 of MABILIN is relatively similar to the TROG, as both tests use a picture selection task. MABILIN, however, deals with problematic issues of the TROG addressed in section 2.5.5. For example, MABILIN does not include structures which seem to involve logical relations such as 'The pencil is not only long but also red' or 'The girl is neither pointing nor running', tested by the TROG. In addition, MABILIN tests relative clauses (RCs) taking into consideration the methodological issues raised by Hamburger and Crain (1982), and widely acknowledged in the literature (cf. Kidd and Bavin, 2002; Adani, in press, and references therein). Contrary to the TROG, the block testing RCs in MABILIN fulfils the felicity conditions which, according to Hamburger and Crain, are necessary for the interpretation of this type of structure. In each picture presented to the child, there are two referents denoted by the noun which is the head of the RC (cf. examples on pages 233 and 234). The function of the RC is, after all, that of restricting the set of potential referents for the NP which serves as the head of the RC. Moreover, MABILIN makes use of test sentences in the past tense, in harmony with Grice's Maxim of Manner (cf. page 44). According to this Maxim, which the TROG violates, the use of the present tense with non progressive aspect is inappropriate in experimental conditions such as the ones at stake.

All in all, it is possible to say that the use of MABILIN is a major improvement in comparison with tests such as the TOLD and the CELF but also more appropriate than the TROG, in spite of the apparent similiarities between the two tests.

More MABILIN modules are under construction and/or standardisation, namely a module focusing on morphosyntactic abilities (gender, number and

person) and a module focusing on argument structure. These additions will certainly contribute to more thorough assessments in future research in Portuguese SLI.

In light of the above, the use of module 1 of MABILIN in conjunction with two tests of non-verbal cogntive abilities (the WISC and the Ravens) was the most adequate diagnostic battery of tests available for use in the present thesis.

6.3 Participants

6.3.1 Children with SLI

In order to recruit children with SLI, approaches were made amongst speech and language therapists' private and public units, university language therapy clinics, mainstream schools with an in-house language therapist and clinical psychology units, mainly in Rio de Janeiro, Brazil (around 300 professionals were contacted). Therapists who agreed to collaborate were asked to select only children with normal hearing and articulation, with Portuguese as their first language, and without a diagnosis of autistic spectrum disorder. A total of six children were identified.

Children undertook an audiometrical test to rule out hearing problems and the non-verbal part of the Brazilian version of the WISC (Wechsler Intelligence Scale for Children) to measure their non-verbal cognitive abilities. Children had to score greater than one standard deviation below the mean (i.e. a standard score greater than 85) on the WISC in order to be considered a potential case of SLI. The Raven's Coloured Progressive Matrices were also used to test the children's non-linguistic cognitive abilities. In addition to the WISC and the Raven's, all children were administered Module 1 of MABILIN.

Two summary tables of the children's age at first testing, sex, social class, type of school attended[43], information on hearing test and scores on WISC, Raven's and MABILIN tests are presented below. Note that, in the column 'Ravens', the first number refers to the individual score of each child with SLI and,

[43] See explanation about the Brazilian educational system on page 49.

for comparison purposes, the number in brackets refers to the mean score of the typically developing children in the same age range as the child with SLI[44].

Table 3: Summary table with information about children with SLI

CODE	AGE	SEX	SOCIAL GROUP	SCHOOL	HEARING TEST
WM	7;03	M	lower income	public	passed
FR	6,01	M	lower income	public	passed
GA	7;06	M	higher income	private	passed
CO	6;4	F	lower income	public	passed
PE	5;5	M	lower income	private	NA
JM	9;1	F	lower income	public	passed

Table 4: Summary table with scores on non-verbal tests and on MABILIN

CODE	WISC standard scores	RAVENS raw scores (mean for child's age)	MABILIN % correct	MABILIN z-scores
WM	108	33 (16.2)	43%	-10.65
FR	106	25 (13.7)	47%	-13.73
GA	98	22 (21.3)	71%	-4.5
CO	80[45]	19 (13.7)	69%	-7.16
PE	NA[46]	17 (20.5)	76%	-3.28
JM	89	23 (20)	74%	-3.84

[44] The scores for the typically developing children are presented here for ease of exposition. They are repeated below with the remaining data about this group, along with standard deviations.

[45] As mentioned in 2.4, CO is a good example of how varied diagnostic measures of SLI can be. She obtained a score of 80 on the WISC, which some researchers consider too low for SLI criteria. CO, on the other hand, obtained a score within normal for her age on the Ravens test.

[46] It was not possible to have PE tested on hearing abilities and the WISC. PE was recruited at a mainstream school I had been visiting to recruit typically developing children for the control group. Although I had asked the school teacher to send consent forms to parents of children without any suspicion of hearing or learning problems, PE's teacher thought it would be interesting to have him included in the study, since she had noticed PE was having difficulties in the classroom, in particular during those activities aimed at preparing students for literacy learning in the next academic year. Testing started with the administration of the Raven's, on which PE scored 17 (relatively low but still within normal for his age: 20.2 (sd 3.5)). Testing continued and it promptly became clear that PE's performance on the language tasks was indeed lower than his age peers. I thus initiated a series of conversations with PE's teacher and the school coordinator, hoping to be able to refer him for professional clinical assessment. The school seemed quite receptive and willing to speak with his family at first, but never took the necessary measures to refer him. Therefore, it was not possible to obtain WISC and audiometrical testing scores for PE, as these required him to be taken to a clinic outside school.

A 2-sided t-test showed that the MABILIN results for each of the six children with SLI are significantly different from those of typically developing children[47]: WM [*t(5)*=7.454, *p*=0.000], FR [*t(5)*=6.84, *p*=0.001], GA [*t(5)*=3.208, *p*=0.024], CO [*t(5)*=3.413, *p*=0.019], PE [*t(5)*=2.414, *p*=0.061], JM [*t(5)*=2.606, *p*=0.0048].

Before describing the control group, it is important to note that, although only six children were included in the SLI group, over 20 other potential cases of SLI were investigated but not included into the study for a variety of reasons. As mentioned above, around 300 speech and language professionals were contacted. Some of these professionals did not respond to my request. Others responded stating that none of their children fit the profile I was looking for. Finally, some speech and language therapists responded stating that they did have children who could potentially fit the profile I was after. I visited several clinics/language units striving to find SLI cases. Most children I tested did extremely well on the MABILIN. Other children were not included for different reasons. I will discuss two cases for illustration purposes. One seven year old girl who was being seen at a university clinic for low income patients performed poorly on MABILIN and, based on the therapist's description, seemed like a potential case of SLI. However, her family stopped taking her to the clinic for therapy before I was able to continue the testing. The therapist insisted the child needed assistance but the family argued they could not take the time off to bring the child for the visits any longer. A different outcome was observed when testing a teenage boy who participated in a single case SLI study in Hermont (2005). I travelled to the State of Belo Horizonte to test this boy. I administered the MABILIN and all the experiments designed for this thesis. The boy performed well in every single task, not hesitating when providing his responses. It could be that his earlier language problems were caused by a delay which was then resolved, since a few years had passed since he was tested for Hermont's study. Given the circumstances just described, it was only possible to include 6 children in the SLI group in the current thesis.

6.3.2 Control group

A control group of typically developing children was recruited for comparison purposes. The children were selected as controls if they had no history of speech

[47] The group average was considered as comparison mean. As will be shown below, none of the variables considered when testing the typically developing children (age; social class) yielded significant results, so data were collapsed into one group.

and language disorder, and no history of hearing impairment. A total of 60 typically developing children were tested, equally divided into two social groups (lower income and higher income), which were, in turn, equally divided into three age groups (five, six and seven years). Recruitment from the two social groups was needed in order to match the profile of the control children as closely as possible to the profile of the children with language impairment. The lower income children were recruited at a public school funded by the Municipality of Rio de Janeiro, whilst higher income children were recruited at schools belonging to the private educational sector of Rio de Janeiro.

In addition to the experiments designed for this thesis, the control children were tested on MABILIN and on the Raven's Coloured Progressive Matrices. Testing took place individually, in a quiet room at the school. Children were given regular breaks. In general, three to four sessions (on different days) per child were needed for all the tasks to be administered. The same order of administration of tasks was attempted with each child, but the time between testing sessions varied depending on the availability of school facilities and the possibility of seeing the child outside normal classroom activities.

Table 5: Summary table with information about low income typically developing children

LOW INCOME CHILDREN			
AGE RANGE	**MEAN AGE (SD)**	**RAVENS (SD)**	**MABILIN (SD)**
5 YEARS	5;7 (2)	18.8 (2.6)	87% (3%)
6 YEARS	6;2 (2.7)	18.6 (3.2)	93% (3.8%)
7 YEARS	7;3 (2.9)	20.8 (3.6)	92% (3.3%)

Table 6: Summary table with information about higher income typically developing children

HIGHER INCOME CHILDREN			
AGE RANGE	**MEAN AGE (SD)**	**RAVENS (SD)**	**MABILIN (SD)**
5 YEARS	5;3 (6.31)	20.2 (3.5)	90% (4.6%)
6 YEARS	6;3 (2)	24 (4.6)	93% (2.9%)
7 YEARS	7;3 (3.2)	25 (5.5)	91% (5.8%)

No significant differences between low income children and high income children were found (independently of different age groups): [[t(59)=-0.334, p=0.740]. The same result holds within each age group: 5 years [t(16)=-1.74, p=0.101], 6 years [t(18)=0.062, p=0.951] and 7 years [t(15)=0.085, p=0.409].

Chapter 7

EXPERIMENTS

The experiments reported in this chapter seek to provide evidence about how gender agreement manifests itself in the language of children with SLI who speak Portuguese in Brazil. In addition to providing data on a language that has been little studied in the field of SLI, the experimental study that follows is discussed under the assumption that our understanding of SLI will only improve if an integrated approach to the disorder is undertaken. Moreover, it is also assumed that difficulties related to the morphological expression of agreement, i.e., what is manifested as a mismatch of gender features in the outcome of production, for example, can potentially be caused by various factors in the course of processing.

The main questions this experimental study sought to answer were the following:

1. At which stage of production does gender processing break down?
    - Are children with SLI able to retrieve the gender of nouns without problems? Or are gender retrieval difficulties the source of gender mismatch in DP production?
    - If gender retrieval is not a problem, what factor(s) cause(s) children with SLI to produce DPs with mismatching gender? Is there a problem in the online processing of agreement? Could it be that children with SLI have difficulties in the encoding of morphophonological information after agreement has taken place?

2. Do children with SLI have more difficulties with nouns that have non-typical endings than nouns with typical endings? In other words, do these children rely on the ending of the noun to produce gender agreement?

3. Do children with SLI have more difficulties with determiner and noun agreement, or with adjectival agreement, or do problems occur equally with both phenomena?

4. What happens when children with SLI encounter a novel noun? Do they behave like typically developing children when assigning gender to a novel noun?

Experiment 1 is a categorisation task that explores children's ability to retrieve the gender of a noun without requiring the production of the noun. Experiment 2 is a grammaticality judgement task in which the effect of a violation of gender between determiners and nouns is tested. Experiments 3 and 4 are elicited production tasks. The first investigates the production of DPs containing a determiner and a noun and the second looks at the production of DPs containing a determiner, a noun and an adjective. Experiment 5 is another elicited production task, but it makes use of novel nouns and it seeks to investigate how children use gender information from the speech signal, as well as the phonological patterns of Portuguese, to assign gender to a novel noun. Experiment 6 is a picture selection task targetting children's abilities to choose a noun solely on the basis of gender information provided by the experimenter.

7.1 EXPERIMENT 1

7.1.1 Introduction

In order to understand what might cause a mismatch in gender agreement between determiner and noun manifested when a sentence is uttered, we need to look at the potentially different stages of production. As seen in chapter 5, many studies provide evidence for the idea that the processing of gender agreement involves different stages. It has been shown that the retrieval of gender features does not necessarily require the retrieval of the phonological form of nouns. A problem in retrieving the gender feature of nouns might impact on the production of correct agreement between determiner and noun.

Experiment 1 is a categorisation task in which children were asked to categorise inanimate nouns into different groups based on their gender. The main rationale for this task was to create a context in which it would be possible for children to demonstrate their knowledge of gender (and gender feature

retrieval in the mental lexicon) without necessarily having to produce any linguistic utterance.

7.1.2 Method

Stimuli

40 nouns grouped into four experimental conditions. All the nouns were controlled for age of acquisition[48] and they all designate concrete objects so that children could easily associate them with the related picture.

1) Masculine gender and typical ending (10 items)
    - e.g. barco (boat$_{masc}$)
2) Masculine gender and non-typical ending (10 items)
    - e.g. jornal (newspaper$_{masc}$)
3) Feminine gender and typical ending (10 items)
    - e.g. banana (banana$_{fem}$)
4) Feminine gender and non-typical ending (10 items)
    - e.g. ponte (bridge$_{fem}$)

Materials

- two baskets
- 32 picture cards for the practice session
- 40 picture cards depicting 40 different test nouns, divided into four groups (see conditions section above)

All the cards showed hand drawn-like images of objects. Care was taken to maintain the same style throughout the cards. Cards were laminated so that children could easily handle them without damaging them.

Procedure

Children were invited to play a game with the experimenter. A total of 72 cards (comprising practice and experimental sessions) was placed on the table. The experimenter showed the cards and baskets to children and said that, in that game, they would need to put cards into two different baskets.

[48] Since there is no extensive database on age of acquisition of lexical items in Portuguese, the Spanish version of the Macarthur-Bates Communicative Development Inventories was used, as Spanish is the closest language to Portuguese.

As the task of categorizing nouns based on their gender is very abstract, a thorough practice session (comprising four stages) was carried out prior to presenting the experimental task. Initially, very concrete examples were used to orientate the child's mind to the task of grouping/categorizing the stimuli into two groups. The first set of cards varied between the categories 'items of clothing' and 'fruit'. The experimenter started the session by saying 'Isso aqui é uma pera. A gente vai colocar esse cartão nessa cesta porque a pera é uma fruta. E a camiseta? É uma fruta? Não, é uma roupa, então a gente vai colocar nessa outra cesta.' ('*This is a pear. So we are going to put this card in this basket because a pear is a fruit. How about this t-shirt? Is it a fruit? No, it is an item of clothing, so we are going to put it in this other basket.'*) Four similar items followed. The second practice session presented pictures of round and square objects and children had to categorise them according to shape. The third stage of the practice session required children to categorise the cards into groups of blue objects and red objects. Children had to succeed on each practice session before moving onto the next one. In the final practice session, the experimenter announced that a language game would start and the child was shown how to group the cards according to the gender of the noun depicted. This was done by the experimenter modeling the first trials, as follows: 'Isso aqui é uma bicicleta. A gente vai colocar nessa cesta porque é *a bicicleta*' *('This is a bicycle. We are going to put it in this basket because it is 'a bicicleta' (the*$_{fem}$ *bicycle*$_{fem}$*)'*) Similar trials followed, until the child showed s(he) had understood the criteria. A total of 12 cards were available in this final practice session.

After the practice session, the actual experiment took place, and the child was asked simply to continue in the same way. The order of presentation of the cards was pseudo-randomized so that not many nouns of the same gender appeared in a sequence. The use of both nouns with typical ending and nouns with non-typical ending rules out the possibility of children performing the task based solely on the phonological properties of the nouns' ending. The experimenter delivered the utterances formed by a DP in a natural way for children of the age range in question. Care was taken to deliver the utterances as clearly as possible but not in a way that compromises the fact that articles are inherently unstressed.

Predictions and possible outcomes

This task sought to answer the question of whether children with SLI are able to retrieve the gender of nouns. The following outcomes can be anticipated:

1. If the difficulties SLI children present with gender are due to a syntactic problem or due to a problem in a later stage of gender processing, and not with the retrieval of the gender feature, children will be able to sort out the nouns depicted on the cards;
2. If the difficulties SLI children present with gender are due to a pre-syntactic problem, it can be argued that they will have difficulties in grouping the nouns depicted on the cards into the baskets appropriately;
3. Due to the demand of metalinguistic abilities on the present task, it can also be argued that children with SLI will not be able to do the task due to difficulties in 'thinking about language'.

7.1.3 Results

This experiment yielded incomplete results. Methodological problems faced during field work made it clear that the metalinguistic abilities demanded by the task made it too difficult for some children, so data collection ceased. Children either understood the criteria required for categorising the nouns, and did the task without any problems (and without making any mistakes), or they did not manage to understand the criteria and grouped the cards randomly (or according to criteria used in the practice session).

In total, four children with SLI and 14 typically developing children were tested. Table 7 below shows their performance. ✓ means the child was able to do the experiment and ✗ means the child did not do the experiment properly.

Table 7: Performance of individual SLI children (referred by their initials) and control children (referred by 'c')

WM	✓	**C1**	✗
FR	✗	**C2**	✗
JM	✗	**C3**	✗
GA	✗	**C4**	✗
		C5	✓
		C6	✓
		C7	✓
		C8	✓
		C9	✓
		C10	✓
		C11	✓
		C12	✓
		C13	✓
		C14	✓

Results were computed on the basis of criteria understanding, and not on the basis of the number of correct responses, as it was previously planned. The table above shows that only one of the children with SLI (WM) was able to do the task. Still, this was only possible after a few unsuccessful trials. The 3 other children with SLI (FR, JM, GA), even after several trials, were not able to understand the criteria required for the task. Four of the typically developing children did not manage to do the task (even after a couple of trials, they did not seem to understand the criteria), while the remaining ten children performed without any problems. The four children who did not understand the criteria (C1 to C4) were five years old at the time of testing and children C5 to C14 were six years old.

7.1.4 Discussion

Originally, Experiment 1 was conceived with the purpose of creating a context for the child to show knowledge of a noun's gender without having to produce any noun, and, thus, without necessarily having to access the noun's phonological form. As children obviously cannot be asked explicitly about the gender of particular nouns — as was done with the adult aphasic patient mentioned in chapter 5 — an indirect way of doing this had to be devised. It turned out, however, that the task contained a higher than expected level of metalinguistic

demand. Bearing this in mind, it is difficult to distinguish between predictions 2 and 3. It can certainly be argued that the difficulties three of the children with SLI and the five year-old typically developing children had with this task were caused by immaturity of metalinguistic abilities. For some reason, these children were not yet able to think about language as well as they use language. Perhaps a requirement for successfully undertaking the task is some knowledge of reading and writing (and, as a consequence, some level of metalinguistic reasoning), something the six year old typically developing children are likely to have.

In addition, as mentioned in the procedure section, a practice section with the use of concrete ways of categorising pictures was carried out before the testing section. The use of this practice phase, however, instead of helping to build the appropriate context for the actual testing, may have had the opposite effect. When noticing some children failed to group the nouns correctly on the basis of their gender, the experimenter asked the child to say why he or she had put a card in a particular basket. In many instances, the child's response suggested that he or she was still grouping the cards based on the more concrete criteria previously used in the testing phases. For example, when grouping a card depicting a bicycle (a_{fem} $bicicleta_{fem}$) some children would say it belonged to a certain basket because it was blue.

Interestingly, however, the age of the children with SLI is either the same (in the case of FR) or higher (in the cases of JM and GA) than the average age of the typically developing children who managed to do the task. So, independently from the methodological problems which obscure findings on gender itself (mainly high load of metalinguistic demands), it seems that three out of the four children with SLI showed a behaviour that was not characteristic of their age peers.

Experiment 2 will look at a different aspect related to gender agreement processing, namely the effect of grammatical gender violation in DPs.

7.2 EXPERIMENT 2

7.2.1 Introduction

As mentioned before, a potential gender mismatch manifested in the production of DPs can be caused by a variety of factors. It is possible that gender difficulties affect only the production of children with SLI. It is also a possibility their

difficulties affect both production and comprehension. According to the model put forward by Corrêa and Name (2003), intact input processing abilities are essential for the acquisition of gender features in Portuguese. The authors argue that it is via agreement that the gender of the determiner is assigned to nouns in the course of acquisition. Difficulties in the input processing of gender agreement within the DP could, thus, affect gender acquisition and, consequently, production.

Assessing the abilities involved in the comprehension of gender agreement is, however, very difficult. A standard picture selection task, for example, would not be appropriate. If we ask a child to select the picture that refers to 'a casa' (the_{fem} $house_{fem}$) in an array of pictures, it is very likely the child's response will be based on his/her lexical knowledge of the item 'casa' and, therefore, knowledge of the agreement relationship between the determiner 'a' and the noun 'casa' would not be tested. Since 'direct' testing of abilities involved in comprehension is virtually impossible, researchers are left with 'indirect' ways of testing these abilities.

The current experiment involved a grammaticality judgment task. Even though grammaticality judgment tasks demand abilities of a metalinguistic nature, requiring, therefore, caution when interpreting results, they can be informative about children's sensitivity towards morphological distinctions if they reveal either correct performance or systematic patterns of errors.

Experiment 2 is a grammaticality judgment task exploring the effects of a gender violation in utterances containing an isolated DP formed by a determiner + a noun. It aims to address the following questions:

1. Are children with SLI sensitive to gender violation?
2. If yes, are they more sensitive to a violation between a determiner and a noun with typical ending (e.g. '*o laranja' [the_{masc} $orange_{fem}$]) than to a violation between a determiner and a noun with non-typical ending (e.g. '*o ponte' [the_{masc} $bridge_{fem}$])?

7.2.2 Method

Children were presented with auditory stimuli via a laptop computer as part of a game that has two dogs (a blue one and a red one) as its main characters.

Children had to judge whether an utterance spoken by one of the dogs was correct or not.

The task used the same nouns used in Experiment 1, combined with determiners to form DPs. Each noun was presented twice, once in the grammatical condition and once in the ungrammatical condition. Half of the nouns were first presented in the grammatical condition while the other half was initially presented in the ungrammatical condition. Otherwise, order of nouns was randomised. In total, 80 utterances (40 grammatical and 40 ungrammatical) were presented to each child.

1) masculine gender in the noun, typical ending and gender matching (10 items)
   e.g. o osso (the_{masc} $bone_{masc}$)
2) masculine gender in the noun, typical ending and gender mismatching (10 items)
   e.g. a garfo (the_{fem} $fork_{masc}$)
3) masculine gender in the noun, non-typical ending and gender matching (10 items)
   e.g. o sol (the_{masc} sun_{masc})
4) masculine gender in the noun, non-typical ending and gender mismatching (10 items)
   e.g. a sorvete (the_{fem} $ice\text{-}cream_{masc}$)
5) feminine gender, typical ending and gender matching (10 items)
   e.g. a mochila (the_{fem} $rucksack_{fem}$)
6) feminine gender, typical ending and gender mismatching (10 items)
   e.g. o laranja (the_{masc} $orange_{fem}$)
7) feminine gender, non-typical ending and gender matching (10 items)
   e.g. a colher (the_{fem} $spoon_{fem}$)
8) feminine gender, non-typical ending and gender mismatching (10 items)
   e.g. o nuvem (the_{masc} $cloud_{fem}$)

Materials

- Compaq nx9010 laptop computer
- 1 microphone

The presentation of the auditory stimuli and the management of the visual devices were conducted by an E-prime software script generated for this experiment. The auditory stimuli were recorded by two female native speakers of Brazilian Portuguese. The entire recording took place in a sound-proof room.

Procedure

Children were invited to play a game. The experimenter told the child to say, by pressing one of two buttons[49], if the dogs said something that "sounds right" or something that "sounds strange". If the child thought the dog said something that "sounds right", the dog got a reward (which varied, according to the dog, between a bone and a bowl of water). If the child thought the dog had said something that "sounds strange", the dog would then produce a sound showing unhappiness. The experiment had three parts – an introduction, a practice phase and an experimental phase. The appearance of each image on the computer screen was controlled by the experimenter, so that the experiment proceeded at the child's pace. The blue dog was always on the left hand side of the screen and the red dog was always on the right hand side. Matching and mismatching utterances were allocated to the two dogs in random order. The experiment was carried out as follows:

On screen – *blue dog sitting and red dog sitting*

Experimenter:

— "Nesse jogo, nós vamos brincar com 2 cachorros – um azul e um vermelho. Esses cachorros estão aprendendo a falar! Cachorros falam? Não, mas esses cachorros estão aprendendo e, como eles estao aprendendo, eles de vez em quando falam de uma maneira esquisita. Você tem que me avisar quando eles falarem esquisito."

(*"This is a game with two dogs – a blue dog and a red dog. These dogs are learning how to speak! Do dogs speak? No, but these dogs are learning and because they are learning, they will speak in a funny way sometimes. You need to tell me when they are speaking in a funny way."*)

On screen – *blue dog standing and red dog sitting*

Experimenter:

[49] The button for "correct utterances" had a sticker with a "smiling face" on it and the button for "incorrect utterances" had a "sad face" on it. Children were instructed on how to use the buttons. In general, children had no problems in sorting the buttons.

— “Quando o cachorro azul estiver falando ele vai estar em pé e vai dizer alguma coisa para você. O cachorro vermelho estará sentado”
(“When the blue dog is talking he will stand up and he will tell you something. The red dog will be sitting.”)

On screen – *red dog standing and blue dog sitting*
Experimenter:
— “A mesma coisa vai acontecer com o cachorro vermelho. Quando o cachorro vermelho estiver falando, ele vai levantar e vai dizer alluma coisa para você. O cachorro azul estará sentado.”
(“The same will happen with the red dog. When the red dog is talking he will stand up and he will tell you something. The blue dog will be sitting.”)

On screen – *blue dog licking a bone and barking and red dog standing*
Experimenter:
— “Quando o cachorro azul falar e você achar que ele falou uma coisa legal, certa, ele vai ganhar um osso e latir. Escute!”
(“When the blue dog speaks and you think that he said something that sounds right, he will be rewarded with a bone and bark! Listen!”)

On screen – *red dog drinking water from a bowl*
Experimenter:
— “Quando o cachorro vermelho falar e você achar que ele falou uma coisa legal, certa, ele vai ganhar um pratinho com água e beber um pouco da água. Escute!”
(“When the red dog speaks and you think that he said something that sounds right, he will be rewarded with a bowl of water and drink from it. Listen!”)

On screen – *red dog and blue dog sitting*
Experimenter:
— “Quando você achar que os cachorros falaram alguma coisa estranha, esquisita, eles vão ficar um pouco chateados e vão reclamar, choramingar.”
(“When you think that what the dogs said sounds strange, they will get a bit upset and they will whine. Listen!”)

On screen – *red dog and blue dog sitting*
Experimenter:

— “Vamos ver como o joguinho funciona?”
(“Let’s see how it works?”)

Predictions and possible outcomes

The current experiment aimed to investigate input processing abilities of the children with SLI via a grammaticality judgement task. The potential outcomes for the task are as follows:

1. If the difficulties children with SLI have with gender are restricted to production, the current task should not pose major problems;
2. If the difficulties children with SLI have with gender also affect their input processing abilities, the task will be problematic for them;
3. Similarly to Experiment 1, it is also possible that the children with SLI will find the task difficult due to its metalinguistic demand.

7.2.3 Results

This experiment yielded incomplete results for technical reasons. The computer which was being used to collect data ceased working while I was in Brazil visiting a school. E-prime, the software needed for the current experiment, was damaged and, due to logistical issues concerning time (school year was coming to an end) and distance, the technician at University College London was unable to solve the problem. Therefore, only little data is available.

Figure 8 below presents the percentage of correct responses. Only four children with SLI and seven typically developing children were tested (mean age 6;78).

Figure 8: Percentage of correct responses in the grammaticality judgement task

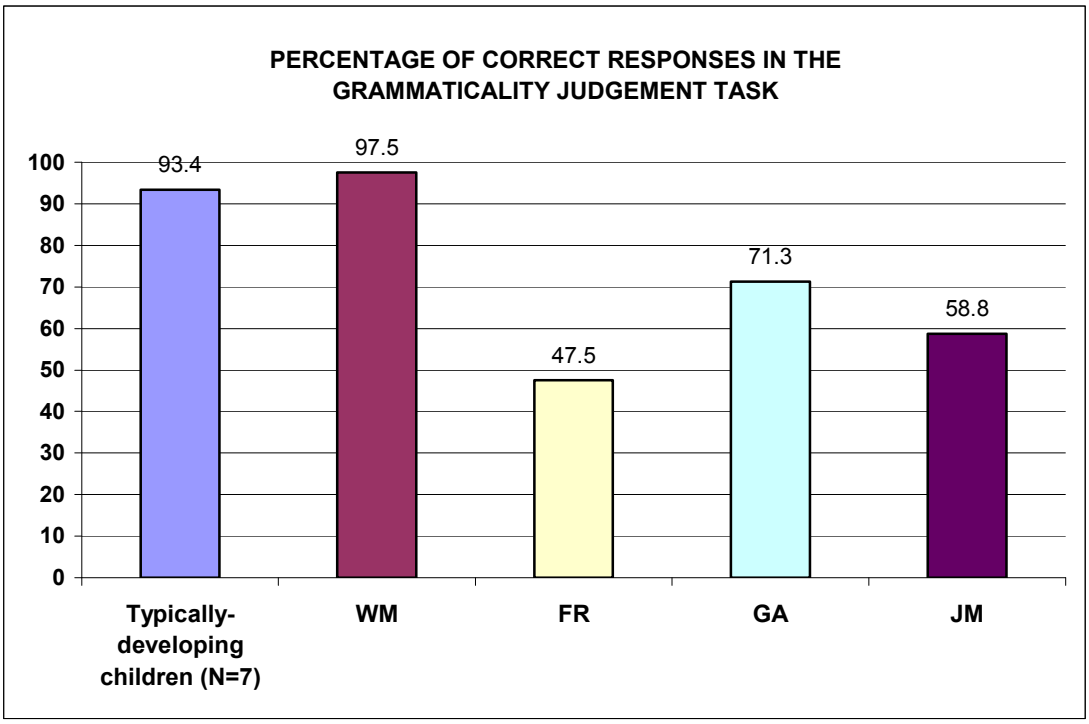


Figure 8 shows that three children with SLI performed more poorly than the typically developing children. One child with SLI, WM, was not successful when he was first administered the experiment, but he then managed to do it without problems. The children FR, JM and GA, even after a couple of trials, continued to perform poorly.

7.2.4 Discussion

This task proved difficult for the children with SLI, with only one child (WM) succeeding.

Interestingly, the three children with SLI who had difficulties with this task also failed in answering questions put by the experimenter after the task was administered. When asked questions like 'When you say '$barco_{masc}$' [$boat_{masc}$], do you say 'o_{masc} $barco_{masc}$' [the_{masc} $boat_{masc}$] or 'a_{fem} $barco_{masc}$' [the $boat_{masc}$]?' In all instances, the children with SLI (with the exception of WM, who succeeded on the task and, thus, was not asked further questions) provided an incomplete answer, producing a bare noun, without any articles, such as 'barco'.

Metalinguistic tasks such as grammaticality judgment involve cognitive demands in addition to linguistic knowledge. Therefore, it is not possible to be

sure if the children with SLI that performed notably less well than the typically developing children fully understood the task. An analysis of the errors produced by these children could potentially shed some light on the issue. If specific response patterns are identified, such as more incorrect answers in the condition with nouns with non-typical endings, this might suggest that children are less sensitive to violations involving a noun with non-typical ending. This, however, was not observed in the data that was collected. In other words, the children did not show a tendency to make more errors in the condition with nouns with non-typical endings.

Note that WM, the child who succeeded on this task, was also the only one who succeeded on Experiment 1. It is also worth noting that WM scored high on both the WISC and the Ravens, the IQ tests used to test the non-verbal skills of the children with SLI who participated in the current study. Although a complete dataset for the current experiment would have provided us with interesting information, WM's performance on the current task, on the IQ tests and on Experiment 1 is indicative that the current task involved a larger load of metalinguistic abilities than ideally needed for a task aiming to assess core language skills. It seems like the task such as the one which was carried out by Jakubowicz and Roulet in their study with French children with SLI (reported in chapter 5) is a better alternative for assessing the input processing of gender within the DP[50].

7.3 EXPERIMENT 3

7.3.1 Introduction

Experiments 1 and 2 explored the retrievement of the gender feature and input processing abilities, respectively. If none of the issues explored in the first two experiments, it is possible that the problem children with SLI have with gender lies in production only. Experiment 3 is an elicited production task that investigates production of gender agreement in isolated DPs formed by a determiner and a noun. This has been investigated in French SLI before, as seen in chapter 5. In Jakubowicz and Roulet's study, children were shown cards with pictures and asked to name what they saw. The French children who participated in the study made 6.9% of errors in gender agreement and omitted around 40% of the determiners in their responses. Since French is a language that does not

[50] Jakubowicz and Roulet's study became available only after data collection for the current thesis was well advanced.

usually allow bare nouns, the children's determiner omissions seem to be a genuine consequence of their problems with DPs. Portuguese, on the other hand, is a language that allows the production of bare nouns in several contexts, such as bare plurals and singular count nouns in argument positions (cf. Schmitt & Munn, 1999). In light of this fact, in order to test the production of DPs by Brazilian children with SLI, a different task from the one reported in Jakubowicz and Roulet had to be designed, namely a task for which the most natural and expected response is a DP formed by a determiner and a noun. This task, reported below, was inspired by the revised version of module 2 of MABILIN (cf. chapter 6).

7.3.2 Method

Children were introduced to three pictures on a computer screen (Compaq nx9010 laptop computer). Each picture was introduced by the experimenter with a DP formed by an indefinite article and a noun. Subsequently, only two pictures remained visible and children were asked to say which picture was missing. By the time the child provided his/her answer, both participants in the speech act (the child and the experimenter) had viewed the picture denoted by the noun. Thus, the most appropriate response was a noun preceded by a definite article.

The same nouns and the same conditions as in Experiment 2 were used here[51]. In order to avoid using an identical set of nouns, 20 filler items were included. These items consisted of pictures targeting plural nouns and therefore required number agreement to be produced by the children. The pictures for the number items appeared as duplicate, in the same frame. The expected response in these cases is a determiner and a noun both marked for plural.

Materials

- Compaq nx9010 laptop computer
- 1 microphone

Procedure

Pictures were presented to the children on a computer using a PowerPoint presentation and animation accordingly. The child was shown three pictures at a

[51] This was mainly due to the lack of feminine nouns without the typical ending "a" that are acquired early and are easily drawn.

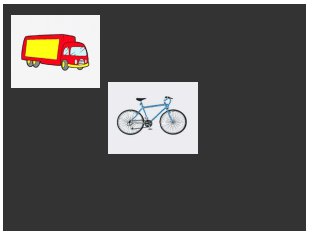
time on the computer screen and was told that one of the pictures would then disappear.

The experiment was carried out as follows:

On screen – picture 1

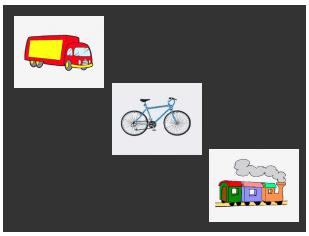
Experimenter: Aqui tem um caminhão (Here there is a_{masc} $truck_{masc}$)

On screen – pictures 1 & 2



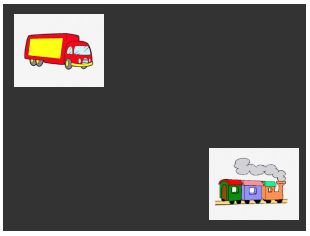
Experimenter: Aqui tem uma bicicleta (Here there is a_{fem} $bicycle_{fem}$)

On screen – pictures 1, 2 & 3



Experimenter: E aqui tem um trem (And here there is a_{masc} $train_{masc}$)

On screen – pictures 1 & 3



Experimenter: O que sumiu? (What has disappeared?)

The appearance of each set of images and the vanishing of the target picture was controlled by the experimenter, so the experiment proceeded at the child's pace.

The task targeted 40 nouns, grouped with the gender properties outlined below:

1) masculine gender and typical ending (10 items)
   e.g. barco (boat$_{masc}$)
2) masculine gender and non-typical ending (10 items)
   e.g. jornal (newspaper$_{masc}$)
3) feminine gender and typical ending (10 items)
   e.g. banana (banana $_{fem}$)
4) feminine gender and non-typical ending (10 items)
   e.g. ponte (bridge$_{fem}$)

Predictions and possible outcomes

This experiment sought to investigate the production of gender in DPs. The following potential outcomes are anticipated:

1. As the task is easy and gender emerges at an early age in typical development (Mills, 1985; Name, 2002), the typically developing children are not expected to have any problems in providing the appropriate answers.

2. With respect to the children with SLI, it is important to note that this task was designed in conjunction with Experiment 1 and the potential outcomes for the current task are more easily understood in the context of Experiment 1. Suppose 1) that Experiment 1 had not presented methodological problems and 2) that the children with SLI had shown that they do know the gender of frequent nouns (by sorting the cards into two different 'gender' baskets). Then, difficulties in the current task would be indication that their problems with gender lie in a later stage of production. In other words, it could be an indication that their problems are not caused by not knowing the gender of nouns (at least not of frequent and early acquired nouns) and not due to difficulties in retrieving the gender of nouns.

7.3.3 Results

Children varied in the type of utterance produced. Although the expected response was a DP formed by a noun preceded by a definite article (given that the two participants of the speech act were supposedly familiar with the noun, as

discussed in the introduction), children did not always produce the article, providing a number of answers with bare nouns.

7.3.3.1 Typically developing children

These children performed at ceiling and did not produce any utterance in which the gender of the determiner mismatched the gender of the noun. In 74% of responses, children provided an answer with an article and a noun, while 26% of the responses contained only a bare noun. As mentioned previously, the most appropriate response in the current task was a noun preceded by a definite article, but an answer containing a bare noun cannot be considered incorrect, as Brazilian Portuguese allows the omission of the determiner in many contexts, unlike other Romance languages.

7.3.3.2 Children with SLI

Like the typically developing children, the children with SLI alternated utterances with a determiner with answers containing a noun only. Their performance with respect to gender marking was high and only a few mistakes (2,5% — six out of 240 items) were made across a large number of items. FR, WM and CO got all the items correct; GA produced two errors: 'a balde' (the_{fem} $bucket_{masc}$) and 'um raquete' (the_{masc} $racket_{fem}$); JM made one error: 'uma tomate' (a_{fem} $tomato_{masc}$); PE produced the largest number of errors, namely three: 'o chave' (the_{masc} key_{fem}), 'o colher' (the_{masc} $spoon_{fem}$) and 'o ponte' (the_{masc} $bridge_{fem}$). Interestingly, WM started three of his answers with the wrong determiner but then corrected himself by re-starting the utterance. For example, WM produced 'um te .. uma televisão' (a_{masc} te …. a_{fem} $television_{fem}$)[52]. WM and PE made one gender error each when presented with filler items: 'dois borracha' (two_{masc} $erasers_{fem}$) and 'dois mala' (two_{mas} $suitcases_{fem}$), respectively[53].

[52] The other items in which he self-corrected were 'dado' ($dice_{mas}$) and 'nuvem' ($cloud_{fem}$).

[53] In these utterances, WM and PE also omit the number morpheme on the noun. Brazilian Portuguese presents dialectal variation in relation to number agreement. Standard Brazilian Portuguese requires number marking in every single determiner, in most nouns (exceptions are cases such as 'ônibus' and 'pires', whose singular forms remain the same in plural contexts) and in adjectives. The dialect of Brazilian Portuguese spoken by working class individuals, however, allows the omission of number morpheme in both nouns and adjectives. Number is only consistently marked in determiner forms.Given that both WM and PE speak the dialect of Brazilian Portuguese that allows the omission of the number morpheme in nouns (and in adjectives), their utterances cannot be considered number marking errors.

7.3.4 Discussion

The main findings of the current experiment were:

1. Typically developing children produced no errors.

2. Contrary to findings in French (Jakubowicz & Roulet, 2007), the children with SLI performed quite well and made only a few errors. Generally, children with SLI did not show any major problems in producing gender agreement between determiner and noun. Nevertheless, although the performance of some of the children with SLI was very high, the occurrence of six errors contrasts with the non-occurrence of errors in the group of typically developing children.

Although the number of errors produced by the children with SLI was very low, it is very clear that there was a tendency for producing masculine determiners instead of feminine targets. In chapter 4, the representation of determiners in the mental lexicon was discussed and two alternatives were proposed: according to alternative 1, determiners would be stored as full forms (masculine and feminine forms independently) and, according to alternative 2, determiners would be represented as a root plus an inflectional slot. In both alternatives, the masculine form of determiners is considered the default option.

It could be argued that the relatively high performance of the children with SLI was triggered by a 'copying procedure': given that the experimenter introduced each picture to the child, producing an utterance in which article and noun matched in gender, children with SLI might have copied what they had just heard, preventing them from making more mistakes. An experiment in which the child has to produce a DP without any sort of modelling beforehand will be able to rule out this possibility and clarify if the copying procedure played any role. Experiment 4 provides a context for production without modelling.

7.4 EXPERIMENT 4

7.4.1 Introduction

In Experiment 1 above, we briefly discussed the idea that gender mismatch can potentially be caused by a variety of factors. It could, as attempted to test in Experiment 1, be caused by difficulties in retrieving the correct gender of nouns. Alternatively, it could be the consequence of problems in establishing agreement

relations between the noun and its accompanying elements. The current experiment was designed to investigate these potential causes.

Like Experiment 3 above, the current experiment tests the elicited production of gender agreement with known nouns. Unlike Experiment 3, which focused on agreement between determiner and noun, it explores agreement relations between three elements: determiner, noun and adjective.

The addition of a third element in the DPs used in the current task allows us to test whether what appears to be a gender mismatch error in production (such as the ones reported in Jakubowicz and Roulet's French study and the few ones reported in Experiment 3 above) results from the selection of the wrong gender feature or from problems in the online processing of agreement between the elements of the DP.

The design of Experiment 4 also gives us the opportunity to investigate potentially two different types of agreement: the literature both within Generative Linguistics and within Psycholinguistc studies, reviewed in chapters 3 and 4 suggests that different phenomena are involved in the processing of agreement between a determiner and a noun and between a noun and an adjective. The task is designed in a way that allows us to contrast the two phenomena.

It is also designed in a way that provides a context in which children can produce an utterance without any modelling by the experimenter, in an attempt to avoid a copying procedure that might have influenced the performance of children in Experiment 3.

The current experiment employed a barrier task to elicit DPs containing an adjective. The child's task was to 'assist' the experimenter in putting her pictures in the same order as the child's pictures. To create a more natural context for the required communication, a barrier was put in between the child and the experimenter, so that each one could see only their own materials. This is a similar task to the one reported in Anderson and Souto (2005)[54].

[54] The two main differences between the current task and Anderson and Souto's task are:
1) Anderson and Souto based their experiment on Lund and Duchan's barrier task (Lund and Duchan, 1993), but decided not to use the barrier between the experimenter and the child because of reported difficulties in maintaining the children's attention. No such difficulties

The specific questions this experiment aimed to address were:

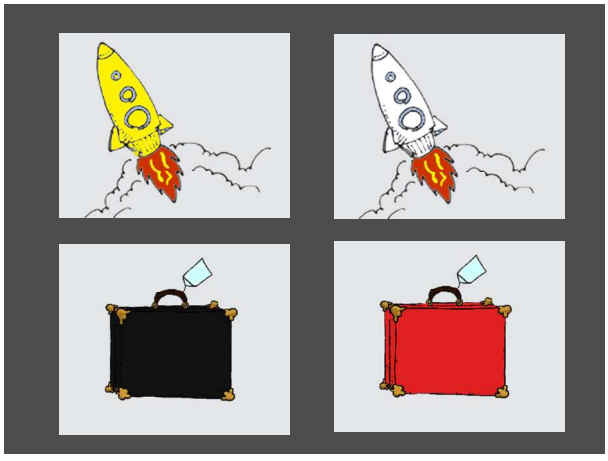
1. Are gender errors more prone to occur in determiners, in adjectives or do they occur equally in both?
2. Do nouns with non-typical endings trigger more errors than nouns with typical endings?
3. Is there a difference in the production of gender agreement on the basis of the gender value of the noun? If there is, is there a tendency for feminine target nouns to cause more difficulties, given that determiners and adjectives accompanying them are marked?

7.4.2 Method

Children were shown pages containing four pictures (two pairs of pictures of objects, where the pictures in the pair represented the same object but with different attributes such as colour or size), while the experimenter had individual cards containing the same pictures as the children but in scrambled order.

Example:

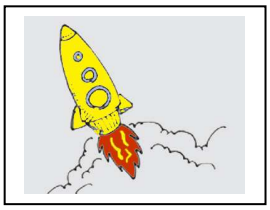
Child's card

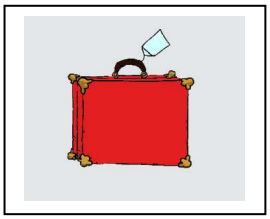


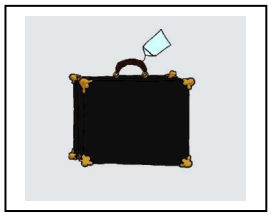
were noticed in the current study, so the barrier was used throughout the administration of the experiment.

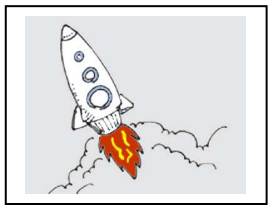
2) Given the crucial problems with Anderson and Souto's selection of lexical items, in particular the misuse of notions such as 'semantic transparency' and 'inherent gender', as reported in chapter 5, the list of lexical items used in this study was not based on the same criteria used in the Spanish study.

Experimenter's cards









The task targeted 20 nouns in four different gender categories:

1) masculine gender, typical ending (5 items x 2)

e.g. o dado branco (the$_{masc}$ dice$_{masc}$ white$_{masc}$)

o dado amarelo (the$_{masc}$ dice$_{masc}$ yellow$_{masc}$)

2) masculine gender, non-typical ending (5 items x 2)

e.g. o foguete amarelo (the$_{masc}$ rocket$_{masc}$ yellow$_{masc}$)

o foguete branco (the$_{masc}$ rocket$_{masc}$ white$_{masc}$)

3) feminine gender, typical ending (5 items x 2)

e.g. a mala vermelha (the$_{fem}$ suitcase$_{fem}$ red$_{fem}$)

a mala preta (the$_{fem}$ suitcase$_{fem}$ black$_{fem}$)

4) feminine gender, non-typical ending (5 items x 2)

e.g. a chave amarela (the$_{fem}$ key$_{fem}$ yellow$_{fem}$)

a chave branca (the$_{fem}$ key$_{fem}$ white$_{fem}$)

Materials

Two sets of cards, containing pictures of the same objects, were used. The child's set contained 10 cards (A4 size) with four pictures each (two pairs of objects, differing in attributes) and the experimenter's set contained 40 cards (10

x 15 cm) with individual pictures. A large paper barrier was used to block the view of the other person's cards.

Procedure

Children were invited to play a game with the experimenter. The set of A4 cards was given to the child, while the experimenter kept the set of small cards. The experimenter asked if the child would be willing to help the experimenter put her cards into the same order as the cards the child had. The instructions reinforced the idea of the child having to "help" the experimenter as the barrier between them prevented the experimenter from seeing the child's cards and vice-versa. The task proceeded as follows:

Experimenter:

— "Nesse jogo, nós vamos brincar com esses cartões. Os seus cartões estão organizados na ordem certa, e os meus estão todos bagunçados! Você pode me ajudar a colocar meus cartões na ordem certa, igual aos seus?"

("*In this game, we are playing with these cards. Your cards are well organised, in the right order, while mine are all messy! Would you help me put my cards in the right order like yours?*")

(give some time for the child to reply)

— "Mas, olha, nesse jogo, tem essa barreira entre a gente e eu nao consigo ver nada do que você esta vendo! Você tem que me dizer direitinho o que você esta vendo para poder me ajudar, OK?"

("*But, look, there is a barrier between us in this game and I cannot see anything that you see! You need to tell me exactly what you are seeing in order to help me, OK?*")

— "Vamos começar?"

("*Let's start?*")

— "O que você esta vendo no seu cartão?"

("*What can you see on your card?*")

— "E agora, o que você está vendo? Me ajuda a colocar os meus cartões na ordem certa!"

("*And now, what can you see? Help me place my cards in the right order!*")

Responses were recorded on a Compaq nx9010 laptop computer, with audio software Audacity.

The barrier in between the child and the experimenter proved to be very efficient: some children would initiate the experiment replying with utterances made only of demonstrative pronouns and pointing to the picture in question, as if the experimenter was able to see what the child was seeing. Whenever that happened, the experimenter reminded the child that she could not see the child's pictures so he/she had to provide complete answers.

Children were expected to produce utterances formed by a determiner + noun + adjective. However, since Brazilian Portuguese allows bare nouns in some positions (cf. page 144), answers without a determiner are also well-formed. Answers were, thus, considered correct as long as there was no gender mismatch between the elements produced, even if a determiner was missing. Answers such as 'mala branca' ('suitcase$_{fem}$ white$_{fem}$') and 'a chave amarela' (the$_{fem}$ key$_{fem}$ yellow$_{fem}$'), for example, were considered correct.

In order to check for reliability of scoring, a native speaker of Brazilian Portuguese independently transcribed the responses of all six children with SLI and six randomly-selected typically developing children (10% of the control group). The level of agreement between the two transcriptions of responses was 97.6%.

Predictions and possible outcomes

This experiment was designed to test the production of DPs formed by a determiner, a noun and an adjective. The following potential outcomes are anticipated:

1. As for Experiment 3, it is expected that the typically developing children will not have any problems in providing the appropriate answers, given that the task is easy and gender emerges at an early age in typical development;
2. The children with SLI, on the other hand, are expected to encounter difficulties with the task. The following outcomes are possible if the task proves difficult for these children:

a. wrong gender in both determiner and adjective in relation to noun
b. wrong gender in determiner only
c. wrong gender in adjective only

If outcome 'a' is reported, it could mean children with SLI have a problem with retrieving the correct gender of nouns but online processing of agreement is intact. If outcome 'b' or 'c' are reported, it could be argued that children with SLI do know the gender of the nouns in question and do not have a problem in retrieving their gender feature.

7.4.3 Results

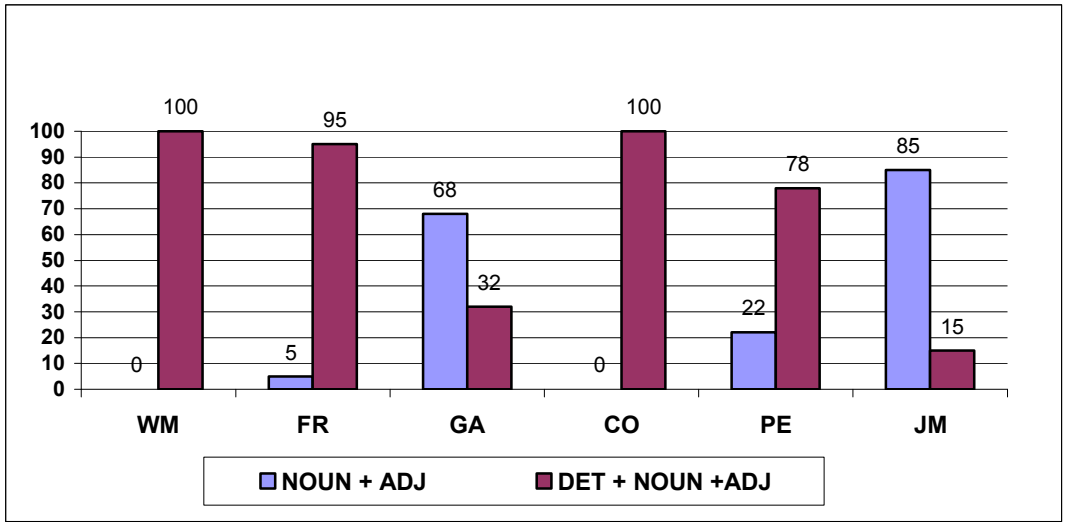
7.4.3.1 Typically developing children

The data from typically developing children will not be presented in detail as they produced only six mistakes out of 2400 test items (60 children x 40 test items). As a whole, 91.6% of responses children provided contained a full DP, while 8.4% were formed by a noun and an adjective only: e.g. 'uma nuvem branca' – 'a_{fem} $cloud_{fem}$ $white_{fem}$' and 'chapéu vermelho' – 'hat_{masc} red_{masc}'. I next look at the performance of the children with SLI.

7.4.3.2 Children with SLI

Like the typically developing children, the children with SLI varied in terms of the type of DP used in their responses. Figure 9 below shows the distribution of responses for each child, followed by a summary with examples.

Figure 9: Breakdown of responses according to DP structure (%)



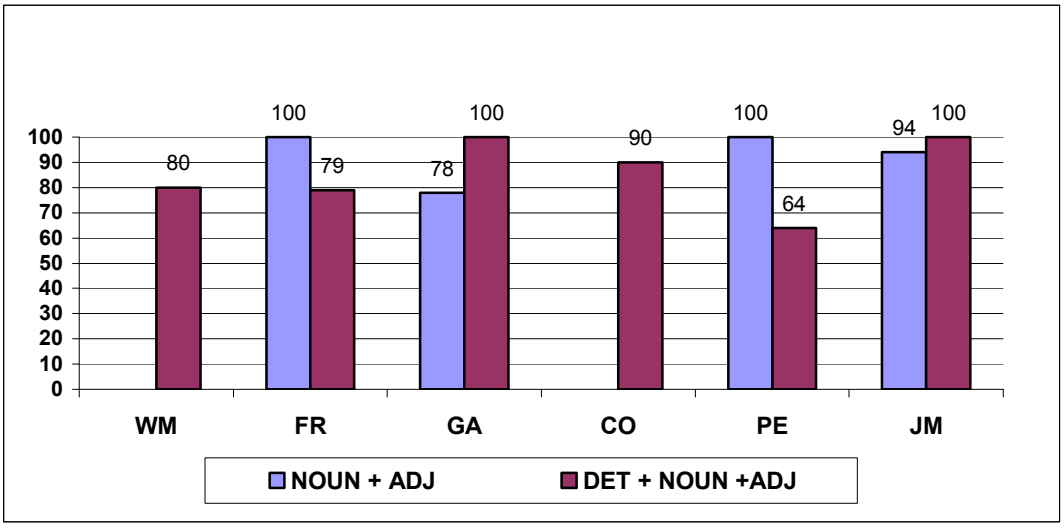
- Four children (FR, GA, PE and JM) alternated their answers containing a full DP with answers that contained only a noun and an adjective.

- Two children (WM and CO) provided all their answers with only one type of DP (det + noun + adj). CO produced around 85% of her answers with a full sentence, instead of an isolated DP. She placed the adjective in copula position, as the following example illustrates: 'um piano é branco' – 'a_{masc} $piano_{masc}$ is $white_{masc}$'). This type of response, with a full sentence, was not given by the other children with SLI, except for a few instances produced by FR (all of which were correct).
- In the answers that contained a determiner, three different types of determiners were observed: definite articles (e.g. 'o telefone preto' – 'the_{masc} $telephone_{masc}$ $black_{masc}$') , indefinite articles (e.g. 'um foguete amarelo' – 'a_{masc} $rocket_{masc}$ $yellow_{masc}$') and demonstrative pronouns (e.g. 'essa colher branca' – '$this_{fem}$ $spoon_{fem}$ $white_{fem}$').

Figure 10 presents the percentage of correct responses according to the DP structure provided by each child with SLI.

It is interesting to note that, for the red columns, the presence of a determiner, which requires gender marking in every possible context in Portuguese and, therefore, could have served as an extra locus for errors, did not trigger more incorrect utterances: the great majority of errors (28 out of 32) with structure determiner + noun + adjective consisted of gender errors in the adjective and not in the determiner.

Figure 10: Percentage of correct responses according to DP structure



Bearing in mind that each participant was presented with 40 test items, raw scores were as follows: WM and FR each made eight errors, GA made five, CO made four, PE made 12, and JM made two errors. A breakdown of the pattern of responses given by the participants with SLI is presented below.

As mentioned previously, all but four of the 32 mistakes produced when a determiner and an adjective were present consisted of utterances in which the adjective mismatched the gender of the noun (e.g. 'uma mala preto' – 'a_{fem} $suitcase_{fem}$ $black_{masc}$'). In other words, the great majority of the incorrect responses these children gave when producing the determiner were incorrect because the gender marking on the adjective was wrong, not the gender marking on the determiner (e.g. 'uma chave amarelo' – 'a_{fem} key_{fem} $yellow_{masc}$' or 'bandeira vermelho' – '$flag_{fem}$ red_{masc}'. The only four responses in which children produced a determiner with the wrong gender were produced by FR and PE: when targeting the masculine noun 'guarda-chuva' (umbrella), which ends with 'a' but is masculine, they produced utterances with both the determiner and the adjective mismatching the noun (e.g. 'uma guarda-chuva preta' – 'an_{fem} $umbrella_{masc}$ $black_{fem}$'). Apart from these four instances with the item 'guarda-chuva', gender marking in the determiner was always correct.

I will now address the issue of gender value, assessing whether or not children performed differently when targeting masculine and feminine nouns. Figure 11 below shows the percentage of correct responses grouped by gender (masculine vs feminine).

Figure 11: Percentage of correct responses according to target gender

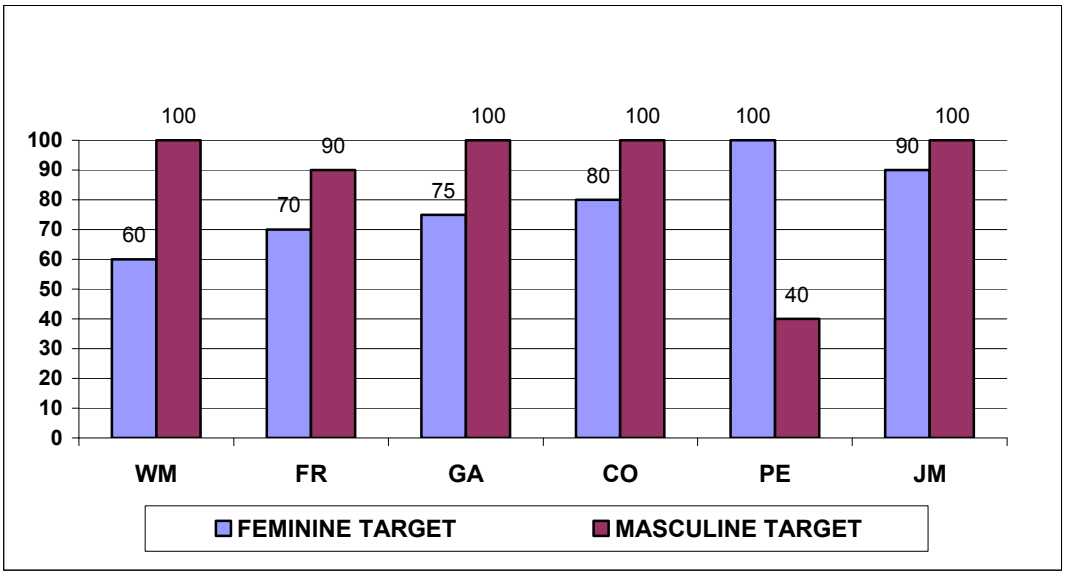
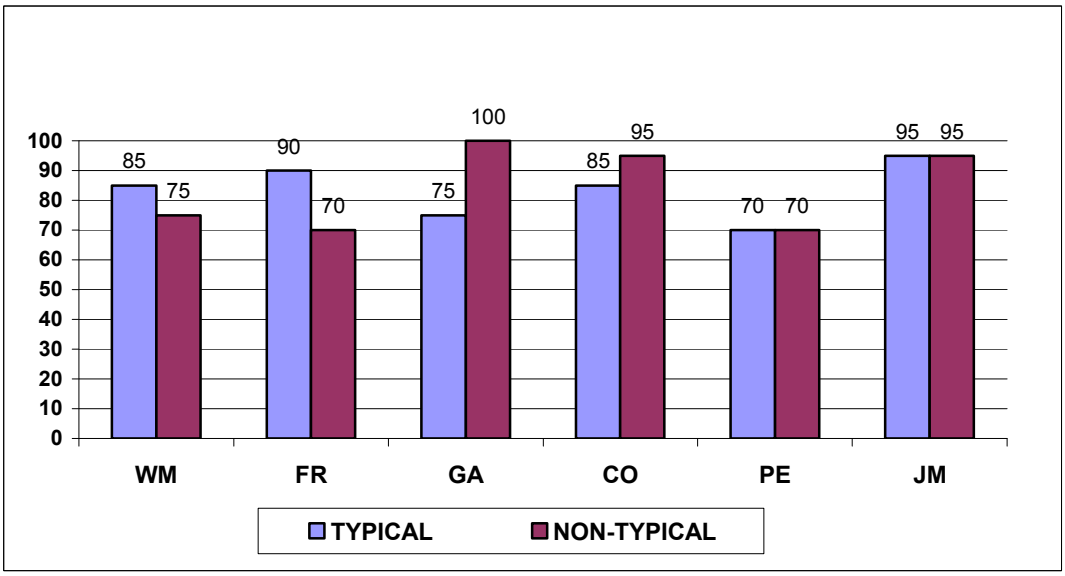


Figure 11 above shows that all children with SLI, except PE, produced a lower percentage of correct responses when the target noun was feminine. Their performance with default unmarked masculine nouns was better. PE, in contrast, performed more poorly with masculine than feminine target nouns. In 60% of the masculine target utterances, he produced a marked feminine form instead of a default unmarked masculine form.

Next I look at the percentage of correct responses grouped by noun ending (typical vs non-typical).

Figure 12: Percentage of correct responses according to type of noun ending



The figure above shows that WM and FR produced a higher percentage of correct responses when the target noun had a typical ending; GA and CO presented the opposite pattern, producing more errors when the noun had a typical ending; PE and JM produced the same number of errors across both conditions. No statistical calculations were carried out due to the low number of test items and participants but it is apparent that differences are negligible.

7.4.4 Discussion

The main findings of this experiment were the following:

1. Typically developing children performed without any problems, producing virtually no incorrect responses (6/2400);

2. The children with SLI, with the exception of PE, performed more poorly when the target noun was feminine than when it was masculine.

3. Most of the errors (28 out of 32, 90%) made by children with SLI consisted of incorrect gender marking on the adjective, not on the determiner. In other words, whenever the determiner was present, it matched the gender of the noun (with only four exceptions to this pattern, all instances with the non-typical noun 'guarda-chuva' (umbrella$_{masc}$);

4. The type of ending did not play a role in the performance of the children with SLI. Any differences were observed between target nouns with typical and non-typical ending were negligible.

Returning to the questions and outcomes outlined in the introduction and prediction sections, it is possible to say the data in Experiment 4 provides convincing evidence against the idea that children with SLI might have a problem in retrieving the gender of known nouns. The data reported here, together with the data in Experiment 3, strongly suggest that children with SLI, at least those of the age range studied here, do know the gender of frequent nouns and do not have a problem in retrieving their gender feature when producing DPs.

Two main questions arise from this pattern of results: what causes children with SLI to encounter more problems with the processing of agreement

between noun and adjective than between determiner and noun? Why is there a strong tendency for default masculine forms to be produced?

The data presented so far suggest that children with SLI do not encounter problems with grammatical gender because of lack of knowledge of the gender feature of particular nouns, or difficulties in retrieving the noun's correct gender feature. In Experiment 3, in which children had to produce DPs with a determiner and a noun only (no adjective) we saw that children with SLI performed almost as well as the typically developing children, though they may have applied a 'copying' strategy, as raised in the discussion. This suggests that these children do know the gender of these nouns. Although the list of nouns used in the current experiment was not the same as the one used in Experiment 3 (some overlap did occur), all the nouns in question are acquired at an early age and are relatively frequent in Portuguese. The fact that children with SLI produced gender correctly in determiners in Experiment 4, where there was no opportunity for any 'copying', do know the gender of frequent nouns and are able to retrieve it during the course of language production. The study carried out by Jakubowicz and Roulet with French speaking children with SLI (discussed extensively in chapter 4) provides additional evidence that children with SLI from the age of 5 or 6 years old onwards (at least those children who are speakers of Romance languages), do know the gender of high frequency inanimate nouns and are able to retrieve it accordingly. In addition, only 10% of the incorrect responses children with SLI produced on Experiment 4 contained a determiner which mismatched the gender of the accompanying noun, which provides further evidence that the children's difficulties are unlikely to be related to problems in accessing the correct gender feature of the noun in the course of language production. If this were the case, they would have produced incorrect responses of the type 'o bandeira vermelho' (the_{masc} $flag_{fem}$ red_{masc}), with both the determiner and the adjective carrying the wrong gender feature. Since this type of error did not occur, we can rule out problems in retrieving the correct gender feature of nouns in Experiment 4.

It seems, therefore, that the children with SLI have all the necessary information to produce DPs, at least up to the level of syntactic encoding. Although there seems to be enough evidence to rule out a few potential explanations for gender problems encountered by children with SLI, discussed above, a number of other alternatives for explaining their difficulties with

adjectives remain. I will focus on three alternatives. They are all based on the configurations discussed in chapters 4 and 5, which I briefly recapitulate below:

- Agreement between determiner and noun is guaranteed either via the probe and goal approach defended in Magalhães (2004) or by the Extended Projection Theory (Grimshaw, 1991). In either proposal, determiners would be fully specified for gender at the end of syntactic encoding;
- Adjective agreement is argued to be based on *theta identification*, which would provide the necessary configuration for this type of agreement. Two possibilities were discussed: 1) *theta identification* would imply full specification of gender for adjectives and the 'weight' of adjective gender production is shared by feature copying processes and the expression of inflectional slots and 2) *theta identification* would solely imply that the two elements are identified, in preparation for a further agreement process and no feature copying processes take place, placing the 'weight' of adjective gender agreement on later processes.
- Two alternatives for the storage of determiners and adjectives in the mental lexicon were discussed: lexicalised and stored as full forms for both masculine and feminine genders or stored as a root plus an inflectional slot.

Alternatives for interpreting Experiment 4:

1. The first alternative is based on the idea according to which gender feature copying would result from *theta identification*. It could be argued that the mistakes children with SLI made with adjective agreement were due to a problem at the level of processing in which the gender feature of the noun is copied to the adjective. Under this alternative, the spell-out rules about resolving the inflectional slot of the adjective are not relevant but would arguably be intact. The configuration outlined under this alternative, nevertheless, has some drawbacks. Although it is a logically possible configuration, it is also somewhat arbitrary and descriptive, and not really explanatory.

2. A second alternative to explain the results of the current experiment focuses on the potential differences between the way determiners and adjectives are represented and accessed in the mental lexicon. In section 5.3.1, we looked

at different ways determiners and adjectives could be represented in the mental lexicon. Two logical possibilities were discussed: 1. items are lexicalised and represented as whole units or 2. items are represented as a stem plus an inflectional slot. If we assume that determiners are lexicalized, given the fact that they are a closed class with few items and that those items are very frequent in speech, and that adjectives are not lexicalised, but possibly represented as a stem plus an inflectional slot that needs to be resolved every time the adjective is produced, a potential explanation can be proposed along the following lines: children with SLI would not have a major problem in producing gender agreement in determiners because there would be no inflectional process involved, i.e., both masculine and feminine forms of determiners would be accessed as single units. Adjectives, on the other hand, would present a relatively more demanding task. Adjectives, in contrast with determiners, would not be lexicalised and would need to go through an inflectional process. Following this line of reasoning, it could be argued that children with SLI have a problem with resolving the inflectional slot of adjectives and at times get it incorrect, with most children resorting to masculine default forms.

3. The third potential explanation I will explore focuses on spell-out rules at the level compatible with vocabulary insertion outlined in chapter 4. Recall that the first alternative discussed for the linguistic configuration of adjective agreement (section 4.5.5) suggested that a relationship between noun and adjective via *theta identification* would hold but no gender feature copying conditioned by TI would arise. Instead, *theta identification* would provide the adjective with a type of 'non-local' spell-out rule by means of which the adjective would get its form. Specifically, this interface rule would read as instructions for the adjective to get its form from information on another item, namely information about the noun which stands in a relation of *theta identification*. This non-local configuration guiding the expression of gender agreement on adjectives would be different from the way gender agreement is expressed on determiners. In the case of determiners, at the level compatible with vocabulary insertion, there is no need to get any information from other elements, as determiners are arguably fully specified for gender at the end of syntactic encoding (see 5.3.1). If the current alternative is pursued, the expression of agreement on adjectives would follow a more complex set of processes in comparison with the expression of agreement on determiners. It could be argued that these potential differences between the processes involved in the expression of agreement in determiners

and in adjectives might be responsible for the distinct patterns observed in the performance of the children with SLI.

It is not clear, however, how this configuration would account for invariant adjectives present in Romance languages (i.e. those adjectives ending in 'e' in Portuguese). In a lexical decision study with French adults, Jakubowicz and Faussart (1998) tested the effects of gender agreement violations in DPs containing invariant adjectives. Two experiments were carried out. In the first experiment, a condition in which the target noun was directly preceded by a gender-marked determiner was contrasted with a condition in which an invariable adjective appeared between the determiner and the noun. Their second experiment compared the effects of gender marking in the latter condition of the first experiment with a condition in which both the determiner and the adjective carried an overt gender mark. The authors found a stronger effect when the target noun was preceded by a determiner plus an invariable adjective than by a determiner only. Moreover. the magnitude of the effect did not vary according to whether or not the adjective carried a phonetically realised mark for gender. Jakubowicz and Faussart interpreted these results as evidence for the automaticity of gender agreement processing, i.e. even invariant adjectives triggered agreement processes.

At the moment, there is not enough evidence to choose between the three alternatives outlined above. Experiment 5 below explores elicited production as well, but makes use of novel nouns, which gives the task a different focus (i.e. assignment of gender rather than retrieval of stored information).

7.5 EXPERIMENT 5

7.5.1 Introduction

As seen in chapter 4, typically developing children acquire the gender system of their native language at an early age. Chapter 4 also reported consistent data suggesting that it is the gender information conveyed by the Determiner that assigns gender to a novel noun, at least in the case of Portuguese (Name, 2002; Corrêa & Name, 2003). These authors claim that phonological information does not play such an important role during the first stages of gender acquisition but can have some influence on the performance of older children, who have already acquired a large number of lexical items and have, therefore, been able to

identify the relatively frequent similarities between the phonological form of the determiner and the noun ending.

In order to investigate how children with SLI assign gender to novel nouns, a task based on the work of Name and Corrêa was carried out. The main question this experiment addressed was which type of information is used more consistently by children with SLI when assigning gender to a novel noun. Is it gender information in determiners or adjectives? Is it phonological information in the noun ending? The current experiment elicits the production of referential expressions requiring gender agreement with novel animate and inanimate nouns[55]. It is important to stress, nevertheless, that caution is needed when interpreting the results given the age of the children who took part in this study. This issue will be discussed in detail in section 7.5.3.2.1.

7.5.2 Method

Children followed stories with imaginary objects and creatures named by pseudo novel nouns. Noun endings were manipulated in order to investigate the extent to which the processing of gender agreement with novel nouns was affected by the phonological form of the novel noun in terms of whether it was concordant, conflicting or neutral with respect to the gender of the determiner.

Test items were 24 pseudo-nouns, grouped into the conditions below, and 18 distractor (known) nouns.

1a) *masculine concordant*: masculine determiner, novel noun ending in –o (4 items)

e.g. um/o puco (a_{masc}/the_{masc} puco)

1b) *feminine concordant*: feminine determiner, novel noun ending in –a (4 items)

e.g. uma/a poca (a_{fem}/the_{fem} poca)

2a) *masculine conflicting*: masculine determiner, novel noun ending in –a (4 items)

e.g. um/o bida (a_{masc}/the_{masc} bida)

[55] While the version used here was heavily inspired by the work of Name and Correa, some changes have been made: several test items were added to the [-animate] list and the materials for the [+animate] test items were different.

2b) *feminine conflicting*: feminine determiner, novel noun ending in –o (4 items)

e.g. uma/a tuco (a_{fem}/the_{fem} tuco)

3a) *masculine neutral*: masculine determiner, novel noun ending in –e (4 items)

e.g. um/o tobe (a_{masc}/the_{masc} tobe)

3b) *feminine neutral*: feminine determiner, novel noun ending in –e (4 items)

e.g. uma/a dobe (a_{fem}/the_{fem} dobe)

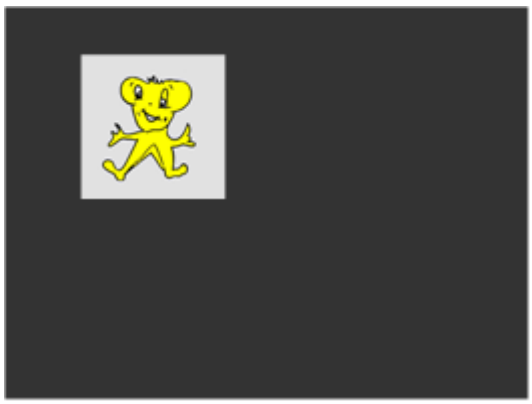
Materials

- Compaq nx9010 laptop computer
- 1 microphone

Procedure

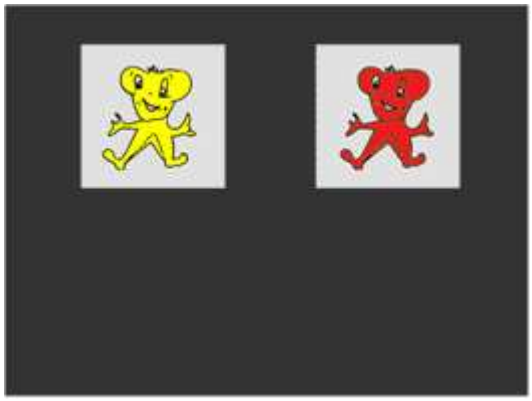
Children were invited to play a game with the experimenter. Slides with story images were presented to the child one by one, as the story was being told. Children had to answer a question relating to the action depicted in the final slide. An example of a mini-story is provided below:

On screen:



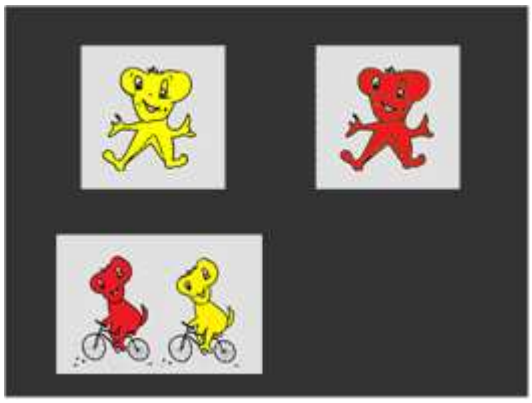
Experimenter: Aqui tem uma bilo (Here there is a_{fem} bilo)

On screen:



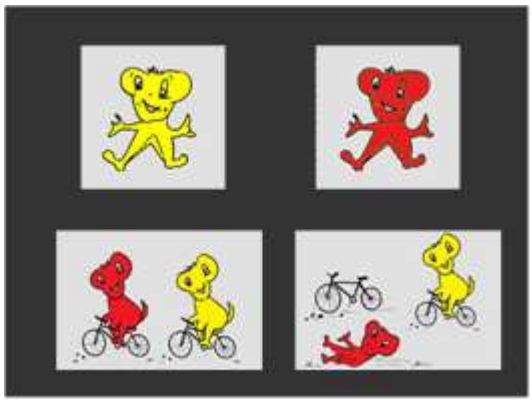
Experimenter: Aqui tem outra bilo (Here there is another$_{fem}$ bilo)

On screen:



Experimenter: As bilos estão andando de bicicleta (The$_{fem}$ bilos are riding a bike)

On screen:



Experimenter: Uma bilo caiu no chão! (A$_{fem}$ bilo fell off the bike!)
Que bilo caiu no chão? (Which bilo fell off the bike?)

Responses were recorded on a Compaq nx9010 laptop computer, with audio software Audacity. Occasionally, children would just point to the imaginary creature who was performing the action (in the case of [+animate] nouns) or the imaginary object upon which the action was carried out (in the case of [-animate] nouns), without providing any oral response. Whenever this happened, the experimenter would ask the child to answer again by speaking instead of pointing.

In terms of scoring, any response containing an item expressing gender was included in the analysis. In other words, children did not have to provide an answer with a full DP (determiner + noun + adjective), as responses containing just a determiner and an adjective or a demonstrative pronoun would show whether they had grasped the correct gender of the novel noun. Therefore, answers such as "o vermelho" (the_{masc} red_{masc}) or "essa" ($this_{fem}$) were accepted as correct, as long as the gender of the items they produced was marked correctly.

As in Experiment 4, a native speaker of Brazilian Portuguese independently transcribed the responses of all six children with SLI and six randomly-selected typically developing children (10% of the control group). The level of agreement between the two transcriptions of responses was 98.5%.

Predictions and possible outcomes

The current experiment sought to answer the question of whether children with SLI are able to use information conveyed by determiners to assign gender to a novel noun. The following outcomes are anticipated:

1. Based on the studies of Name and Corrêa, it was predicted that the typically developing children would have no problems in accomplishing this task, with the exception of the conflicting condition. In this condition, given the relatively advanced age of the children, it was expected that a correlation between the phonological form of the determiner and the final vowel of the noun would interfere in the processing of gender agreement. As previously seen in chapter 5, it is thought that this correlation starts to interfere with the processing of gender agreement in production as children's vocabulary expands and they start exploring the morphological marking of gender in nouns;

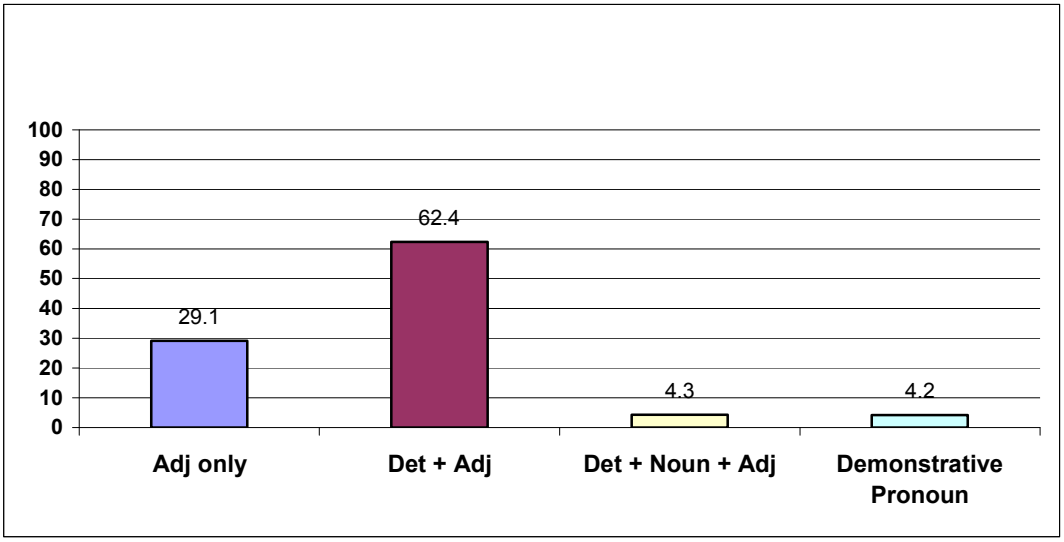
2. In the case of children with language impairment, if the way they assign gender to a novel noun follows the same pattern observed for the typically developing children, it was expected that their performance in this task will be similar. If, however, the children with SLI do not use the gender cues in the determiner in the same way the typically developing children do, they will likely make mistakes across conditions, and not just in the condition where there is a conflict between the phonological form of the determiner and the final vowel of the noun.

7.5.3 Results

7.5.3.1 Typically developing (TD) children

Typically developing children varied in terms of the type of response they provided. Figure 13 below presents the distribution of responses across the four types they produced.

Figure 13: Percentage of responses in terms of DP structure for TD children



This figure shows that in most responses, children produced an utterance formed by a determiner and an adjective (e.g. 'o amarelo' – 'the$_{masc}$ yellow'), but utterances with an adjective only were also relatively frequent (e.g. 'branco' – white$_{masc}$). The other two types of responses were rare. In other words, children did not generally reproduce the recently-learned noun in their response (i.e. answers of the type 'o paco amarelo' – 'the$_{masc}$ paco yellow$_{masc}$' were given rarely). Responses containing a demonstrative pronoun only (e.g. esse – 'this$_{masc}$') were also rare. It is worth noting that these children produced responses with adjectives that show gender marking in over 95% of response items.

The overall performance of the two groups of typically developing children was compared and it was observed that the lower-income group scored 92.5% correct while the higher-income group did so in 94.4% of the occurrences.

As this difference between the two control groups is not significant [*t(46)*=1.217, *p*=0.23], they were treated as a single group in further analyses. Figure 14 below presents the percentage of correct responses according to the value of the gender introduced by the determiner (masculine vs feminine) and figure 15 shows the percentage of correct responses according to noun ending (concordant vs conflicting vs neutral)

Figure 14: Percentage of correct responses in terms of gender for TD children

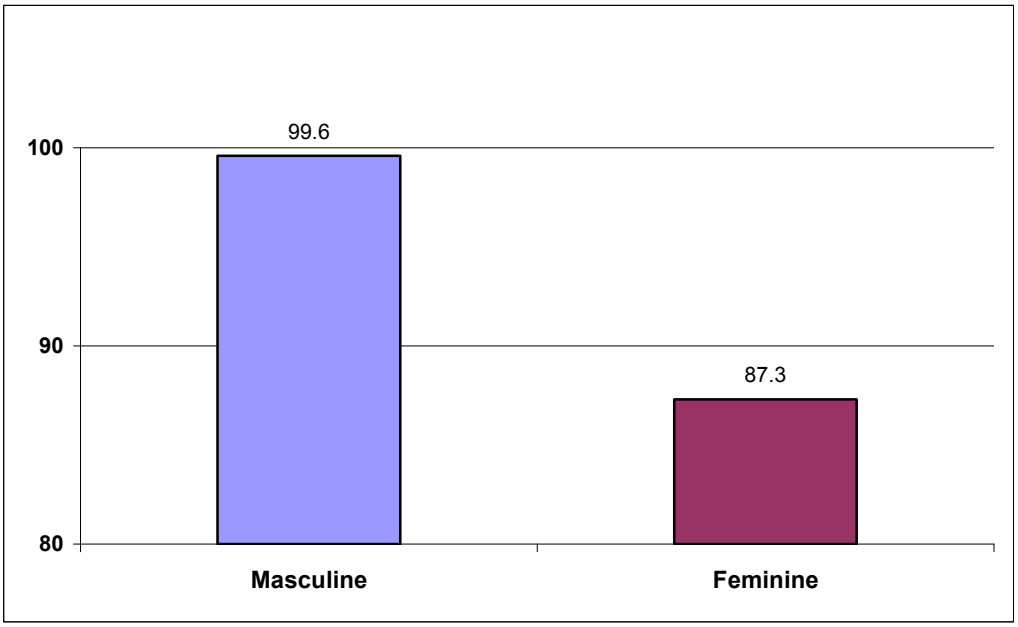
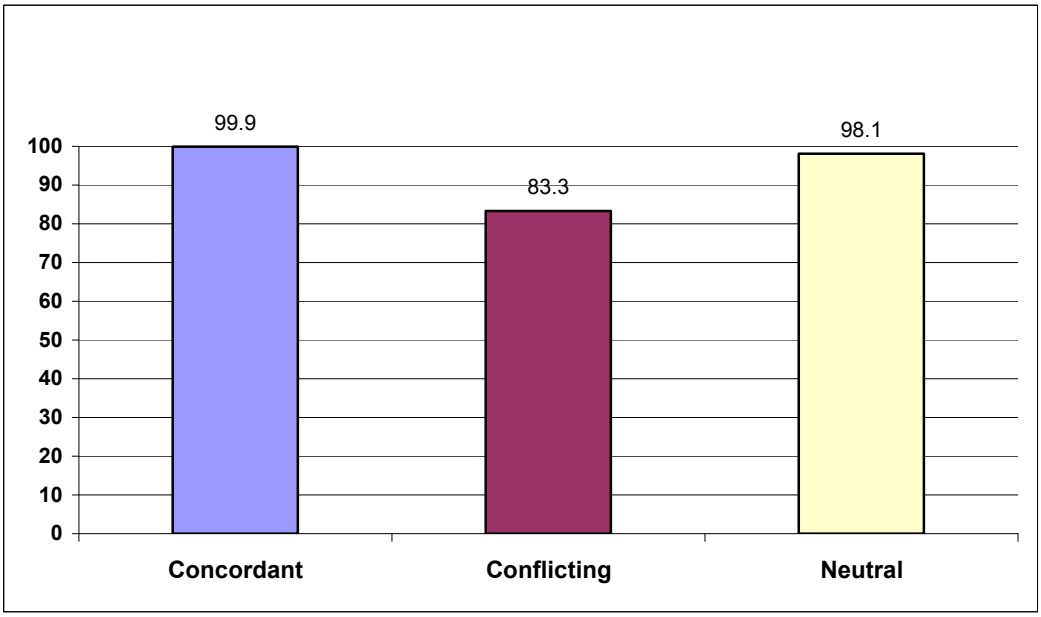


Figure 15: Percentage of correct responses in terms of noun ending for TD children



The difference in correct responses between masculine and feminine is statistically significant [*t(30)*=2.977, *p*=0.006]. Figure 15 shows that the conflicting condition, in which the phonological form of the novel noun is not concordant with the phonological form of the determiner, triggered more incorrect responses than the concordant and neutral conditions. The differences between the conflicting condition and both the concordant and neutral condition are statistically significant, respectively [*t(32)*=3.416, *p*=0.002] and [*t(38)*=3.109, *p*=0.002], whereas the difference between concordant and neutral conditions is not statistically significant [*t(50)*=0.407, *p*=0.69].

Out of the 87 errors the typically developing children made, 83 consisted of errors in which the value of the gender introduced by the determiner accompanying the novel noun was feminine (therefore, requiring an utterance with a marked feminine form). In other words, in 83 instances, children produced responses such as 'o vermelho' ('the$_{masc}$ red$_{masc}$'), 'amarelo' ('yellow$_{masc}$') or 'esse' ('this$_{masc}$') when the context required feminine responses. Only four responses requiring a masculine form were registered as incorrect, i.e., children produced utterances such as 'a vermelha' ('the$_{fem}$ red$_{fem}$'), 'amarela' ('yellow$_{fem}$') or 'essa' (this$_{fem}$), when the context required masculine responses.

Before reporting the findings for the children with SLI, let us discuss the results of the typically developing children.

7.5.3.1.1 Discussion

Typically developing children performed well overall. Virtually no errors were made on the *concordant* and on the *neutral* conditions. They only produced a substantial number of incorrect responses (17%) in the *discordant* condition, in which the phonological form of the recently-learned noun did not match the phonological form of the determiner and these errors tended to be with feminine target utterances. The performance of the typically developing children on this task is compatible with the findings obtained in the study carried out by Corrêa and Name (2003), on which the current experiment was based. In Corrêa and Name, two groups of typically developing children were tested on a virtually identical task: a group of children younger than three years of age (mean age 31 months) and a group of children older than three (mean age 54 months). They showed that children produced significantly more errors in the *discordant* condition than in the *concordant* and *neutral* ones. In other words, children demonstrated that they were sensitive to a correlational phonological pattern between determiner and noun. Nevertheless, their results showed a developmental trend: the group of older children produced a higher percentage of incorrect responses in the *discordant* condition than the group of younger children, indicating that sensitivity to the phonological pattern increased with age. The vulnerability to congruence effects in the older group is argued by Corrêa and Name to be due to these children's greater lexical knowledge. A critical mass of lexical representations in the children's mental lexicon is presumably required in order for a phonological pattern to be identified. In spite of the errors observed in the *discordant* condition, the lack of significant differences between *concordant* and *neutral* conditions provides convincing evidence for the idea that a phonology-gender co-relation is not the basis for the establishment of gender agreement. Rather, gender agreement in Portuguese seems to be acquired on the basis of information conveyed by the functional category D(eterminer): the identification of morphophonological information within the closed class of determiners and the parsing of agreement between determiner and noun are arguably the crucial factors in assigning gender to novel nouns in acquisition.

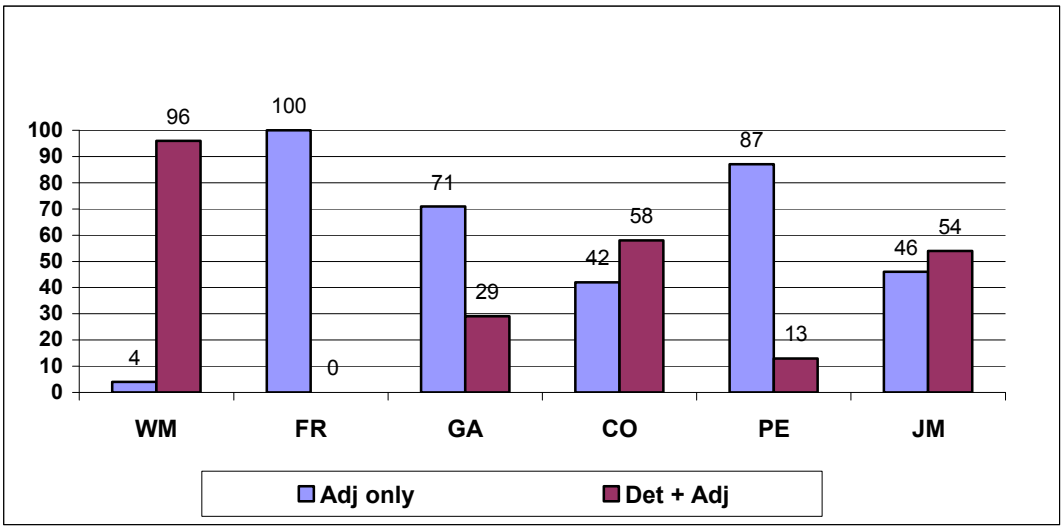
Given the explanation put forward by Corrêa and Name, it seems highly plausible that the typically developing children who took part in the current study behaved similarly. In other words, it seems that their mistakes in the *discordant* condition resulted from the fact that these children have already acquired a

lexicon large enough for phonological patterns to be established. Nevertheless, they were well above chance, even in this condition.

7.5.3.2 Children with SLI

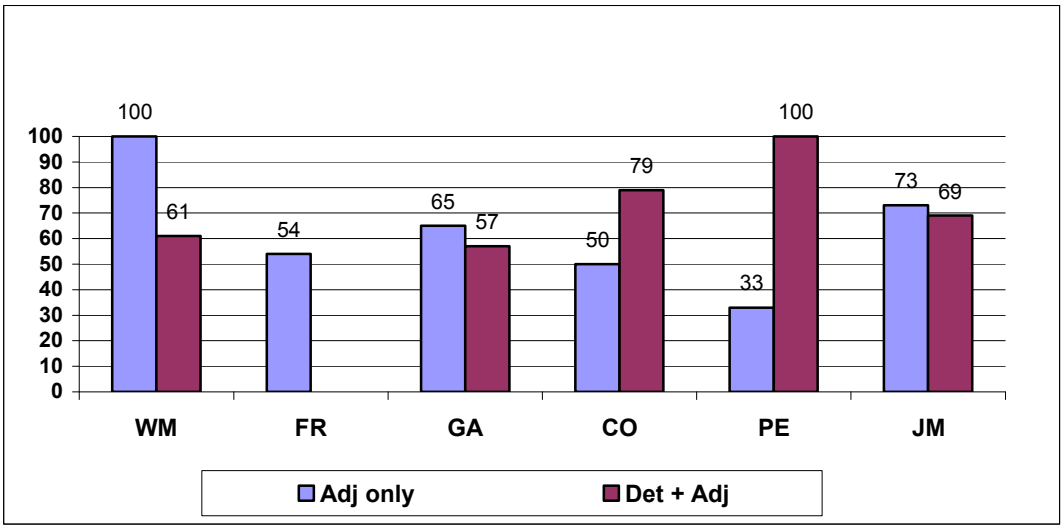
The children with SLI also varied their responses in terms of the type of DP used. However, unlike the typically developing children, they never reproduced the novel noun that was introduced by the experimenter nor did they use demonstrative pronouns in any of their responses. The figure below presents the distribution of their answers across the two types of responses they produced.

Figure 16: Percentage of responses in terms of response type for children with SLI



In general, errors occurred across both types of utterance reported in the figure above. A figure with the percentage of errors each child produced according to the items of the DP included in the utterance is presented below. It is important to note that, unlike the previous experiment, where the determiner was almost always correct, the incorrect utterances in which both a determiner and an adjective were produced were wrong because both the gender of the determiner and the gender of the adjective were incorrect in relation to the novel noun introduced by the experimenter (e.g. 'o vermelho' 'the_{masc} red_{masc}' when referring to a character or object introduced as feminine, like 'uma dobe' 'a_{fem} dobe').

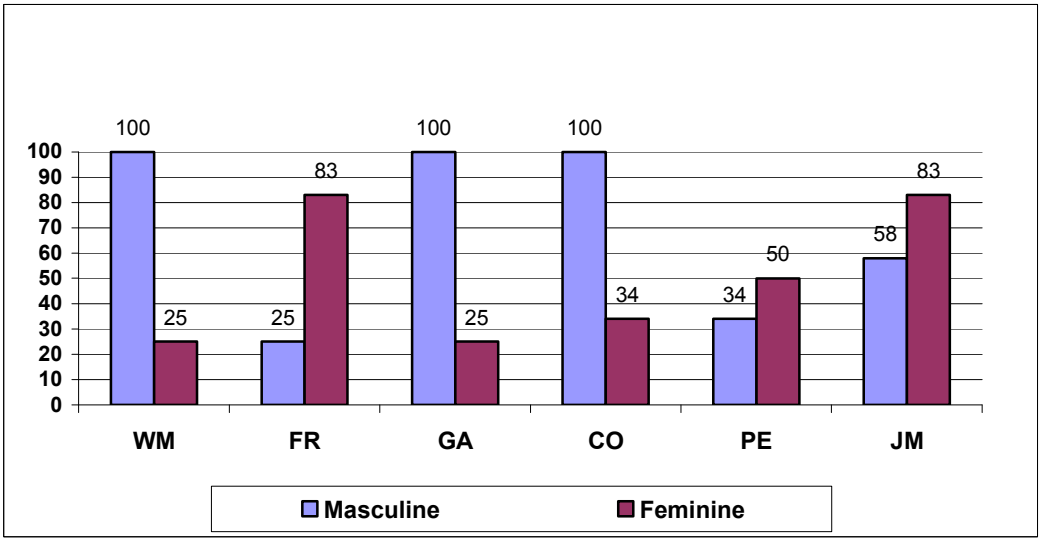
Figure 17: Percentage of correct responses in terms of response type for children with SLI



Recall that each participant was presented with 24 test items. The raw error scores for each child are as follows: WM and GA got 15 items correct, FR got 13 of them right, CO gave 16 correct answers, PE produced 10 correct items, and, finally, JM got 17 correct responses. A breakdown of the pattern of responses is presented below.

Figure 18 presents a breakdown of correct responses according to target gender of the novel noun (masculine vs feminine) as indicated by the gender of the determiner presented to the child.

Figure 18: Percentage of correct responses in terms of target gender for children with SLI



This figure shows that WM, GA and CO made errors only when they were supposed to produce an utterance with a marked feminine form, while FR, PE and JM, even if they did make errors when the target response was feminine, produced more errors when the target response was masculine.

With the masculine target, WM, CO and GA are performing above chance level (p-value<0.01), while PE and FR are at chance level. In the feminine target all children are performing at chance level, apart from FR and JM who are performing above chance level (but only at a p-value<0.05).

Figure 19 presents the percentage of correct responses according to noun ending.

Figure 19: Percentage of correct responses in terms of noun ending for children with SLI

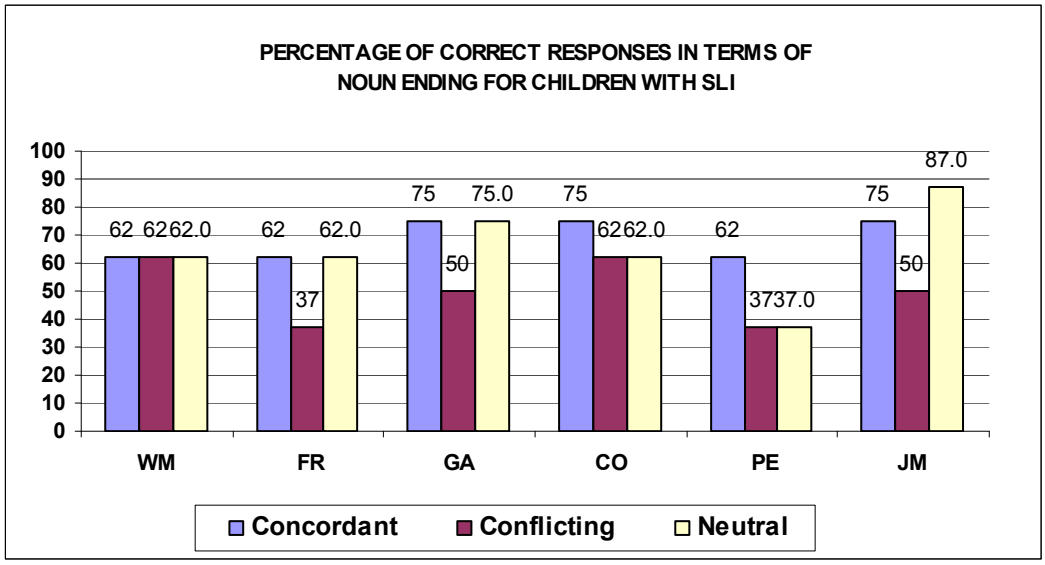


Figure 19 shows that the children with SLI produced many mistakes across conditions, and these were not confined to the *conflicting* condition. The percentage of correct responses in the *conflicting* condition is lower for some children, but the difference between conditions is not nearly as striking as it was for the typically developing children.

7.5.3.2.1 Discussion

The main findings of this experiment were the following:

1. The overall performance of the children with SLI was poor

2. Three of the children with SLI (WM, GA and CO) performed more poorly when the target utterance was feminine than when it was masculine.

3. While, in general, children with SLI produced more errors in the *discordant* condition, these children encountered difficulties across conditions.

4. Children never repeated the recently-learned noun and provided mostly answers with a determiner and an adjective or with a single adjective.

5. All the errors made by the children with SLI in this task consisted of a wrong gender being assigned to the recently-learned noun.

6. No gender violation between determiner and adjective was observed in the production of children with SLI. Incorrect responses containing a determiner and an adjective always matched in gender. They were considered incorrect not because of a gender violation between the elements produced, but because the gender of the elements produced by the children was wrong in relation to the target introduced by the experimenter.

The current experiment yielded interesting and informative results. As anticipated earlier, caution is needed when interpreting the results given the relative advanced age of the participants. Nevertheless, it is undeniable that the performances of the two groups of children were strikingly different, indicating that children with SLI deal with gender information in novel encounters in an unstable manner, with the same child alternating correct and incorrect instances.

A series of related questions regarding the performance of the children with SLI can be raised. Does their relatively poorer performance mean they do not notice the input being provided by the experimenter? Are they performing at chance level? Could their poor performance be the result of a general lack of attention to the task? As the performance of the children with SLI was not homogeneous as a group, a more detailed look at individual or subgroup data is warranted. WM, GA and CO showed a very strong tendency to use the default masculine form (cf. Figure 17). All their responses for a masculine target were masculine and more than half of their responses for a feminine target were mistakenly masculine. This pattern, nevertheless, gives us some indication that, although these three children tended to provide masculine responses, they did

notice the input being given by the experimenter, even if perhaps only sporadically: these three children never produced a feminine form when the context required a masculine form. It seems reasonable to say that if WM, GA and CO had not noticed the input, a more random pattern of responses would have been registered. The difficulties of these children, therefore, seem to be with the marked feminine form.

FR, PE and JM showed a different pattern compared to WM, GA and CO. These children showed a tendency to produce feminine responses. They, however, produced a large number of incorrect responses with both target genders, indicating that they might not be using the input at all. Their bias towards producing feminine responses was, nevertheless, unexpected.

Although Experiment 4 and the current experiment use different methodologies, it is interesting to establish some comparisons. The fact that the two tasks used different types of words as stimuli (known nouns vs novel nouns) may assist us in raising relevant questions and shed light on the behaviour of children with SLI when faced with novel nouns. Recall that, in Experiment 4, almost none of the errors made by the children with SLI consisted of a wrong gender being retrieved for a given known noun. All but four of the incorrect responses these children produced were wrong because the gender of the adjective accompanying the noun did not agree with the gender of the noun. In the current experiment, in contrast, all the errors that the children with SLI made consisted of a wrong gender being retrieved for (or assigned to) a recently learned noun. Whenever both a determiner and an adjective were produced, they matched in gender. It seems, therefore, that the difficulties showed by children with SLI were distinct in each task. In Experiment 4, the context led children to produce the known nouns that were targeted and children used stored information about these nouns in their responses. In Experiment 5, children did not produce the recently-learned nouns (the experimental context did not require it) and there was no stored information about the nouns to be used because they were all novel nouns. Interestingly, typically developing children were very often able to grasp enough gender information from the few utterances provided by the experimenter and correctly assign gender to the novel noun in question. In other words, the performance of the typically developing children was robust even with only a few encounters with the novel word. The children with SLI, on the other hand, often failed to assign the correct gender. What do the current results tell us

about children with SLI? In Experiment 4, we saw that these children used stored information about known nouns, while, in the current experiment, the children with SLI were not able to systematically use information provided by the experimenter to assign the correct gender to novel nouns. How long it took for these children to store information about the nouns in Experiment 4 is an open question. Taking into consideration the results of the current experiment, it seems reasonable to suggest that children with SLI might need more exposure to utterances with novel nouns in order to extract the relevant information.

Why would children with SLI need exposure to more instances of the same novel word in order to be able to assign its gender? As we saw in chapter 4, the acquisition of gender agreement under normal circumstances arguably takes the following path: the identification of morphophonological alterations within the closed class of determiners and the parsing of the DP bootstraps the grammatical operation of agreement enabling children to assign gender to a novel noun (Name, 2002; Corrêa and Name, 2003).

According to Name (2002), the characterisation for gender acquisition outlined above opens the possibility for some sort of fluctuation of gender feature while it is not yet fully stored (or while a given lexical entry is not yet fully stored with a gender feature of a given value). Name reports that such fluctuation was observed for two participants in her study with typically developing children:

> "A child (age 3;1) was introduced to the image of a new object which was named 'the$_{fem}$ puco". The same object then appeared in a different color, within a brief story. When asked which "puco" had fallen down, the child replied: "Essa aqui" ("This$_{fem}$ one"), referring to the gender conveyed by the determiner. At the end of the experiment, the child wanted to show the images to his/her mother, naming the object as "o puco" ("the$_{masc}$ puco"). Something similar happened with another child. It is as if the lexical entry were unstable, in working memory and, when phonological information is not accessible any longer (**a** puco), other strategies would help. With multiple exposure, the entry is stored in the lexicon with its gender value fixed." (Name, 2002: 142; my translation)

Could something similar to what is reported in Name be applicable to children with SLI but on a larger scale?

Additional support to the idea that children with SLI might take longer than the typically developing children to accomplish the storage of gender values to novel nouns comes from two sources: 1. Experiment 3 of the current thesis registered very few incorrect responses. However, among the 6 errors made, 3 were given by child PE, who was the youngest of the participants with SLI (age 5;5); 2. in the French study conducted by Jakubowicz and Roulet (2007), the group of children with SLI was divided into two smaller groups according to age. The group of younger children produced a higher percentage of gender violation errors than the older group.

It could also be argued that the children with SLI performed poorly in the current experiment due to difficulties in processing the input provided by the experimenter, i.e., gender agreement in comprehension modality. On the basis of the model put forward by Name and Corrêa, the ability to process agreement between determiner and noun is crucial for assigning gender to a novel noun. This, however, does not seem to be a problem for the children with SLI, or at least not their major difficulty. If children could not process agreement, it is likely they would be relying on noun endings, which did not happen (either in the current experiment or in any other task carried out for this thesis or in any of the experiments reported by Jakubowicz and Roulet). In addition, the two input processing experiments conducted by Jakubowicz and Roulet (cf. chapter 5) also showed children with SLI did not rely on noun endings when providing their answers.

Interestingly, gender in the adjective, which was a major problem for children in Experiment 4, was not a particular issue in the current experiment. In Experiment 4, the great majority of errors children with SLI produced consisted of gender mismatch in the adjective but not in the determiner (i.e. the gender in the adjective was wrong in relation to the noun and determiner). Here, on the other hand, a different pattern occurred: children did make mistakes with adjectives but, whenever that happened, the gender of the determiner was incorrect as well. Mistakes consisted of incorrect gender in determiner and adjective in relation to the recently-learned noun. Are the explanations put forward to explain problems in adjective agreement in Experiment 4 compatible with these results? Why is it that children make mistakes with adjectives when a known noun is present but do not when a noun is absent? It seems reasonable to suggest that, given the fact that nouns in Experiment 5 were novel nouns, just recently introduced to the

children, an incorrect response should not be characterised in terms of an 'agreement' error. Three of the children with SLI (WM, GA and CO) made errors only when the target novel noun was feminine. In other words, they provided a masculine response when the correct one should have been a feminine response. Since their responses did not contain a noun and the noun they were referring to was not part of their stored lexicon (at least not in the same way as a known noun), it is possible that these mistakes are the result of children simply resorting to masculine default forms, without actual agreement between elements taking place.

The performance of the other three children with SLI did not follow the pattern above. FR, PE and JM made more errors when the target novel noun was masculine, with FR showing a relatively strong tendency to provide feminine responses. Could it be that FR used feminine forms as a sort of default in this experiment[56]?

7.6.6 EXPERIMENT 6

7.6.1 Introduction

As discussed in Experiment 2, testing the input processing of gender agreement is difficult. An attempt to test them by grammaticality judgment was reported in Experiment 2 and we saw that, in addition to the logistical problems I encountered, the methodology itself does not seem to work with the age group. Experiment 6 is a second attempt to test the input processing abilities of the children with SLI. Given the unfeasibility of 'directly' testing these abilities, the current experiment aimed to investigate the extent to which children with SLI would be able to select the target item in a picture pair on the basis of a description provided by the experimenter, where grammatical gender was the only cue to resolve a potential ambiguity.

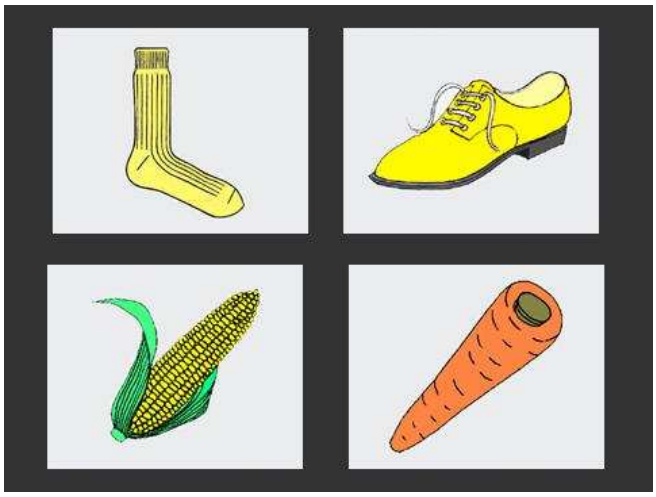
7.6.2 Method

On a computer screen, children were shown pairs of cards depicting single objects. The two objects in each pair were similar to each other in that they shared properties such as colour, shape or function (e.g. corn and carrot, which are edible and have an elongated shape). The experimenter described a single object, always providing two different physical characteristics. Two pairs of

[56] See the discussion regarding the notion of default on page 93.

objects appeared at a time on the computer screen (Compaq nx9010 laptop computer), as shown below.

Example:



Procedure

Children were invited to play a game with the experimenter, in which he/she had to guess what the experimenter was describing. The task proceeded as follows:

Experimenter:

— "Agora nós vamos brincar de 'o que é o que é'. Eu vou falar algumas coisas coisas para você e você tem que advinhar do que estou falando."

(*"Now we are going to play a 'guess what it is' game. I am going to say certain things to you and you have to guess what it is that I am describing*")

— "Vamos começar?"

("*Let's start?*")

— "O que é o que é? A gente usa ela no pé. È amarela.

(*"Guess what it is. We wear it$_{fem}$ on our feet. [It] is yellow$_{fem}$*")

(give some time for child to reply)

— "O que é o que é? A gente come ela. È comprida.

(*"Guess what it is. We eat it$_{fem}$. [It] is long$_{fem}$"*)

Children were, thus, required to use gender cues present in the pronoun and in the adjective produced by the experimenter in order to select the adequate picture.

The task targeted 20 nouns in four different gender categories:

1) masculine gender, typical ending (5 items)

e.g. milho ($corn_{masc}$)

2) masculine gender, non-typical ending (5 items)

e.g. lápis ($pencil_{masc}$)

3) feminine gender, typical ending (5 items)

e.g. cenoura ($carrot_{fem}$)

4) feminine gender, non-typical ending (5 items)

e.g. moto ($motorcycle_{fem}$)

Predictions and possible outcomes

This experiment was designed as an attempt to test gender input processing abilities. In addition to investigating whether children with SLI are able to process determiner and noun agreement in the comprehension modality, the task seeks to examine whether noun ending affects the children's performance. To my knowledge, no task similar to this one has been carried out before. Gender information referring to inanimate nouns is quite abstract and never used as a way to disambiguate between referents in the way it might be with animate nouns. Thus, the methodology, while interesting because it allows for inanimate nouns to be used, might not be ideal. Nevertheless, the following outcomes can be anticipated:

1. If the typically developing children perform poorly, contrary to expectation, this is ambiguous between difficulties due to methodology and difficulties with gender.;

2. If the children with SLI are able to use input processing gender information to complete the task, this could indicate that their gender comprehension skills are intact;

3. If the performance of the children with SLI is affected by the type of noun ending, it is expected that most errors would occur with non-typical noun endings.

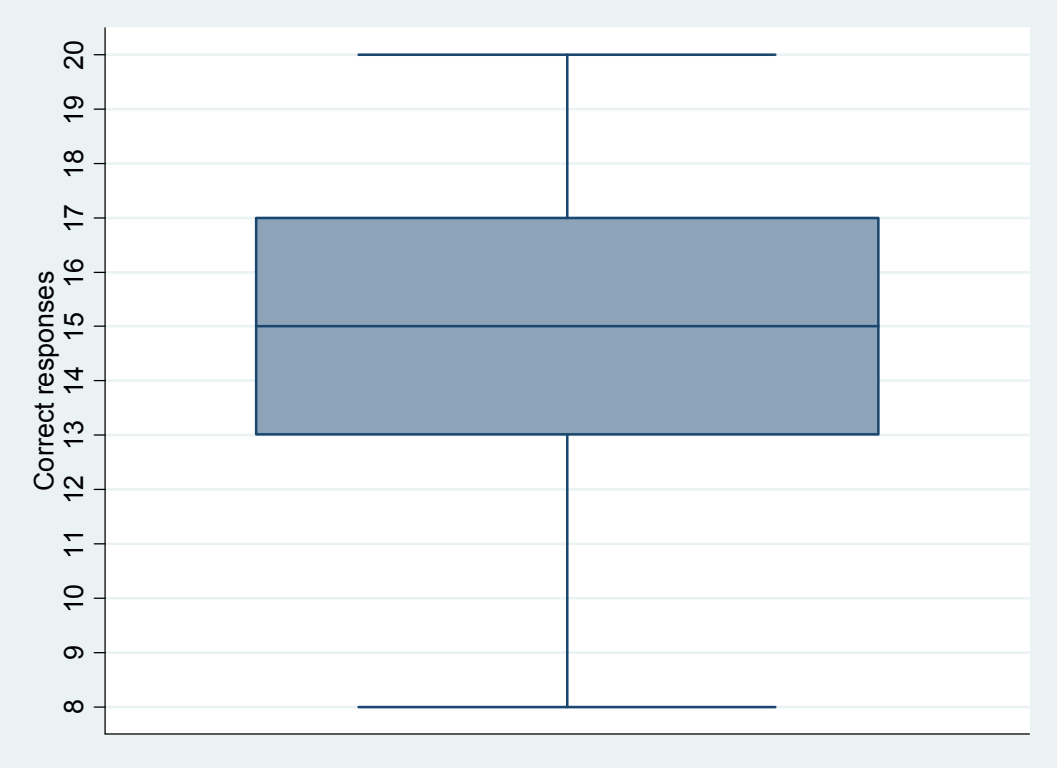
7.6.3 Results

7.6.3.1 Typically developing children

The performance of the typically developing children in the current task varied considerably. Some children completed the task without making any errors while others performed poorly. However, no significant difference between the control groups of higher and lower income children was found [$t(30)=0.60$, $p=0.5522$], so

they were treated as a single group. Figure 20 below illustrates the variability in the performance of the typically developing children.

Figure 20: Mean scores for typically developing children



The typically developing children produced a mean of 15.12 correct responses (sd 2.84). In other words, seventy-six per cent of their responses were correct, i.e. the right picture was selected. Out of the 24% incorrect responses, 47% consisted of feminine targets (mean 2.29, sd 1.71, e.g. 'cenoura' – 'carrot$_{fem}$') while 53% consisted of masculine targets (mean 2.58, sd 1.61, e.g. 'milho' – 'corn$_{masc}$'). Fifty-three per cent of all errors were made when the targeted noun had a typical ending (mean 2.58, sd 1.84, e.g. ('cenoura' – 'carrot$_{fem}$' or 'milho' – 'corn$_{masc}$') and 47% when it had a non-typical ending (mean 2.3, sd 1.51, e.g. 'lápis' – 'pencil$_{masc}$' or 'moto' – 'motorcycle$_{fem}$').

7.6.3.2 Children with SLI

The raw scores for each child with SLI are as follows:

Table 8: Raw scores for children with SLI

WM	9	CO	11
FR	12	PE	11
GA	13	JM	10

Figure 21 below shows the percentage of correct responses produced by the children with SLI according to gender (masculine vs feminine).

Figure 21: Percentage of correct responses in terms of gender

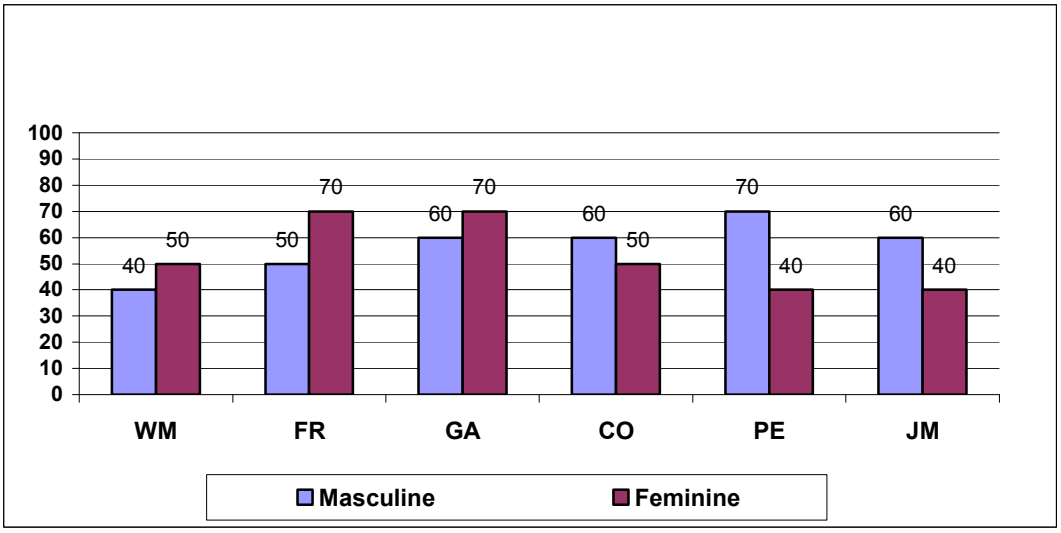
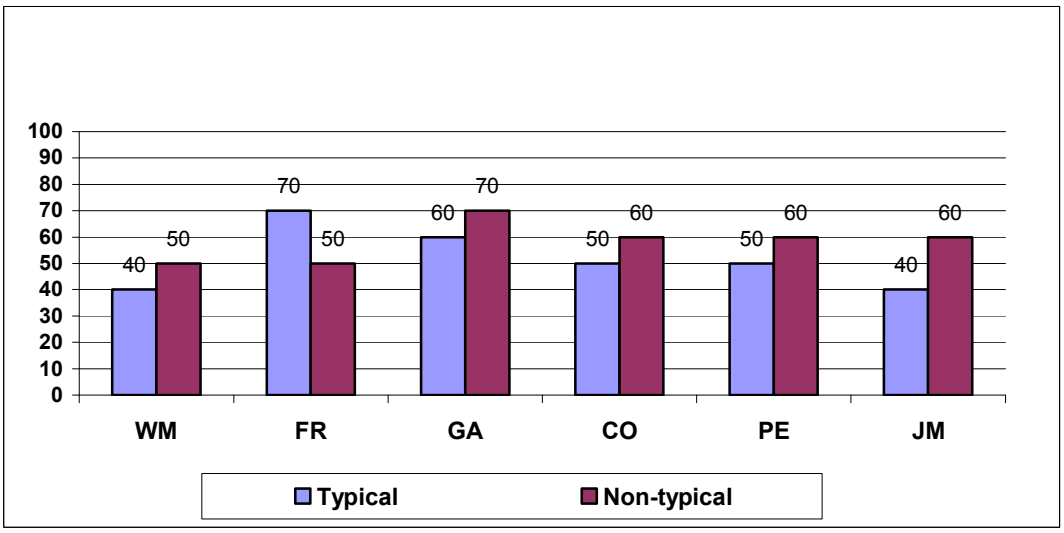


Figure 21 above shows that the children had problems with both masculine utterances and feminine utterances.

The figure below presents the percentage of correct responses according to the typicality of the ending of the noun being targeted by the experimenter (typical vs non-typical).

Figure 22: Percentage of correct responses in terms of noun ending



7.6.4 Discussion

The current task was designed in an attempt to provide a suitable context for investigating the input processing of grammatical gender. As discussed in the introduction, creating such a context is challenging and a standard picture selection task cannot be used. The methodology applied here involved a substantial level of abstract metalinguistic reasoning, as gender information in inanimate nouns is not commonly used to disambiguate between referents and, as such, the task presented difficulties even for typically developing children. Performance within the group of typically developing children varied considerably and not all children found the task demanding. Age did not seem to be a factor influencing their performance. All the six children with SLI, on the other hand, performed poorly on the task. In spite of the fact that the group of children with SLI is extremely small, it is interesting that every single child found the task demanding. It could be argued that such performance is due to their low sensitivity to linguistic stimuli, particularly morphological distinctions. But it should be born in mind that the typically developing children overlapped with the SLI group despite having robust gender distinctions in output tasks.

7.7 Summary

This chapter presented the results of six experimental tasks conducted with children with SLI and a group of typically developing children. Two groups of typically developing children were originally recruited on the basis of social background but were later combined as no significant differences were found. The experiments sought to investigate aspects of gender agreement processing in Brazilian Portuguese SLI. Tasks were designed in order to test different structural relationships within the DP, as well as different stages in the processing of a gender marked utterances.

An overall analysis of the experiments reported in this chapter shows roughly two different types of outcome. Generally speaking, the experiments investigating the production of grammatical gender yielded clear differences between the SLI and the typically developing children. In contrast, the experiments exploring the input processing skills of grammatical gender yielded variation in the group of typically developing children and occasionally in the SLI group, with overlap between the two groups, suggesting that they are less informative about SLI. It was argued that these tasks require a high level of

metalinguistic and cognitive abilities rather than tap basic language skills. I take up the issues raised by this pattern of findings in the final chapter.

Chapter 8
DISCUSSION AND CONCLUSION

8.1 Summary

This thesis investigated Specific Language Impairment in children. Part I offered a critical analysis of SLI, evaluating the fluctuation in the use of the term language among various researchers in the field. Part II reported an experimental study exploring grammatical gender within the Determiner Phrase.

We saw that the field of SLI has been approached by a variety of disciplines in the last few decades. These disciplines share an interest in the human capacity of acquiring a language but do not necessarily share a similar concept of the term *language* itself. On the one hand, we have researchers who use the term *language* as a term for the concept of communication, including in their research program a broad range of phenomena. It is questionable whether some of these phenomena are part of the language domain. On the other hand, there are researchers who work with a much more restricted conception of the term *language*, leaving out of their investigations factors that are related to language but not necessarily part of it. While the latter group of researchers focuses on what could be characterised as 'basic' language skills, i.e. those skills acquired relatively early, without formal instruction, by any typically developing child, the former group includes a wide range of aspects involved in verbal communication, those acquired without formal instruction but also those which are learned via instruction and/or relatively later.

I showed that this fluctuation in the use of the term language has had a strong impact on the way research on SLI is carried out, with problems mainly residing in inappropriate diagnostic tools used to recruit children. Moreover, I argued that many of the controversies that surround the field of SLI actually originate in the variability of interpretation of the term *language*. A thorough analysis of tests that are often used by clinicians and researchers, such as the TOLD, the CELF and the TROG, showed that SLI is indeed a heterogeneous 'label', as many researchers have claimed, but it is also a label currently being used to place together children with potentially distinct difficulties.

We saw that the TOLD's and the CELF's subtests present many problems in their design. The subtests assessing vocabulary, for example, make use of a

large quantity of lexical items that are formal, infrequent and arguably dependent on formal schooling. An additional problem identified among the vocabulary subtests refers to the need to draw on encyclopaedic knowledge in order to provide appropriate responses. This is the case with the TOLD's subtest entitled *oral vocabulary*, where children are supposed to give definitions of words. We saw that the subtests testing grammatical structures are also subject to criticisms. One of the main problems of such subtests reside in the fact that many test items involve sentences that are very difficult to represent pictorially, demanding a high level of inference from the children for successful completion. In addition to the issues in the actual tests, we saw that their manuals are problematic. A look at the CELF's manual, for example, indicated that its authors make use of a variety of different terms to refer to the population they wish to target, such as 'learning disabled children', 'language and learning disabled student', 'learning disabled adolescents', 'language disabled student' and 'language and learning disabled high school student'. Interestingly, the authors often compare the performance of 'disabled children and adolescents' with the performance of 'their academically achieving age peers'. This fluctuation in terminology and the use of academic-related nomenclature cast doubts on the adequacy of the CELF for identifying potential cases of SLI.

An analysis of the TROG showed that, although the test is much less subject to criticisms than the TOLD and the CELF, it still presents crucial problems. The analysis of the TROG focused mainly on the way relative clauses (RCs) are assessed and it showed that this violates both the felicity conditions, which, according to Hamburger and Crain (1982), are necessary for the interpretation of RCs; and Grice's Maxim of Manner, which, also according to Hamburger and Crain, needs to be fulfilled in the experimental setting. These problems arguably weaken the appropriateness of the test.

Following the analysis of the current state of the field of SLI, I proposed two alternative scenarios for the disorder: one according to which SLI is not a basic language disorder per se but a condition affecting academic performance and another according to which SLI is a valid category but only if approached in its narrow conception.

Under Scenario 1, I claimed that what is usually characterised as SLI should, instead, be defined as a failure in achieving the necessary

communication skills for educational success, bringing a socio-cultural perspective into the debate on SLI. This proposal was supported by a number of other findings in addition to the analysis of the main tests used in the clinical and research contexts. We saw that the literature on SLI reports a relatively high percentage of test bias, with most minority populations in the United States performing below the 'white' majority in many popular tests. While it was outside the scope of this thesis to thoroughly analyse socioeconomic and cultural issues impacting on linguistic and academic performance, we briefly discussed widely reported data concerning the performance of minority groups in educational settings, which provide supporting evidence to the claim that SLI might be better characterised as a condition affecting school performance. It has been extensively documented that minority groups perform less well on tests for university admission, have generally lower grades, present higher rates of school dropout, and lower rates of high school completion, college transition and completion. Interestingly, however, education attainment of Mexican immigrants improves considerably once factors such as generation, language and social capital are controlled for. We also saw that one exception to the tendency of minority groups performing less well in educational settings is the Asian population in the United States. In general, Asian immigrants perform even better than white students. Notably, none of the Asian participants in the classic SLI study carried out by Tomblin and colleagues were found to present SLI. Contrary to other minority students in the United States, however, Asian students are extremely advantaged in terms of parental education levels. Moreover, we saw that HIP Asia (highly performing Asia — Japan, South Korea, Singapore, Hong Kong and Taiwan) seem to culturally pay more attention to education than some other regions. For instance, while the public sectors of both Latin America and HIP Asia countries spend similar amounts on education, Latin Americans concentrate their expenses on higher education and HIP Asian countries give more importance to primary education. I suggested that the low rates of SLI among the Asian population within the United States can perhaps be explained by the high parental education levels of Asian descendants, along with the importance the culture gives to education.

I also brought into the discussion of the socio-cultural-educational aspects of SLI the issue of law and school policies in North American schools (and possibly British schools). According to the authors of the TOLD themselves, the test can be useful in evaluating children's progress in prescribed remedial

programs, an important component of special education, and often required by law or by school policy. I argued that an overdiagnosis of SLI might be taking place, as has been suggested in the case of ADHD (Attention Deficit/Hyperactivity Disorder). If this is true, it would take place in countries such as the United States and the United Kingdom, where government school systems are relatively homogeneous and dedicated funding for special education programmes is available. In the case of countries like Brazil, no such pressure for extra funding and, therefore, no pressure to 'label' students according to guidelines, exist, as schools need to deal with more basic issues, such as insufficient numbers of teachers.

Under Scenario 2 outlined in chapter 2, SLI could be a valid category. The scenario was grounded in a narrow conception of the term *language* and the idea that an impairment in basic language skills, those which are independent of formal instruction and naturally acquired under normal circumstances, exists as a logical possibility and therefore should be explored, at least as a starting point. SLI under Scenario 2 would affect a considerably lower percentage of school age children than the 7% that is estimated by the current literature. Scenario 1 would, thus, be partially applicable. The problems that many of the children who are currently diagnosed as SLI encounter would be accounted for under an alternative explanation, possibly one focusing on problems related to communication skills needed for school achievement. The proposal of Scenario 2 is consistent with the view that different kinds of information are handled by different systems. Under this view, it is reasonable to consider that the spontaneous and unintentional acquisition of a native language is not governed by the same mechanisms involved in learning information via formal instruction. I argued that research on SLI must be able to differentiate between problems in these different domains, which, at the moment, are conflated.

Scenario 2, which I have advocated, clearly shares a modular view of language with the current hypotheses of SLI grouped together under 'linguistic hypotheses'. However, it was argued that the hypotheses formulated by authors such as Clahsen, Wexler and van der Lely reveal some inconsistencies between the arguments promoted and the way subjects are diagnosed in their research. While these authors use a narrow conception of the term *language*, and their research does focus on relevant linguistic structures, when recruiting children for their studies, they make frequent use of standardised tests that were criticised in

chapter 2 of this thesis. In other words, they make claims that are at least partially incompatible with their recruitment methods, as the tests in question involve a very wide range of skills and not just those related to these authors' view of SLI. Based on this finding, I argued for the following: while these authors formulate hypotheses about the nature of SLI, it is possible that they are providing no more than descriptions of the patterns that emerge in the data they collect. Highly specific and selective claims about SLI are being made, but these claims are often based on samples that are recruited on relatively broad terms. In my view, only theoretically-motivated and well-grounded diagnostic procedures, from the beginning of any study, can identify participants with a consistent and appropriate profile.

The second part of this thesis presented a behavioral study conducted with Brazilian Portuguese-speaking children with SLI, investigating their processing of grammatical gender. This experimental study attempted to explore the view of SLI outlined under Scenario 2, as discussed in Part I. The study sought to take a narrow approach to SLI from its start: children were recruited on the basis of a test which arguably avoids many of the problems associated with other assessments often used in the field. The test which was used, referred to as MABILIN, focuses on core linguistic skills and was constructed on the basis of carefully-controlled variables. It was argued that MABILIN is a more appropriate test for use in SLI studies than widely used tests such as the TOLD, the CELF and the TROG. Moreover, the investigation reported in Part II aimed to test grammatical gender, a very specific linguistic phenomenon, and the goal was to carry out tasks tapping linguistic knowledge rather than wide-ranging knowledge assessed by tests such as the TOLD and the CELF. Furthermore, the study was conducted under the assumption that an integrated and conciliatory approach to SLI is preferred, one in which linguistic and processing accounts to the disorder complement each other. A proposal based on Marr's (1982) work on vision cognition was discussed as an interesting way to deal with the relationship between Linguistics and Psycholinguistcs, whereby linguistic and psycholinguistic theories complement each other.

An extensive literature review of aspects concerning grammatical gender agreement was carried out. In chapter 4, we saw that gender is adequately defined on grammatical grounds and its determining criterion is *agreement*. The literature review on the linguistic aspects of gender revealed that the status of

adjectives is still controversial. After an analysis of the main accounts of adjectival placement within the DP, the approach according to which adjectives are analysed as adjuncts (base generated to the right in most cases in Romance languages) was considered the most appropriate in capturing the behaviour of predicative adjectives in Portuguese. With respect to the phenomenon of gender agreement, the literature review indicated that the *probe* and *goal* approach, based on the latest developments of the minimalist program, works in the case of determiner and noun agreement but is not able to account for adjectival agreement. We also looked at Grimshaw's extended projection theory, as an alternative way to account for determiner and noun agreement and presented a tentative proposal to account for adjectival agreement. This tentative proposal used as a basis Higginbotham's theta identification theory and developed the idea of noun and adjective configured as constituents whose theta-roles are identified. The thesis assumed that masculine forms are unmarked forms with respect to gender in Portuguese, such that feminine is a feature and masculine, in turn, is equivalent to the absence of a feature. Two different possibilities on the basis of Higginbotham's theory were explored: one according to which, as a result of *theta identification*, adjectives are not specified for gender and another according to which theta identification results in adjectives being specified for gender. Under the first alternative, no gender feature copying conditioned by *theta identification* would take place but *theta identification* would provide the context for a spell-out rule according to which, at the level of vocabulary insertion, the adjective needs to get its form by looking at the noun which stands in this relation of *theta identification*. Under the second alternative, on the other hand, a morphological rule of feature copying conditioned by *theta identification* would apply.

The literature review continued in chapter 5, with an overview of the psycholinguistic aspects of gender agreement. First we saw that little is known about gender in SLI, although some research has been carried out in Portuguese, French and Spanish. Although very preliminary, the data from these studies indicated that gender agreement is an area of potential difficulties for children with SLI, particularly adjectival agreement. We also looked at gender acquisition by typically developing children, which is reported to evolve smoothly, with few errors, and at a relatively early age in many languages. While most studies to date assume that grammatical gender is somewhat idiosyncratic and that its acquisition depends on general learning processes, sensitive to

frequency, phonological cues and semantic properties, recent studies by Corrêa and Name take a different perspective, namely one which views gender acquisition as a process dependent on syntactic mechanisms. According to these authors, children (at least those acquiring Portuguese and, possibly, other Romance languages) use information conveyed by the closed class of Determiners in order to assign gender to a novel noun. The parsing of morphologically marked gender classes within the category D would 'bootstrap' the syntactic operation of agreement, enabling the gender of the determiner to be assigned to the noun.

The final section on the psycholinguistic aspects of gender reviewed studies on processing of gender agreement within the DP in adults, both neurologically healthy adults and patients with aphasia. I discussed the different stages that need to be incorporated in a model of gender production, taking into account the considerations of linguistic aspects of agreement in chapter 4. In addition, we looked at evidence suggesting that agreement between determiner and noun is a different phenomenon compared to agreement between noun and adjective.

The critical review of literature on the linguistic and psycholingustic aspects of gender provided the background for the 6 experiments reported in chapter 7. For example, the design of the tasks that were carried out relied heavily on the status of the Determiner as proposed by Abney (1987), and discussed in section 4.4.1. Abney's determiner hypothesis is compatible with the gender acquisition model, put forward by Name and Corrêa (cf. section 5.2.2), according to which children use morphophonological information within the DP to assign gender to a novel noun. Name and Corrêa's work served as a basis for Experiment 5 of the current thesis. In addition, the notions of *intrinsic* and *optional* genders guided the construction of the tasks, avoiding, for example, the problems which were previously identified with Anderson and Souto's study with Spanish-speaking children with SLI (cf. section 5.1.2). Another example of how the review of literature on the linguistic aspects of gender provided the background for the experimental study was the discussion about the potential differences between determiner/noun agreement and noun/adjective agreement in chapter 4. Moreover, the idea that determiners are fully specified for gender at the end of the syntactic derivation (per Magalhães' probe and goal proposal and Grimshaw's Extended Projection theory, cf. section 4.5) was incorporated into the

psycholinguistic model of gender agreement production sketched in section 5.3.1 In terms of the review of the literature on psycholinguistic aspects of gender, the evidence suggesting determiner and adjective agreement are processed differently also guided the formulation of tasks whose experimental contexts allowed us to test this in the performance of the children with SLI.

Experiment 1 aimed to investigate whether children with SLI have a problem in retrieving the gender feature of nouns, which could influence the production of correct agreement between determiner and noun. The experiment used a categorisation task in which children had to group pictures depicting inanimate nouns. It was designed to create a context in which children could retrieve the gender feature of the target lexical items without necessarily having to produce any utterance. Administration of the experiment revealed methodological problems when testing both typically developing children and children with SLI, so data collection ceased. Those who undertook the experiment either were not able to understand the criteria and grouped the cards randomly (or according to the criteria used in the practice session) or understood the criteria required for categorising the noun and completed it without problems or mistakes. The limited data available, nevertheless, seem to suggest that successful completion of the task requires knowledge of reading and writing skills. It is, thus, possible to say that this task did not yield a satisfactory means to test basic knowledge of gender.

Experiment 2 was a grammaticality judgement task which aimed to test whether the children with SLI were sensitive to an effect of gender violation between determiners and nouns and, if so, whether they were more sensitive to violations involving a DP whose noun has a typical ending. Incomplete data was collected due to technical reasons but the limited data available indicate children with SLI might not be as sensitive to gender violation in grammaticality judgement tasks as typically developing children tend to be.

Experiment 3 involved an elicited production task in which children had to produce isolated DPs formed by a determiner and a noun. Previous literature on gender in French and Spanish SLI had indicated that the structure determiner and noun could be an area where children with the disorder would encounter problems. The children who participated in Experiment 3, however, made only a few mistakes, showing that gender agreement between determiner and noun

does not seem to be a major problem in Portuguese. They did, nevertheless, produce a few incorrect responses, something that did not occur with any typically developing children in the control group. This experiment was efficient and informative, as it demonstrated that the SLI children were able to produce agreement between determiner and noun.

Experiment 4 also involved an elicited production task, but it explored DPs containing a determiner, a noun and an adjective. The addition of an adjective provided the context for testing whether what seems to be a gender mismatch error in production (i.e. the few incorrect responses reported in Experiment 3 above and the mistakes given by the French children with SLI in the work of Jakubowicz and Roulet) is caused by problems in selecting the correct gender feature or by problems in processing agreement between elements of the DP. The experiment also allowed us to investigate two potentially different types of agreement: determiner and noun vs noun and adjective. The task was inspired by the work by Anderson and Souto (2005) but the version in this thesis made use of a list of test items put together on the basis of well-founded criteria, avoiding the theoretical misunderstandings present in Anderson and Souto's list. The typically developing children performed at ceiling. The children with SLI, on the other hand, made a considerable number of mistakes, mostly regarding adjective agreement. In other words, whenever a determiner was present, children almost always got its gender feature correct, but provided many utterances with a feature mismatch between noun and adjective. These results were discussed and three alternative explanations were outlined. Experiment 4 was very efficient in testing the gender processing abilities of the children with SLI. It provided further evidence that determiner agreement was not a problematic area for these children, while it demonstrated that adjective agreement did pose difficulties for them.

Experiment 5, like Experiments 3 and 4, involved an elicited production task, but it is strikingly different from those as it made use of novel nouns. Imaginary characters participated in brief stories and children were asked to provide an answer with a gender marked item. By manipulating the endings on the nouns used to name the imaginary characters, it was possible to examine the type of information the children with SLI relied on when encountering a situation in which they had to assign gender to a novel noun. Results showed that the children with SLI performed very differently from the group of typically developing

children. While the latter produced only a few mistakes, which were mostly made in the 'conflicting condition' (where gender in the determiner and final vowel on the novel noun mismatched phonologically), children with SLI produced mistakes across conditions.

As Experiment 4, Experiment 5 was very efficient. No methodological problems were observed and very informative results were obtained. Taken together, results of Experiments 4 and 5 demonstrated that the children with SLI were able to produce determiner and noun agreement but found it difficult to assign gender to a recently-encountered noun. These results were interpreted in terms of greater exposure children with SLI might need to achieve stability and a robust knowledge of gender.

Finally, Experiment 6 was designed to test whether children with SLI would be able to resolve a potential ambiguity on the basis of grammatical gender. They were asked to guess what the experimenter was describing and select the appropriate target item in a picture pair. A standard picture selection task is not suitable for testing inanimate nouns, so the current design was used in a tentative way to test input processing abilities. Results of Experiment 6 showed that 76% of the typically developing children's responses were correct, i.e. the right picture was selected. However, their individual performance varied considerably, with some children scoring 100% of correct responses and some others performing randomly. None of the children with SLI performed well. Experiment 6, as Experiment 2, is another illustration of how difficult it is to test input processing abilities concerning grammatical gender

Returning to the research questions formulated in the beginning of chapter 7 and repeated here, we can now consider the major findings of the experiments conducted for this thesis.

The main questions this experimental study sought to answer were the following:

1. At which stage of production does gender processing break down?
    - Are children with SLI able to retrieve the gender of nouns without problems? Or are gender retrieval

difficulties the source of gender mismatch in DP production?

- If gender retrieval is not a problem, what factor(s) cause(s) children with SLI to produce DPs with mismatching gender? Is there a problem in the online processing of agreement? Could it be that children with SLI have difficulties in the encoding of morphophonological information after agreement has taken place?

2. Do children with SLI have more difficulties with nouns that have non-typical endings than nouns with typical endings? In other words, do these children rely on the ending of the noun to produce gender agreement?

3. Do children with SLI have more difficulties with determiner and noun agreement, or with adjectival agreement, or do problems occur equally with both phenomena?

4. What happens when children with SLI encounter a novel noun? Do they behave like typically developing children when assigning gender to a novel noun?

The major findings of the experiments can be summarised as follows:

1. Gender retrieval did not seem to be a problem for the children with SLI. Although the experiment that attempted to test this specifically (Experiment 1) was not completed due to methodological problems, data from other experiments indicate that gender retrieval is not a problem. Alternative accounts for the gender mismatch in the performance of children with SLI were explored and their weaknesses and strengths were assessed. One alternative involved the possibility that problems at the level of processing in which the gender feature of the noun is copied to the adjective caused gender mismatch. A second alternative considered that determiners and adjectives are represented and accessed differently in the mental lexicon, and the production of

adjectives needs to go through an inflectional process, something that would not happen with determiners. A third alternative explored the idea that adjective agreement is dependent upon a 'non-local' spell-out rule provided by *theta identification* by means of which the adjective would get its form. This interface rule could be viewed as instructions for the adjective to get its form from information on the noun.

2. Generally speaking, it is possible to say that children with SLI did not use the ending of nouns as a cue for gender agreement. This finding is highly compatible with the French data in Jakubowicz and Roulet's studies.

3. The current thesis provided strong evidence that children with SLI had substantially more difficulties with adjectival agreement than with agreement between determiner and noun. This is also compatible with the French study mentioned above but, in Portuguese, the rate of errors with determiners was even lower than in French.

4. Children with SLI reacted differently from typically developing children when faced with a DP containing a novel noun. Their performance when assigning gender to a recently-encountered noun was much poorer than that of the typically developing children.

8.2 Evaluation of investigation

As noted at the end of chapter 7, an overall analysis of the six experiments conducted in the current thesis indicates the experiments assessing production skills fulfilled their aims more appropriately than the experiments exploring input processing skills. As anticipated, testing comprehension skills concerning gender agreement proved a very difficult task. The attempts made in this thesis turned out to require an excessive level of metalinguistic abilities, reducing at least part of their potential to inform us about basic gender agreement processing. The input processing tasks and Experiment 1, which also required a high level of metalinguistic skills, are not, however, uninformative. It is interesting to note, for example, that WM was the only child with SLI who performed well in both the

card-grouping task in Experiment 1 and in the grammatically judgement task in Experiment 2. The same WM also scored extremely high in both the Ravens and WISC, reported in chapter 6. Meanwhile, his performance in the production tasks in Experiments 4 and 5 was poor. WM perhaps represents a case of extreme mismatch, with limitations in basic language skills tapped by production tasks, but able to draw on intact knowledge (the gender of the determiner, evidenced by Experiment 1) together with his cognitive skills (evidenced by nonverbal tests) to succeed on the cognitively more demanding task of categorising nouns according to gender. Although some of the tasks did not fulfill their objectives completely, they can still be seen as a positive outcome of the current thesis. Arguably, these findings constitute additional evidence to the idea that SLI is a valid and informative category if the focus is on basic language skills, those acquired by any typically developing child without any former instructions, as proposed in Part I.

With respect to the production tasks, it is interesting to note that the typically developing children, by and large, performed at ceiling. One exception to this pattern is their performance on the discordant condition in Experiment 5, which was highly predictable and does not, by any means, constitute evidence against the finding that typically developing children have a solid and robust knowledge of the gender system in Portuguese. It is interesting to note that the typically developing children performed consistently at ceiling even though the group is not a consistent group. Recall that children from two different social classes took part in the study. Also, although the typically developing children performed within normal ranges in both non-verbal skills tests (Ravens and WISC) and on MABILIN, there was some within group variation and not many children performed at ceiling[57]. The children with SLI, on the other hand, presented a different performance pattern compared to the typically developing children. In most cases, the children with SLI made a considerable number of mistakes across different tasks and different conditions. Their performance, nonetheless, was not extremely low, which could lead one to question if results are really informative. Such questioning, however, does not seem to hold. Considering that gender is such a core property of Portuguese, easily acquired by typically developing children early in the acquisition process, and that the

[57] While MABILIN has been argued to be a better measure for SLI studies than the TROG, the CELF and the TOLD, it is, given the nature of the task it uses, tapping more skills than the gender production tasks.

typically developing children in both social groups performed at ceiling, the errors produced by the children with SLI, even if the errors were not very numerous and the children were few, demonstrate that the production of gender agreement, particularly adjective agreement, was an area of difficulty for them. It could, therefore, be argued that the production tasks used in this thesis are particularly good in picking up SLI in Portuguese and possibly in other languages which have a gender system.

Importantly, no significant differences between the two social samples of typically developing children who were recruited for the control groups were found in the experiments. The same is true for the results of the test MABILIN. This is indicative that the tasks used in the current thesis were, in general, able to control, at least to a certain extent, for cognitive factors dependent on educational level/achievement. In other words, it seems that the assessments used here (with the exception of the input processing tasks, as stated before) successfully tapped phenomena that are not (at least not heavily) dependent on experience. Such findings suggest that the testing used for this study reached its objective of focusing on basic language skills, contrary to many of the tests often used in the study of SLI.

One question that leads from the pattern of results observed in this thesis refers to the type of deficit that affects the children with SLI, namely, whether these children are delayed in the process of acquiring the gender system of Portuguese (and possibly other languages) or whether their problems consist of a deviance in normal behaviour (deviance is used to refer to a pattern of errors that is not observed in the acquisition process of typically developing children, not even in the early stages). Taken together, the experiments reported here, along with additional evidence from French SLI, indicate that children with SLI from the age of around 6 years old have a relatively solid knowledge of the gender features of nouns in their language, in that determiner agreement presents little problem. Once the gender feature of nouns is stored in their mental lexicon, these children do not seem to have a problem in retrieving such information. The time and amount of exposure that is needed for a child with SLI to acquire the gender of a novel noun is an open question, but it is likely that they need longer exposure than typically developing children, as the data from Experiment 5 seem to suggest. Recall also the results obtained by Name (2002): children as young as two years of age were able to assign gender to novel nouns. Considering that

typical gender acquisition takes place virtually without errors, we might have a case for deviance with respect to determiner and noun agreement. Given the evidence available at present, however, the case is not particularly strong. In contrast, adjective agreement seems to be an area of difficulty for children with SLI and the errors produced by these children do not correspond to the type of data observed for typically developing children of any age. Adjectival agreement is, thus, possibly a stronger case of deviance.

Before moving on to questions about future research, a final note about the low number of children who participated in the current study is due. In spite of the efforts outlined in chapter 6, only six children with SLI were recruited to take part in the experiments. At this stage, we are in a better position to ask why so many children who were potential cases of SLI according to their speech and language therapists (cf. section 6.3.1) performed within normal range on MABILIN and, therefore, were not included in the SLI group. It seems reasonable to suggest that MABILIN is a much more stringent measure for identifying children with SLI than most other tests used in the clinical and research settings. As mentioned before, MABILIN focuses on relevant linguistic structures (unlike the TOLD and the CELF) and it incorporates findings from recent studies with typically developing children in a way that the TROG does not (e.g. the way the two tests assess the comprehension of relatives clauses).

8.3 Questions about future research

Although not conclusive by any means, the results of the current thesis raise important issues regarding future SLI research. The findings suggesting that the children with SLI of the age range studied here do know the gender features of nouns (at least of frequent nouns) but might take longer than typically developing children to acquire those, combined with the findings indicating that adjective agreement is a problem even when accompanying nouns for which the children know the gender (as evidenced by their relatively intact performance on determiner agreement), are compatible with the type of approach to SLI research argued for in chapter 2. Such findings call attention to the need to look at SLI from a wide perspective, one which brings together developmental models and models of adult-stage processing. If the findings of Experiment 5 are replicated, i.e., if future studies provide additional evidence that children with SLI have difficulties in extracting relevant information concerning the gender of novel nouns, it could mean that something similar takes place outside the experimental

setting as well. In other words, it could be that, at an earlier age, when the acquisition of the gender feature of most frequent nouns is in progress, children with SLI need more exposure to relevant input than typically developing children. This, in turn, could result in the production of DP utterances with a gender mismatch between determiner and noun. Indeed, as mentioned before, PE, the youngest child to participate in the current study, made half of the few errors observed in Experiment 3. The same can be said about the younger group of children with SLI who took part in Jakubowicz and Roulet's French study. Most of the incorrect responses in their production task were made by the younger children. Moreover, SLI is a developmental disorder, not an acquired disorder that affects a system which is already stable. By investigating SLI without considering a developmental viewpoint, an important aspect of the puzzle is left out. As stated in 3.2.2, by 'developmental', I do not mean accounts of the state of children's knowledge at various points in time, i.e., descriptions of what children can and cannot do at different ages, but a procedural account of how children go from one stage in the process to another, partly along the lines of the proposal advanced by Karmiloff-Smith and colleagues. The excerpt below, from Christophe (2001), on 'how to study development', quoting Mehler and Christophe illustrates this point: "Consider scientists interested in the problem of physical growth: it is obviously useless to measure each growing child, regardless of the virtues of the yardstick. Eventually the news that children get taller with age reaches the crowds as does the fact that this process tapers off after puberty. Why then, one may ask, should anyone pursue such observations? We do not know … Whether we measure the expansion of the lexicon, memory span, attention span, the ability to solve logical puzzles, the facts are similar: children generally get better with age. In the absence of an *explanation* of the observed phenomena, this kind of study does not contribute data of great value" (p. 259).

8.4 Implications beyond SLI research

In 3.2, I discussed the implications of adopting a tripartite model for the study of language cognition on the basis of Marr's (1982) work on vision cognition. According to this model, language must be approached at three levels: computational, representational/algorithmic and implementational, where the computational level is abstractly characterised in terms of the task to be performed (linguistics), the representational/algorithmic level is described in terms of the steps that must be followed for a task to be carried out

(psycholinguistics) and the implementational level refers to the system in its physical realization (neurological activity) .

The finding that the children with SLI performed differently with respect to determiner/noun agreement and noun/adjective agreement has implications that extend beyond SLI research. This raises questions about the processing of adjective agreement as opposed to determiner agreement. The findings reported in the current thesis are compatible with previous research on adult language processing reviewed in chapter 5, which suggested that different phenomena are involved in agreement within the DP. The data reported here are compatible with the processing model which was sketched in 5.3.1, which, in turn, is compatible with linguistic models that accommodate the idea that different mechanisms are involved in determiner/noun agreement and noun/adjective agreement. For example, the findings of the current thesis can be said to be compatible with the proposal put forward in 4.5.5 which is based on Higginbotham's theta identification theory, but not with the proposal by Magalhães (2004), reviewed in 4.5.3.

8.5 Conclusion

In sum, the current thesis sought to contribute to the debate about what is known as Specific Language Impairment. Two possible scenarios for the field were put forward and an approach to the disorder on the basis of a narrow conception of the term *language* was proposed as the only way SLI could be considered a valid category. An experimental study which attempted to put this approach into practice was carried out. It investigated grammatical gender abilities in Brazilian Portuguese. The theoretical discussion and the experimental findings demonstrated that, by around the age of six, the children with SLI had acquired the gender of frequent nouns, as illustrated by their robust performance on tasks involving determiner and noun agreement. In contrast, adjective agreement posed problems for these children, suggesting that different mechanisms are involved in determiner/noun agreement and noun/adjective agreement. The children with SLI performed poorly when encountering novel nouns, which led us to conjecture that children with SLI need more exposure to acquire gender than typically developing children.

Within the context of SLI studies in Portuguese and, more generally, in Romance languages, the results obtained here provide us with informative data about this little-explored group of languages.

It seems possible to say, therefore, that an approach to SLI under a narrow conception of the term *language* (Scenario 2), as discussed and put forward in Part I of the current thesis, proved to be a productive course of action for the field. It led to the identification of tasks that differentiated core difficulties from difficulties related to inferencing, encyclopaedic knowledge and metalinguistic abilities. That does not imply, by any means, that tests such as the CELF and the TOLD and studies of broader aspects of language and of metalinguistic abilities are not relevant. I hope to have convinced the reader that different phenomena are involved in each case.

REFERENCES

Abels, K. & Neeleman, A. (2009) Universal 20 without the LCA. In. J. Brucart, A. Gavarró, J. Solà (Eds) *Merging Features: Computation, Interpretation, and Acquisition*. Oxford University Press.

Abney, S. (1987) The English Noun Phrase in its sentential aspect, PhD, MIT.

Adani, F. (in press) Rethinking the acquisition of relative clauses in Italian: towards a grammatically-based account. *Journal of Child Language*.

Alexiadou, A. (2001) Adjective syntax and noun raising: word order assymmetries in the DP as the result of adjective distribution. *Studia Linguistica*, Vol. 53, p. 217-248.

Alexiadou, A., Haegeman, L. & Stavrou, M. (2007) *Noun Phrase in the Generative Perspective*. Mouton de Gruyter.

Anderson, R & Souto, S. (2005) The use of articles by monolingual Puerto Rican Spanish-speaking children with specific language impairment. *Applied Psycholinguistics*, Vol. 26, p. 621 – 647.

Antón-Mendez, I., Nicol, J. & Garrett, M. (2002) The relation between gender and number agreement processing. *Syntax*, Vol. 5, p. 1-25.

Aram, D., Morris, R. & Hall, N. (1993) Clinical and Research Congruence in Identifying Children With Specific Language Impairment. *Journal of Speech and Hearing Research*, Vol. 36, p. 580-591

Archibald, L. & Gathercole, S. (2006) Visuospatial Immediate Memory in Specific Language Impairment. *Journal of Speech, Language, and Hearing Research* Vol. 49, p. 265-277

Badecker, W., Miozzo, M. & Zanuttini, R. (1995) The two-stage model of lexical retrieval: evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, Vol. 57, p. 193-216.

Barber, H. & Carreiras, M. (2005) Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience*, Vol. 17, p. 137-153.

Barrett, C. & Kurzban, R. (2006) Modularity in Cognition: Framing the Debate. *Psychological Review*, Vol. 113, p. 628-647.

Bauman, J. (1996) The Materials Writers Newsletter. The Newsletter of the Materials Writers' National Special Interest Group of the Japan Association of Language Teachers, Vol. IV.

Beard, R. (1995). *Lexeme-morpheme base morphology*. Albany: SUNY Albany Press.

Bedore, L. & Leonard, L. (2001) Grammatical Morphology Deficits in Spanish-Speaking Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, Vol. 44, p. 905–924.

Bernstein, J. (1991) DPs in French and Walloon: Evidence for Parametric Variation in Nominal Head Movement. *Probus*, Vol. 3, p. 101-126.

Bernstein, J. B. (1993) *Topics in the Syntax of Nominal Structure across Romance*. Doctoral Thesis. CUNY.

Bishop, D.V.M. (1994) Is Specific Language Impairment a Valid Diagnostic Category? Genetic and Psycholinguistic Evidence. Philosophical Transactions. *Biological Sciences*, Vol. 346, No. 1315, The Acquisition and Dissolution of Language, p. 105-111.

Bishop, D.V.M. (2003) *Test for Reception of Grammar*. Harcourt Assessment.

Bloom, L. (1991) *Language development from two to three*. Cambridge University Press.

Bobaljik (2008). Where's ϕ? Agreement as a post-syntactic operation**.** In D. Harbour, D. Adger, & S. Béjar (Eds) *Phi-Theory: Phi features across interfaces and modules*, Oxford University Press, p. 295-328.

Bridges, A. & Smith, J. (1984) Syntactic comprehension in Down's syndrome children, *British Journal of Psychology*, Vol. 75, p. 187-196.

Brinton, B., Spackman, M., Fujiki, M. & Ricks, J. (2007) What Should Chris Say? The Ability of Children With Specific Language Impairment to Recognize the Need to Dissemble Emotions in Social Situations. *Journal of Speech, Language, and Hearing Research*, Vol. 50, p. 798-811

Burgemeister, B.B., Blum, L.H. & Lorge, I. (1972) *Columbia Mental Maturity Scale*. The Psychological Corporation.

Caramazza, A. & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, Vol. 64, p. 309-343.

Carney, A. (2002) *Syntax: A Generative Introduction*. Blackwell Publishers.

Carrow, E. (1973) *Screening Test for Auditory Comprehension of Language*. Austin, TX: Learning Concepts.

Carstairs-McCarthy. A. (1992) *Current Morphology*. Routledge.

Chomsky, N. (1957) *Syntactic Structures*. The Hague/Paris: Mouton.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1986) *Knowledge of Language, its nature, origin and use*. New York: Praeger.

Chomsky, N. (1995) *The Minimalist Program*. Cambridge, Mass: The MIT Press

Christophe, A. (2001) How to study development. In E. Dupoux (Ed) *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo (1990) *Types of A'-Dependencies*. Cambridge, Mass.: MIT Press.

Cinque, G. (1994) On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.Y. Pollock, L. Rizzi & R. Zanuttini (Eds) *Paths Towards Universal Grammar*. Georgetown University Press.

Cinque, G. (2004) A Phrasal Movement Analysis of the Romanian DP. In A. Minuţ & E. Munteanu (Eds.) *Studia Linguistica et Philologica in Honorem D.Irimia*. Editura Universităţii "Al.I.Cuza", Iaşi, p. 129-142.

Cinque, G. (2005) Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry*, Vol. 36, p. 315-332.

Clahsen, H. (1989) The grammatical characterization of developmental dysphasia. *Linguistics*, Vol. 27, p. 897-920.

Clahsen, H., Bartke, S. & Göllner, S. (1997) Formal features in impaired grammars: a comparison of English and German SLI children. *Journal of Neurolinguistics*, Vol. 10, p. 151-171.

Clahsen, H. & Almazan, M. (1998) Syntax and morphology in Williams syndrome. *Cognition*, Vol. 68, p. 167-198.

Conti-Ramsden, G. & Durkin, K. (2008) Language and Independence in Adolescents With and Without a History of Specific Language Impairment (SLI). *Journal of Speech, Language, and Hearing Research*, Vol 51, p. 70–83.

Corbett, G. (1991) *Gender*. Cambridge. Cambridge University Press.

Corbett, G. (2006) *Agreement*. Cambridge University Press.

Corrêa, L.M.S. (1999) Aquisição da Linguagem: Uma retrospectiva dos Últimos Trinta Anos. *Revista D.E.L.T.A*, Vol. 15, p. 339-383.

Corrêa, L.M.S. (2000a) Categorias Funcionais e Processos Flexionais na Emergência da Linguagem e um Instrumental Psicolingüístico para o Diagnóstico de Manifestações Morfossintáticas do DEL. Projeto FAPERJ.

Corrêa, L.M.S. (2000b). Uma hipótese para a identificação do gênero gramatical com particular referência para o português. Paper presented in the V Encontro Nacional e I Encontro International sobre Aquisição da Linguagem, Porto Alegre, 2-6 Oct.

Corrêa, L.M.S. (2005a) On the processing of gender and number agreement in early language acquisition. Séminaires du Lundi, Laboratoire de Psychologie Expérimentale (LPE), Universite Paris V.

Corrêa, L. M. S. (2005b) The proposal of an instrument for a principled diagnosis of SLI. Poster presented at EUCLDIS 2005.

Corrêa, L. M. S. (2006) Conciliando processamento lingüístico e teoria de língua no estudo da aquisição da linguagem e dos desvios do desenvolvimento lingüístico. In : L. M. S. Corrêa (Ed). *Estudos sobre Aquisição da Linguagem e*

Desvios no Desenvolvimento Lingüístico: Trabalhos oriundos do II Instituto de Inverno em Língua e Cognição.e deficiências da Linguagem. Ed. da PUC-Rio.

Corrêa, L.M.S., Freitas, M.C., Lima, C.M.C. (2003) Crianças com queixas de linguagem e procedimentos usuais de avaliação de habilidades lingüísticas In *Caleidoscópio*: Programa de Pós-Graduação em Lingüística Aplicada, Porto Alegre, Editora da UNISINOS, Vol. 1, p. 43-67.

Corrêa, L.M.S. & Name, M.C (2003) The Processing of Determiner – Noun Agreement and the Identification of the Gender of Nouns in the Early Acquisition of Portuguese. *Journal of Portuguese Linguistics*, Vol. 2, p. 19 – 43.

Corrêa, L.M.S, Almeida, D. & Porto, R. S. (2004) On the representation of Portuguese gender-inflected words in the mental lexicon. *Brain and Language*, Vol. 90, p. 63-73.

Corrêa, L.M.S. & Augusto, M.R.A. (2005) Possible Loci of SLI From a Both Linguistic and Psycholinguistic Perspective. In Abstracts of EUCLDIS 2005, IX EUCLDIS Conference, Royaumont.

Corrêa, L.M.S. & Augusto, M.R.A. (2006) Computação lingüística no processamento on-line: em que medida uma derivação minimalista pode ser incorporada em modelos de processamento? In *Livro de Resumos do XXI Encontro Nacional da ANPOLL*.

Craig, H. (1991) Pragmatic characteristics of the child with specific language impairment: An interactionist perspective. In T. M. Gallagher (Ed). *Pragmatics of Language: Clinical Practice Issues*. Singular: San Diego, CA, p. 11-42.

Craig, H. & Evans, J. (1993) Pragmatics and SLI: Within-Group Variations in Discourse Behaviors. *Journal of Speech and Hearing Research*, Vol 36, p. 777-789.

Crystal, D. (1997) *A Dictionary of Linguistics and Phonetics*. Blackwell Publishers.

Culatta, B. Page, J. & Ellis, J. (1983) Story Retelling as a Communicative Performance Screening Tool. *Language, Speech, and Hearing Services in Schools*, Vol. 14, p. 66-74.

Dell, G. (1986) A Spreading activation theory of retrieval in sentence production. *Psychological Review*, Vol. 93, p.283-321.

Delsing, L-O. (1993) *The internal structure of noun phrases in the Scandinavian languages: A comparative study*. Unpublished Ph.D. dissertation, University of Lund.

Demonte, V. (1999) 'El Adjetivo: Clases y usos. La posición del Adjetivo en el Sintagma Nominal'. In I. Bosque & V. Demonte (Eds) *Gramática descriptiva de la lengua española*, Espasa-Calpe, Madrid, p. 129-215.

Di Domenico, E. (1997) *Per uma teoria del genere grammaticale*. Padova: Unipress.

Dunn, L.M.(1965). *Peabody Picture Vocabulary Test*. Ed. American Guidance Service Inc., Minnesota.

Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, J. (1997) *British Picture Vocabulary Scales*, 2nd edition. Windsor: NFER Nelson.

Ebbels, S. & van der Lely, H. K. J. (2001) Meta-syntactic therapy for children with severe persistent SLI using visual coding. *International Journal of Language and Communication Disorders,* Vol. 36, p. 345-350.

Emonds, J. (2000) *Lexicon and Grammar: The English Syntacticon*. Berlin: Mouton de Gruyter.

Evans, J. (2001) An emergent account of language impairments in children with SLI: implications for assessment and intervention. *Journal of Communication Disorders*, Vol. 34, p. 39-54

Faussart, C. (2000). *Précédence linéaire ou dépendance syntaxique dans l'effet de discordance en français*. Unpublished manuscript, Université Paris 5, France.

Ferrand, L. (2001). Grammatical gender is also on the tip of French tongues. *Current Psychology Letters*, Vol. 5, p. 7-20.

Figueira, R. A. (1996) O erro como dado de eleição nos estudos de aquisição da linguagem. In: De Castro (Org.) *O Método e o Dado no Estudo da Linguagem*. Campinas: Editora da Unicamp.

Figueira, R.A. (2001) Marcas insólitas na aquisição de gênero gramatical: a propriedade reflexiva da linguagem na fala da criança. In: Anais do 5º Econtro Nacional sobre Aquisição da Linguagem e do 1º Encontro Internacional sobre Aquisição da Linguagem. *Letras de Hoje*, Vol. 36, p. 313-320.

Fodor, J. (1983) *The Modularity of Mind*. MIT Press, Cambridge, MA.

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S. & Frauenfelder, U. (2008) The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, Vol 23, p. 329 — 374.

Friederici, A. & Jacobsen, T. (1999) Processing Grammatical Gender During Language Comprehension. *Journal of Psycholinguistic Research*, Vol. 28, p. 467-484.

Fromkin, V. (1997) Some thoughts about the brain/mind/language interface. *Lingua*, Vol. 100, p. 3-27.

Fukuda, S. & Fukuda, S.E. (1994) Developmental language impairment in Japanese: A linguistic investigation. *McGill Working Papers in Linguistics 10*, 150-177.

Gathercole, S.E. & Baddeley, A.D. (1990) Phonological memory deficits in language disordered children: is there a causal connection? *Journal of Memory and Language*, Vol. 29, p. 336-360.

Giusti, G. (2002) The functional structure of noun phrases: A bare phrase structure approach. In G. Cinque (Ed) *Functional Structure in DP and IP*, Oxford University Press, p. 54-90.

Gonzaga, M (2004) The structure of the DP in European Portuguese: evidence from Adjectives and Possessives. *Harvard Working Papers in Linguistics*, Vol. 10.

Gopnik, M. (1990) Feature Blindness: A Case Study. *Language Acquisition*, Vol. 1, p. 139-164.

Graig, H. & Evans, J. (1993) Pragmatics and SLI: within group variations in discourse behaviours. *Journal of Speech and Hearing Research*, Vol. 36, p. 777-789.

Grice, P. (1975) Logic and conversation. In P. Cole and J. L. Morgan (Eds) *Speech Acts*. New York: Academic Press.

Grimshaw, J. (2005) Words and Structures. Leland Stanford Junior University. Center for the Study of Language and Information - Lecture Notes n^{o} 151.

Grodzinsky, Y. (1990) *Theoretical Perspectives on Language Deficits*. Cambridge, MA: MIT Press.

Hagoort, P. & Brown, C. (1999) Gender Electrified: ERP Evidence on the Syntactic Nature of Gender Processing. *Journal of Psycholinguistic Research*, Vol. 28, p. 715-728.

Halle, K. & Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In K. Hale and S. J. Keyser (Eds) *The View from Building 20*, p. 111-176. Cambridge MA: MIT Press.

Hammer, C., Pennock-Roman, M., Rzasa, S. & Tomblin, B. (2002) An Analysis of the Test of Language Development – Primary for Item Bias. *American Journal of Speech and Language Pathology*, Vol. 11, p. 274-284.

Hauser, M., Chomsky, N. & Fitch, T. (2002) The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, Vol. 298, p. 1569-1579.

Haeusler (2005) *A Estrutura Argumental de Verbos na Produção Eliciada de Crianças com Queixas de Linguagem e Manifstações do Déficit Especificamente* Lingüístico. Master's Dissertation. PUC-Rio.

Henzl, V. (1975) Acquisition of grammatical gender in Czech. *Papers and Reports on Child Language Development*, 10, 188-200.

Hermont, A. (2005) *Aquisição de Tempo e Aspecto no Déficit Especificamente Lingüístico*. Doctoral Thesis. UFRJ.

Higginbotham, J. (1985) On Semantics. Linguistic Inquiry, Vol. 16 , p. 547-593.

Jackendoff, R. (1977) *X-bar syntax*. MIT Press.

Jakubowicz, C. & Faussart, C. (1998) Gender agreement in the processing of spoken french. *Journal of Psycholinguistic Research*, Vol. 27, p. 597-617.

Jakubowicz, C., Nash, L., Rigaut, C. & Gérard, C-L. (1998) Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition*, Vol. 7, p. 113-160.

Jakubowicz, C. & Roulet, L. (2007) Narrow syntax or interface deficit? Gender agreement in French SLI. In: J.M. Liceras, H. Zobi & H. Goodluck (Eds) *The role of formal features in second language acquisition*. Lawrence Erlbaum Associates.

Kao, G. & Thompson, J. (2003) Racial and Ethnic Stratification in Educational Achievement and Attainment. *Annu. Rev. Socio*, Vol. 29, p. 417-442.

Karmiloff-Smith, A. (1979) *A functional approach to child language. A study of determiners and reference*. Cambridge: Cambridge University Press.

Karmiloff-Smith. A. (1998) Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, Vol. 10, p. 389-398.

Kaufman, S.S., Kaufman, N.L. (1990) *Kaufman Brief Intelligence Test*. Circle Pines, MN: AGS.

Kayne, R. (1994) *The Antisymmetry of Syntax*. MIT Press.

Kidd, E. & Bavin, E. (2002) English-Speaking Children's Comprehension of Relative Clauses: Evidence for General-Cognitive and Language-Specific Constraints on Development. *Journal of Psycholinguistic Research*, Vol. 31, p. 599-617.

Knittel, M. L. (2005) Some remarks on adjective placement in the French NP. Probus, Vol. 17, p. 185–226.

Laenzlinger, C. (2005) French adjective ordering: perspectives on DP-internal movement types. *Lingua*, Vol. 115, p. 645–689.

Lamarche, J. (1991) Problems for N°-Movement to Num-P. *Probus*, Vol. 3, p. 215-236.

Lee-Williams, C. & Fernald, A. (2007) Young Children Learning Spanish Make Rapid Use of Grammatical Gender in Spoken Word Recognition. *Psychological Science*, Vol. 18, p. 193-198.

LeFever, G., Arcona, A. & Antonuccio, D. (2003) ADHD among American Schoolchildren: Evidence of Overdiagnosis and Overuse of Medication. *The Scientific Review of Mental Health Practice*, Vol. 2.

Leonard, L. (1998) *Children with Specific Language Impairment*. MIT Press.

Levelt, W.J.M. (1989) *Speaking: From Intention to Articulation*. MIT Press.

Macacchero, N. (2004) *Aquisição das categorias funcionais tempo e aspecto no défict especificamente lingüístico*. Master's Dissertation. UFRJ.
Magalhães, T. (2004) A Valoração de Traços de Concordância dentro do DP. *Revista D.E.L.T.A.*, Vol. 20, p. 149-170.

Marr (1982) *Vision: A computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman.

Marshall, C. & van der Lely, H. (2006) A challenge to current models of past tense inflection: the impact of phonotactics. *Cognition*, Vol. 100, p. 302-320.

Marshall, C. & van der Lely, H. (2008). Recognition of gated verbs by children with Grammatical-Specific Language Impairment: Effects of inflection and frequency. *Journal of Neurolinguistics*, Vol. 22, p.433-451.

Marshall, C. (2004) *The morpho-phonological interface in Specific Language Impairment*. PhD Thesis. University College London.

Marshall, E. (2000) Duke Study Faults Overuse Of Stimulants for Children. *Science*, New Series, Vol. 289, p. 721.

McGregor, K. K. & Leonard, L. B. (1995) Intervention for word-finding deficits in children. In: Fey, M. E.; Windsor, J.; Warren, S. F. (Ed.). *Language intervention: preschool through the elementary years*. Baltimore: Paul Brookes Publishing, p. 85-105.

Mehler, J. & Dupoux, E. (1990) *Nascer Humano*. Lisboa: Instituto Piaget.
Miozzo, M. & Caramazza, A. (1999) The Selection of Determiners in Noun Phrase Production. *Journal of Experimental Psychology*, Vol. 25, p. 907-922.

Moura Castro, C. & Verdisco, A. (2002) Lessons from a Diagonal Dialogue: HIP Asia and Latin America. In C. Moura Castro & A. Verdisco (Eds) *Making Education Work: Latin American Ideas and Asian Results*. Inter-American Development Bank.

Name, M. C. (2002*) Habilidades perceptuais e lingüísticas no processo de aquisição do sistema de gênero no português*. Doctoral Thesis. PUC-Rio.

Name, M. C. L, Corrêa, L. M. S. (2001) Young children sensitivity to determiners and the identification of the gender system in Portuguese. In: GALA CONFERENCE 2001 – Generative Approaches to Language Acquisition, 2001, Palmela. Gala. p.180-187.

Neeleman, A. & van de Koot, H. (2009) Theoretical Validity and Psychological Reality of the Grammatical Code. In: M. Everaert, T. Lentz, H. De Mulder, Nilsen, & A. Zondervan,A. (Ed.) *The Linguistics Enterprise: From Knowledge of Language to knowledge in Linguistics (to appear)*. Linguistik Aktuell/Linguistics Today series. John Benjamins.

Newcomer, P., & Hammill, D. (1991). *Test of language development–Primary: 2*. Austin, TX: Pro-Ed.

Nilsen, O. (2003) *Eliminating Positions*. Ph.D Thesis. Utrecht University.

Pérez-Pereira, M. (1991) The acquisition of gender: what Spanish children tell us. *Journal of Child Language*, Vol. 18, p. 571-590.

Perovic, A. (2003) *Knowledge of Binding in Down Syndrome; Evidence from English and Serbo-Croation*. PhD Thesis. University College London.

Pesetsky, D. & Torrego, E. (2004) Probes, Goals and Syntactic Categories. *Proceedings of the 7th annual Tokyo Conference on Psycholinguistics* (Keio University, Japan).

Phillips, C. (1990) Order and Structure. PhD Theis. MIT.

Phillips, C. and Wagers, M. (2007) Relating Structure and Time in Linguistics and Psycholinguistics. In G. Gaskell (Ed.) *Oxford Handbook of Psycholinguistics*. Oxford University Press.

Picallo, M. C. (1991) 'Nominals and nominalization in Catalan' *Probus*, Vol. 3, p. 279-316.

Polišenská, K. (2006) *Gender Acquisition in Child L1 and 2L1 Czech. An Experimental Study*. MA Thesis. University of Amsterdam.

Pysz, A. (2006) The structural location of adnominal adjectives: Prospects for Old English. SKASE *Journal of Theoretical Linguistics*, Vol. 3, p. 59-85.

Radford, A. (1997) *Syntax: A Minimalist Introduction*. Cambridge University Press.

Raven, J., Raven, J.C., & Court, J.H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Section 1: General Overview. San Antonio, TX: Harcourt Assessment.

Rice, M.L., Wexler, K. & Cleave, P.L. (1995) Specific Language Impairment as a period of Extended Optional Infinitive. *Journal of Speech and Hearing Research*, Vol. 38, p. 850-863.

Rice, M. & Wexler, K. (1996). Toward Tense as a Clinical Marker of Specific Language Impairment in English-speaking Children, *Journal of Speech and Hearing Research*, Vol. 39, p. 850-863.

Rice, M., Wexler, K. & Hershberger, S. (1998) Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, Vol. 41, p. 1412-1431.

Ritter, E. (1993) Where's gender? *Linguistic Inquiry*, Vol. 24, p. 795-803.

Rodrigues, E. (2006) *Processamento da concordância de número entre sujeito e verbo na produção de sentenças*. Doctoral thesis. PUC-Rio.

Roehrs, D. (2006) *The Morpho-Syntax of the Germanic Noun Phrase: Determiners Move into the Determiner Phrase*. Doctoral Thesis. Indiana University.

Roulet-Amiot, L. & Jakubowicz, C. (2006) Production and perception of gender agreement in French SLI. Advances in Speech–Language Pathology, Vo. 8, p. 335–346.

Sanchez, M. & Grinstead, J. (2004) Nominal Number & Gender Agreement in Child Spanish SLI. Poster presented at Generative Approaches to Language Acquisition North America, University of Hawai'i, Manoa.

Schmitt, C. (1996) *Aspect and the Syntax of Noun Phrases*. Doctoral Thesis. University of Maryland.

Schmitt, C. & Munn, A. (1999) Against the nominal mapping parameter: bare nouns in Brazilian Portuguese. Proceedings of North East Linguistc Society (NELS) 29.

Schriefers, H. & Jescheniak, J. (1999) Representation and Processing of Grammatical Gender in Language Production: A Review. *Journal of Psycholinguistic Research*, Vol. 28, p. 575-600.

Semel, E., Wiig, E. & Secord, W. (1987) *Clinical Evaluation of Language Fundamentals*. Psychological Corporation.

Shady, M. (1996) *Infants' sensitivity to function morphemes*. PhD Dissertation. University of Buffalo.

Shafer, V., Shucard, D., Shucard, J. & Gerken, L. (1998) An Electrophysiological Study of Infants' Sensitivity to the Sound Patterns of English Speech. *Journal of Speech, Language and Hearing Research,* Vol. 41, p. 874-886.

Silveira, M. (2002) *O Déficit Especificamente Lingüístico (DEL) e uma avaliação preliminar de sua manifestação em crianças falantes de português*. Master's Dissertation. PUC-Rio.

Sirin, S. (2005) Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, Vol. 75, p. 417-453.

Slobin, D.I. (1985) Crosslinguistic Evidence for the Language–Making Capacity. In: D.I. Slobin (Ed.) *The Crosslinguistic Study of Language Acquisition*, Vol.2: Theoretical Issues. Hillsdale, N. J.: Lawrence Erlbaum.

Smith, N. & Tsimpli, I-M. (1995) *The Mind of a Savant*. Blackwell.

Sproat, R. (1985) *On deriving the lexicon.* PhD dissertation. MIT.

Sproat, R. & Shih, C.(1988) Prenominal Adjectival Ordering in English and Mandarin, Proceedings of the 18th Annual Meeting of the North East Linguistics Society, p. 465-489.

Tallal, P. & Piercy, M. (1973) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, Vol. 241, p. 468-469.

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W.M., Merzenich, M. M., (1996) Language Comprehension in Language-Learning Impaired Children Improved with Acoustically Modified Speech, *Science*, V. 271, p. 81-84.

Thomas, M. & Karmiloff-Smith, A. (2005) Can Developmental Disorders Reveal the Component Parts of the Human Language Faculty? *Language Learning and Development*, Vol. 1, p. 65-92.

Ticio, M. E. (2003) *On the Structure of DPs*. Doctoral Thesis. University of Connecticut.

Tomblin, B., Records, N. & Zhang, X. (1996) A System for the Diagnosis of Specific Language Impairment in Kindergarten Children. *Journal of Speech and Hearing Research*, Vol. 39, p. 1284-1294.

Tomblin. B., Records, N., Buckwalter, P., Zhang, X. Smith, E. & O'Brien, M. (1997) Prevalence of Specific Language Impairment in Kindergarten Children. *Journal of Speech, Language and Hearing Research*, Vol. 40, p. 1245–1260

van Berkum, J. (1997) Syntactic processes in speech production: the retrieval of grammatical gender. *Cognition*, Vol. 64, p. 115 – 152.

van der Lely, H. K. J. (1997). Narrative discourse in Grammatical specific language impaired children: A modular language deficit? *Journal of Child Language*, Vol. 24, p. 221-256.

van der Lely, H. (1998) SLI in children: movement, economy, and deficits in the computational-syntactic system. *Language Acquisition*, Vol. 7, p. 161-192.

van der Lely, H. K. J. (2004). Evidence for and implications of a domain-specific grammatical deficit. In L. Jenkins (Ed.) *The genetics of Language. Linguistic Variations series*. Series editors J. Rooryck and P. Pica. Elsevier, Oxford: Chapter 6, p 117-145.

van der Lely, H. K. J. (2005b). Grammatical-SLI and the computational grammatical complexity hypothesis. *Revue Frequences*, Vol 17, p. 13-20.

van der Lely H K. J. (2005a). Domain-specific cognitive systems: Insight from Grammatical specific language impairment. *Trends in Cognitive Sciences*, Vol. 9, p. 53-59

van der Lely, H. K. J., & Stollwerck, L. (1997). Binding theory and specifically language impaired children. *Cognition*, Vol. 62, p. 245-290.

van der Lely, H. K. J. & Battell, J. (2003). Wh-movement in children with grammatical SLI: a test of the RDDR Hypothesis. *Language*, Vol. 79, p. 153-181.

van der Lely, H. K. J., Rosen, S. & Adlard, A. (2004).Grammatical language impairment and the specificity of cognitive domains: Relations between auditory and language abilities. *Cognition*, Vol. 94, p. 167-183.

Vane, J. (1975) Vane Evaluation of Language Scale. *Archives of Behavioral Sciences*, Vol. 49, p. 3-33.

Wechsler, D. (2004) *Wechsler Intelligence Scale for Children*. Harcourt Assessment.

Wiig, E. & Semel, E. (1980) *Language assessment and intervention for the learning disabled*. Columbus : C.E. Merrill Pub. Co.

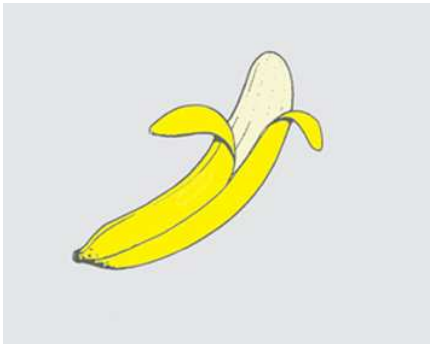
APPENDIX

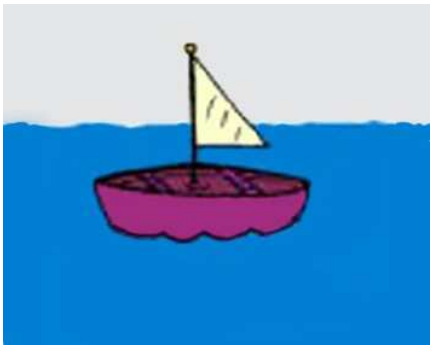
EXPERIMENT 1 – Basket task

LIST OF WORDS:

1. mochila	backpack
2. colher	spoon
3. olho	eye
4. garfo	fork
5. pente	comb
6. martelo	hammer
7. ponte	bridge
8. espelho	mirror
9. raquete	racket
10. cama	bed
11. nariz	nose
12. mão	hand
13. bola	ball
14. barco	boat
15. nuvem	cloud
16. mesa	table
17. casa	house
18. pão	bread
19. televisão	television
20. osso	bone
21. morango	strawberry
22. dedo	finger
23. banana	banana
24. árvore	tree
25. balde	bucket
26. queijo	cheese
27. sol	sun
28. coração	heart
29. cruz	cross
30. trem	train
31. chave	key
32. flor	flower
33. sorvete	icecream
34. faca	knife
35. pipa	kite
36. tambor	drum
37. ovo	egg
38. meia	sock
39. laranja	orange
40. telefone	telephone

EXAMPLES OF CARDS:





EXPERIMENT 2 – Grammaticality judgment task

LIST OF UTTERANCES:

1. a casa	the$_{fem}$ house$_{fem}$	33.o martelo	the$_{masc}$ hammer$_{masc}$
2. a nariz	the$_{fem}$ nose$_{masc}$	34.a meia	the$_{fem}$ sock$_{fem}$
3. o ovo	the$_{masc}$ egg$_{masc}$	35.o nuvem	the$_{masc}$ cloud$_{fem}$
4. a laranja	the$_{fem}$ orange$_{fem}$	36.a faca	the$_{fem}$ knife$_{fem}$
5. o televisao	the$_{masc}$ television$_{fem}$	37.a sorvete	the$_{fem}$ icecream$_{masc}$
6. a banana	the$_{fem}$ banana$_{fem}$	38.o ponte	the$_{masc}$ bridge$_{fem}$
7. o queijo	the$_{masc}$ cheese$_{masc}$	39.o morango	the$_{masc}$ strawberry$_{masc}$
8. a pente	the$_{fem}$ comb$_{masc}$	40.a telefone	the$_{fem}$ telephone$_{masc}$
9. o olho	the$_{masc}$ eye$_{masc}$	41.o trem	the$_{masc}$ train$_{masc}$
10. o flor	the$_{masc}$ flower$_{fem}$	42.a garfo	the$_{fem}$ fork$_{masc}$
11. o colher	the$_{masc}$ spoon$_{fem}$	43.a mao	the$_{fem}$ hand$_{fem}$
12. a sol	the$_{fem}$ sun$_{masc}$	44.a osso	the$_{fem}$ bone$_{masc}$
13. o espelho	the$_{masc}$ mirror$_{masc}$	45.o casa	the$_{masc}$ house$_{fem}$
14. a bola	the$_{fem}$ ball$_{fem}$	46.o sol	the$_{masc}$ sun$_{masc}$
15. o mao	the$_{masc}$ hand$_{fem}$	47.a ponte	the$_{fem}$ bridge$_{fem}$
16. o chave	the$_{masc}$ key$_{fem}$	48.o laranja	the$_{masc}$ orange$_{fem}$
17. a pao	the$_{fem}$ bread$_{masc}$	49.a raquete	the$_{fem}$ racket$_{fem}$
18. o dedo	the$_{masc}$ finger$_{masc}$	50.a chave	the$_{fem}$ key$_{fem}$
19. o cruz	the$_{masc}$ cross$_{fem}$	51.a flor	the$_{fem}$ flower$_{fem}$
20. a cama	the$_{fem}$ bed$_{fem}$	52.o pao	the$_{masc}$ bread$_{masc}$
21. a balde	the$_{fem}$ bucket$_{masc}$	53.a nuvem	the$_{fem}$ cloud$_{fem}$
22. a mochila	the$_{fem}$ backpack$_{fem}$	54.a martelo	the$_{fem}$ hammer$_{masc}$
23. a coracao	the$_{fem}$ heart$_{masc}$	55.o cama	the$_{masc}$ bed$_{fem}$
24. o garfo	the$_{masc}$ fork$_{masc}$	56.o sorvete	the$_{masc}$ icecream$_{masc}$
25. o osso	the$_{masc}$ bone$_{masc}$	57.a espelho	the$_{fem}$ mirror$_{masc}$
26. a tambor	the$_{fem}$ drum$_{masc}$	58.a morango	the$_{fem}$ strawberry$_{masc}$
27. o barco	the$_{masc}$ boat$_{masc}$	59.a arvore	the$_{fem}$ tree$_{fem}$
28. a trem	the$_{fem}$ train$_{masc}$	60.o pipa	the$_{masc}$ kite$_{fem}$
29. o raquete	the$_{masc}$ racket$_{fem}$	61.o pente	the$_{masc}$ comb$_{masc}$
30. a mesa	the$_{fem}$ table$_{fem}$	62.a cruz	the$_{fem}$ cross$_{fem}$
31. o arvore	the$_{masc}$ tree$_{fem}$	63.o meia	the$_{masc}$ sock$_{fem}$
32. a pipa	the$_{fem}$ kite$_{fem}$	64.o faca	the$_{masc}$ knife$_{fem}$

LIST OF UTTERANCES (CONT):

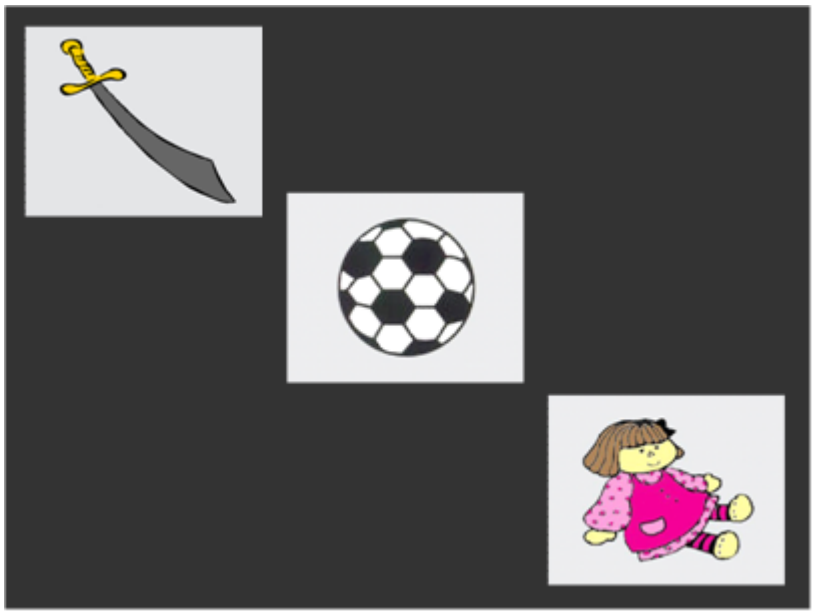
65.	a olho	the_{fem} eye_{masc}
66.	a televisao	the_{fem} $television_{fem}$
67.	o tambor	the_{masc} $drum_{masc}$
68.	a ovo	the_{fem} egg_{masc}
69.	a colher	the_{fem} $spoon_{fem}$
70.	a queijo	the_{fem} $cheese_{masc}$
71.	o banana	the_{masc} $banana_{fem}$
72.	o nariz	the_{masc} $nose_{masc}$
73.	o mochila	the_{masc} $backpack_{fem}$
74.	o mesa	the_{masc} $table_{fem}$
75.	o coracao	the_{masc} $heart_{masc}$
76.	o telefone	the_{masc} $telephone_{masc}$
77.	o bola	the_{masc} $ball_{fem}$
78.	a dedo	the_{fem} $finger_{masc}$
79.	o balde	the_{masc} $bucket_{masc}$
80.	a barco	the_{fem} $boat_{masc}$

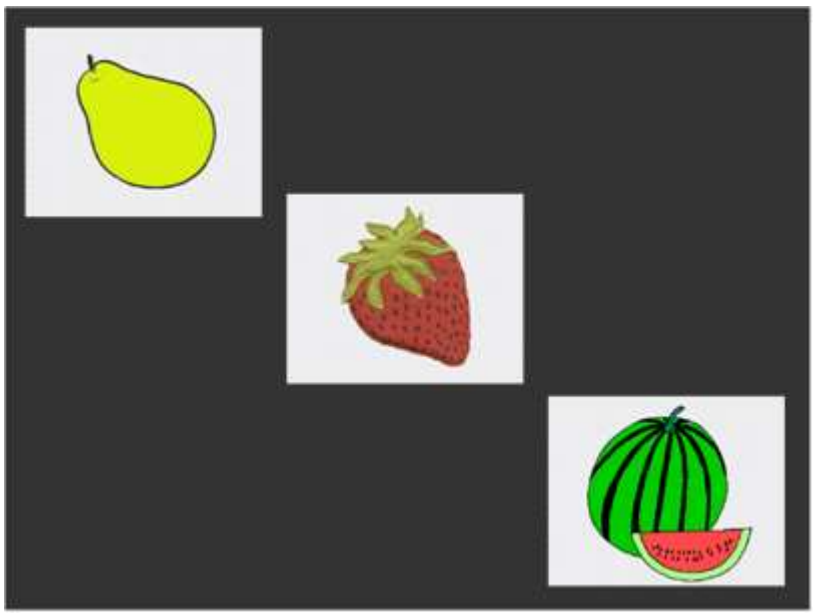
EXPERIMENT 3 – Elicited production task (determiner + noun)

LIST OF WORDS:

Same as Experiment 1

EXAMPLES OF POWERPOINT SLIDES:



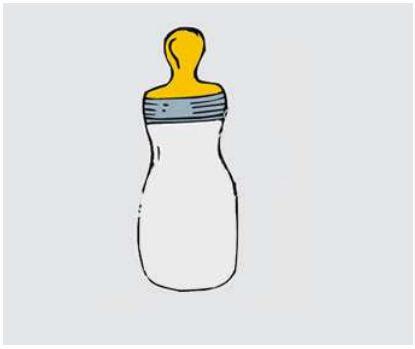


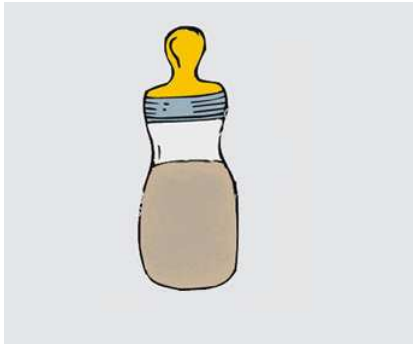
EXPERIMENT 4 – Elicited production task (determiner + noun + adjective)

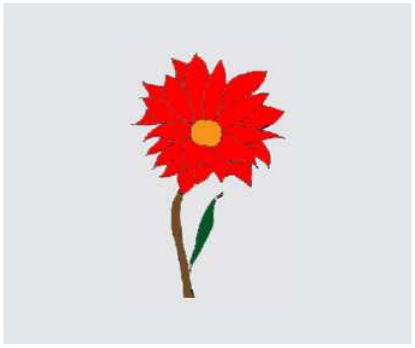
LIST OF TARGET SENTENCES:

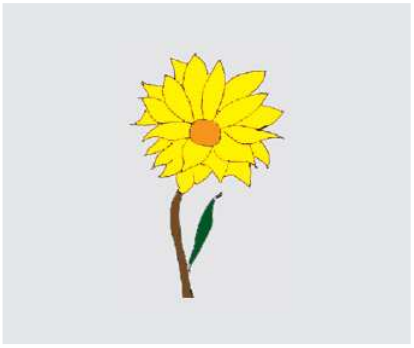
1. Um guarda-chuva preto	A black umbrella
Um guarda-chuva vermelho	A red umbrella
2. Um dado amarelo	A yellow dice
Um dado branco	A white dice
3. Uma nuvem branca	A white cloud
Uma nuvem preta	A black cloud
4. Uma janela aberta	An open window
Uma janela fechada	A closed window
5. Um chapéu branco	A white hat
Um chapéu preto	A black hat
6. Um balanço preto	A black swing
Um balanço vermelho	A red swing
7. Uma flor amarela	A white flower
Uma flor vermelha	A red flower
8. Um piano branco	A white piano
Um piano preto	A black piano
9. Uma mamadeira cheia	A full baby bottle
Uma mamadeira vazia	An empty baby bottle
10. Uma mesa quadrada	A square table
Uma mesa redonda	A round table
11. Um livro aberto	An open book
Um livro fechado	A closed book
12. Uma televisão ligada	A television on
Uma televisão desligada	A television off
13. Uma mão limpa	A clean hand
Uma mão suja	A dirty hand
14. Um foguete amarelo	A yellow rocket
Um foguete branco	A white rocket
15. Um telefone amarelo	A yellow telephone
Um telefone preto	A black telephone
16. Uma mala preta	A black suitciase
Uma mala vermelha	A red suitcase
17. Um pão quadrado	A square loaf of bread
Um pão redondo	A round loaf of bread
18. Um martelo amarelo	A yellow hammer
Um martelo vermelho	A white hammer
19. Uma chave amarela	A yelllow key
Uma chave branca	A white key
20. Uma bandeira amarela	A yellow flag
Uma bandeira vermelha	A red flag

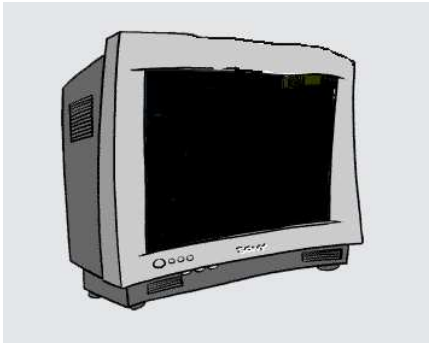
EXAMPLES OF PICTURES:











EXPERIMENT 5 – Elicited production (novel nouns)

LIST OF STIMULI (INANIMATE NOUNS):

1. Aqui tem uma teka e aqui tem outra teka. As tekas estão na rua. Um carro atropelou uma teka. Que teka o carro atropelou?
Here there is a_{fem} teka and here there is $another_{fem}$ teka. The_{fem} tekas are in the street. A car ran over a_{fem} teka. Which teka did the car run over?

2. Aqui tem um bida e aqui tem outro bida. Os bidas estão na escada. Um bida caiu da escada. Que bida caiu da escada?
Here there is a_{masc} bida e here there is $another_{masc}$ bida. The_{masc} bidas are on the stair. A_{masc} bida fell off the stairs. Which bida fell off the stairs?

3. Aqui tem um bafe e aqui tem outro bafe. Os bafes estão em cima da árvore. O pássaro pegou um bafe. Que bafe o pássaro pegou?
Here there is a_{masc} bafe and here there is $another_{masc}$ bafe. The_{masc} bafes are up on the tree. The bird grabbed a_{masc} bafe. Which bafe did the bird grab?

4. Aqui tem uma tuco e aqui tem outra tuco. As tucos estão na caixa. A menina pegou uma tuco. Que tuco a menina pegou?
Here there is a_{fem} tuco e here there is $another_{fem}$ tuco. The_{fem} tucos are in the box. The girl grabbed a_{fem} tuco. Which tuco did the girl grab?

5. Aqui tem um dabo e aqui tem outro dabo. Os dabos estão dentro do armario. Um dabo sumiu. Que dabo sumiu?
Here there is a_{masc} dabo and here there is $another_{masc}$ dabo. The_{masc} dabos are in the closet. A_{masc} dabo disappeared. Which dabo disappear?

6. Aqui tem uma mile e aqui tem outra mile. As miles estão na piscina. Uma mile saiu da piscina. Que mile saiu da piscina?
Here there is a_{fem} mile and here there is $another_{fem}$ mile. The_{fem} miles are in the swimming pool. A_{fem} mile left the pool. Which mile left the pool?

7. Aqui tem um tobe e aqui tem outro tobe. Os tobes estão na praia. Um tobe caiu na agua. Que tobe caiu na agua?
Here there is a_{masc} tobe and here there is $another_{masc}$ tobe. The_{masc} tobes are on the beach. A_{masc} tobe went to the water. Which tobe go to the water?

8. Aqui tem um puco e aqui tem outro puco. Os pucos estão na grama. O menino pegou um puco. Que puco o menino pegou?
Here there is a_{masc} puco and here there is $another_{masc}$ puco. The_{masc} pucos are on the grass. The boy grabbed a_{masc} puco. Which puco did the boy grab?

9. Aqui tem uma lalo e aqui tem outra lalo. As lalos estão no sofa. O gato lambeu uma lalo. Que lalo o gato lambeu?
Here there is a_{fem} lalo and here there is $another_{fem}$ lalo. The_{fem} lalos are on the sofa. The cat licked a_{fem} lalo. Which lalo did the cat lick?

10. Aqui tem uma poca e aqui tem outra poca. As pocas estao em cima da cama. Uma poca caiu no chao. Que poca caiu no chao?
Here there is a_{fem} poca and here there is $another_{fem}$ poca. The_{fem} pocas are on the bed. A_{fem} poca fell on the floor. Which poca fell on the floor?

11. Aqui tem um depa e aqui tem outro depa. Os depas estão na grama. Um depa foi para o balde. Que depa foi para o balde?
Here there is a_{masc} depa and here there is $another_{masc}$ depa. The_{masc} depas are on the grass. A_{masc} depa went inside the bucket. Which depa went inside the bucket.

12. Aqui tem uma dobe e aqui tem outra dobe. As dobes estão na rua. Uma dobe foi para a lixeira. Que dobe foi para a lixeira?
Here there is a_{fem} dobe and here there is $another_{fem}$ dobe. The_{fem} dobes are on the street. A_{fem} dobe went in the garbage. Which dobe went in the garbage?

LIST OF STIMULI (ANIMATE NOUNS):

1. Aqui tem um paco e aqui tem outro paco. Os pacos estão no jardim. Um paco subiu na pedra. Que paco subiu na pedra?
Here there is a_{masc} paco and here there $another_{masc}$ paco. The_{masc} pacos are in the garden. A_{masc} paco climbed on the rock. Which paco climbed on the rock?

2. Aqui tem um fole e aqui tem outro fole. Os foles estão na fazenda. Um fole pegou uma maçã. Que fole pegou uma maçã?
Here there is a_{masc} fole and here there is $another_{masc}$ fole. The_{masc} foles are in the farm. A_{masc} fole grabbed an apple. Which fole grabbed an apple?

3. Aqui tem uma peta e aqui tem outra peta. As petas estão nadando no mar. Uma peta se escondeu atrás da planta. Que peta se escondeu?
Here there is a_{fem} peta and here there is $another_{fem}$ peta. The_{fem} petas are swimming in the sea. A_{fem} peta hid behind the plant. Which peta hid behind the plant?

4. Aqui tem uma tile e aqui tem outra tile. As tiles estão tomando sol na piscina. Uma tile caiu na piscina. Que tile caiu na piscina?
Here there is a_{fem} tile and here there is $another_{fem}$ tile. The_{fem} tiles are sun bathing in the pool. A_{fem} tile went in the pool. Which tile went in the pool?

5. Aqui tem um diba e aqui tem outro diba. Os dibas estão comendo. Um diba acabou de comer. Que diba acabou de comer?
Here there is a_{masc} diba and here there is $another_{masc}$ diba. The_{masc} dibas are eating. A_{masc} diba finished eating. Which diba finished eating?

6. Aqui tem uma bilo e aqui tem outra bilo. As bilos estão andando de bicicleta. Uma bilo caiu no chão. Que bilo caiu no chão?
Here there is a_{fem} bilo and here there is $another_{fem}$ bilo. The_{fem} bilos are riding their bikes. A_{fem} bilo fell on the floor. Which bilo fell on the floor?

7. Aqui tem uma lole e aqui tem outra lole. As loles estão no rio. Uma lole subiu no barco. Que lole subiu no barco?
Here there is a_{fem} lole and here there is $another_{fem}$ lole. The_{fem} loles are in the river. A_{fem} lole went on the boat. Which lole went on the boat?

8. Aqui tem um keko e aqui tem outro keko. Os kekos estão no supermercado. Um keko subiu no carrinho. Que keko subiu no carrinho?
Here there is a_{masc} keko and here there is $another_{masc}$ keko. The_{masc} kekos are in the supermarket. A_{masc} keko climbed in the shopping cart. Which keko climbed in the shopping cart?

9. Aqui tem um bado e aqui tem outro bado. Os bados estão brincando no parquinho. Um bado desceu no escorrega. Que bado desceu no escorrega?
Here there is a_{masc} bado and here there is $another_{masc}$ bado. The_{masc} bados are playing in the playground. A_{masc} bado went down the slide. Which bado went down the slide?

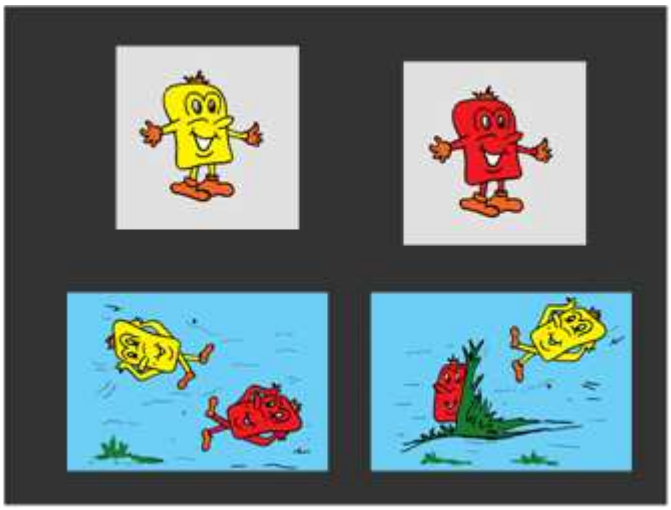
10. Aqui tem uma tule e aqui tem outra tule. As tules estão no jardim. Uma tule pegou uma flor. Que tule pegou a flor?
Here there is a_{fem} tule and here there is $another_{fem}$ tule. The_{fem} tules are in the garden. A_{fem} tule grabbed a flower. Which tule grabbed a flower?

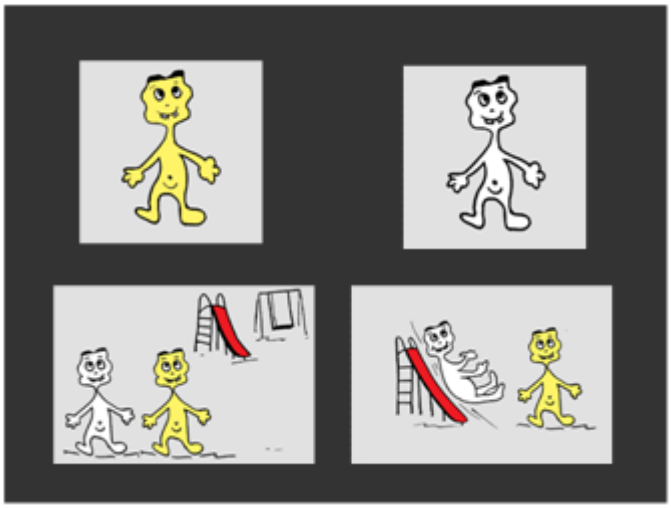
11. Aqui tem um beke e aqui tem outro beke. Os bekes estão brincando no skate. Um beke caiu do skate. Que beke caiu do skate?

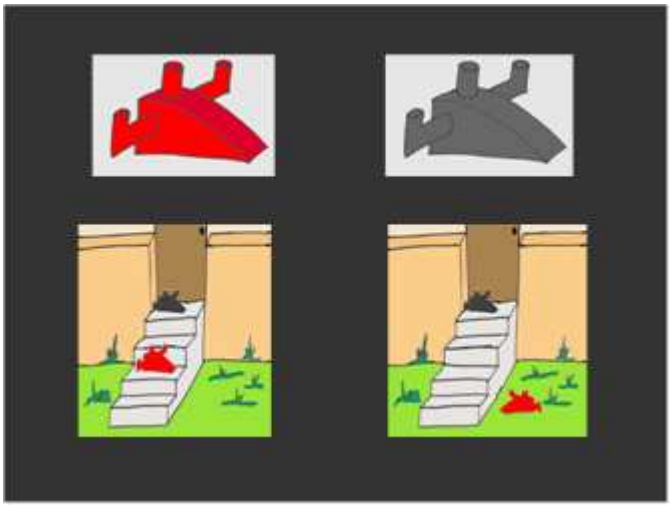
Here there is a_{masc} beke and here there is $another_{masc}$ beke. The_{masc} bekes are playing with their skates. A_{masc} beke fell off the skate. Which beke fell off the skate?

12. Aqui tem uma kebo e aqui tem outra kebo. As kebos estão na rua. Uma kebo caiu no buraco. Que kebo caiu no buraco?
Here there is a_{fem} kebo and here there is $another_{fem}$ kebo. The_{fem} kebos are in the street. A_{fem} kebo fell in the hole. Which kebo fell in the hole?

EXAMPLES OF POWERPOINT SLIDES:







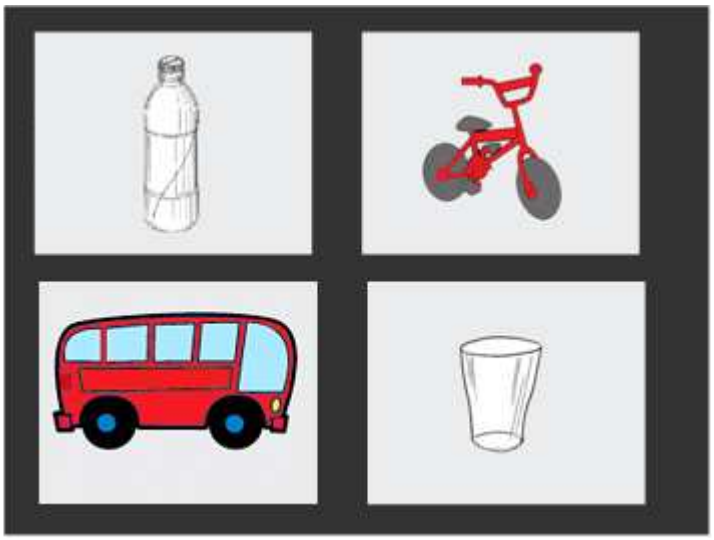
EXPERIMENT 6 – Input processing task ('guess what it is')

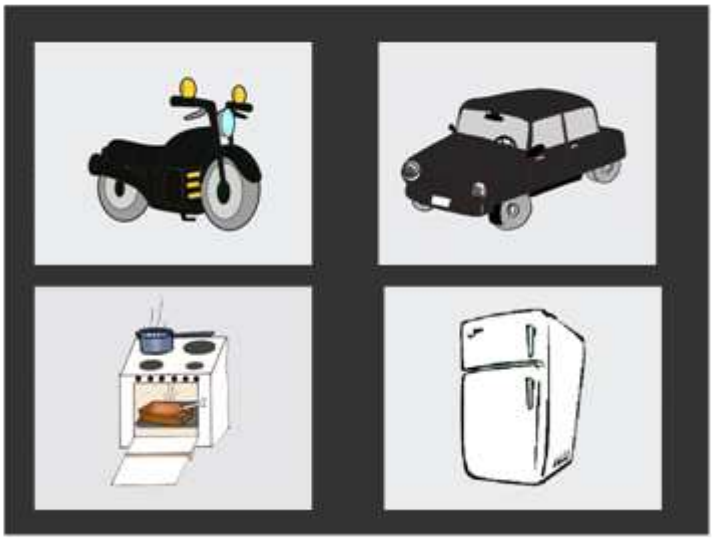
LIST OF WORDS:

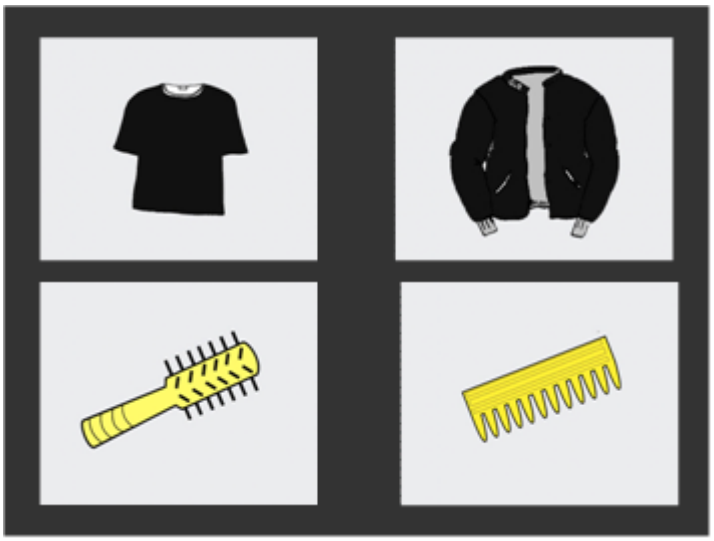
PRIMEIRA RODADA	**FIRST ROUND**
1. LARANJA X BOLO	1. ORANGE X CAKE
A gente come ela	We eat it
É redonda	It's round
2. VESTIDO X SAIA	2. DRESS X SKIRT
Só mulher usa ele	Only women wear it
É amarelo	It's yellow
SEGUNDA RODADA	**SECOND ROUND**
1. PULSEIRA x RELÓGIO	1. BRACELET X WATCH
As pessoas põem ela no braço	People wear it around their arm
É vermelha	It's red
2. PANELA X PRATO	2. PAN X PLATE
A gente põe comida dentro dela	We put food in it
É redonda	It's round
TERCEIRA RODADA	**THIRD ROUND**
1. MOTO X CARRO	1. MOTORCYCLE X CAR
A gente usa ela para ir de um lugar para outro	We use it for going from one place to the other
É preta	It's black
2. FOGÂO X GELADEIRA	2. STOVE X FRIDGE
A gente encontra ele na cozinha	We find it in the kitchen
É branco	It's white
QUARTA RODADA	**FOURTH ROUND**
1. LÁPIS X CANETA	1. PENCIL X PEN
As pessoas usam ele para escrever	People use it to write
É vermelho	It's red
2. QUEIJO X BANANA	2. CHEESE X BANANA
A gente come ele	We eat it
É amarelo	It's yellow
QUINTA RODADA	**FIFTH ROUND**
1. COLHER X GARFO	1. SPOON X FORK
A gente usa ela quando come	We use it when we eat
É branca	It's white
2. TELEVISÂO X RÁDIO	2. TELEVISION X RADIO
A gente liga ela na tomada	We plug it into the socket
É preta	It's black
SEXTA RODADA	**SIXTH ROUND**
1. COPO X GARRAFA	1. CUP X BOTTLE
A gente usa ele para botar água	We use it to put water
É redondo	It's round
2. ÔNIBUS X BICICLETA	2. BUS X BICYCLE
A gente usa ele para ir de um lugar para o outro	We use it for going from one place to the other

É vermelho	It's red
SÉTIMA RODADA	**SEVENTH ROUND**
1. MEIA X SAPATO	1. SOCK X SHOE
A gente usa ela no pé	We wear it on our foot
É amarela	It's yellow
2. MILHO X CENOURA	2. CORN X CARROT
A gente come ele	We eat it
É comprido	It's long
OITAVA RODADA	**EIGHTH**
1. CHAVE X CADEADO	1. KEY X LOCK
A gente usa ela para trancar a porta	We use it for locking the door
É amarela	It's yellow
2. SOFÁ X CADEIRA	2. SOFA X CHAIR
A gente senta nele	We sit on it
É vermelho	It's red
NONA RODADA	**NINTH ROUND**
1. CAMISETA X CASACO	1. T-SHIRT X COAT
A gente veste ela	We wear it
É preta	It's black
2. PENTE X ESCOVA	2. COMB X BRUSH
A gente usa ele quando arruma o cabelo	We use it when doing our hair
É amarelo	It's yellow
DÉCIMA RODADA	**TENTH ROUND**
1. ÁRVORE X PRÉDIO	1. TREE X BUILDING
A gente encontra ela na rua	We find it in the street
É alta	It's tall
2. CADERNO X MOCHILA	2. NOTEBOOK X BACKPACK
A gente leva ele pra escola	We bring it to school
É vermelho	It's red

EXAMPLES OF PICTURES:







MABILIN MODULE 1

LIST OF STIMULI:

1. O sapo
The frog
2. O urso
The bear
3. A bailarina
The ballerina
4. Menino de botas
Boy wearing boots
5. Menina que pula corda
Girl who is rope skipping
6. O cachorro roeu o osso
The dog gnawed on a bone
7. A girafa machucou o urso
The giraffe hurt the bear
8. Quem a bailarina penteou
Who did the ballerina comb?
9. A mãe da gatinha lambeu ela
The mother of the little cat licked her
10. A tartaruga que chamou o macaco comeu a folha
The turttle that called the monkey ate the leaf
11. A formiga foi beijada pela borboleta
The ant was kissed by the butterfly
12. Quem chamou o urso?
Who called the bear?
13. A vaca que o elefante molhou comeu o capim
The cow that the elephant wetted ate the grass
14. O pai do macaquinho coçou ele
The father of the little monkey scratched him
15. O macaco empurrou a vaca e pisou a flor
The monkey pushed the cow and stepped on the flower
16. A girafa beijou o elefante que pegou a flor
The giraffe kissed the elephant that grabbed the flower
17. A televisão foi carregada pelo menino
The television was carried by the boy

18. Que gato o porco beijou?
    Which cat did the pig kiss?
19. O cachorro que lambeu o gato derrubou a cadeira
    The dog that licked the cat tipped the chair
20. A menina e o menino derrubaram a cadeira
    The girl and the boy tipped the chair
21. O cachorro que o urso chamou pulou a cadeira
    The dog that the bear called jumped over the chair
22. O coelho foi jogado do muro
    The rabbit was pushed from the wall
23. O pai do menino vestiu ele
    The boy's father dressed him
24. Que macaco empurrou o cachorro?
    Which monkey pushed the dog?
25. O gato lambeu a pata
    The cat licked the paw
26. A tartaruga chutou o cachorro que pulou a poça
    The turttle kicked the dog that jumped the puddle
27. Quem o tigre pulou?
    Who did the tiger jump?
28. O Coelho que chamou o cachorro chutou a bola
    The rabbit that called the dog kicked the ball
29. O sapo espetou o cachorro e sujou o pé
    The frog poked the dog and dirtied the foot
30. A mãe da menina se penteou
    The girl's mother combed herself
31. O gato foi pisado pelo coelho
    The cat was stepped on by the rabbit
32. O cavalo que o elefante machucou comeu a maçã
    The horse that the elephant hurt ate the apple
33. Quem empurrou a garota?
    Who pushed the girl?
34. O menino abriu a porta
    The boy opened the door
35. O coelho viu que o macaco se molhou
    The rabbit saw that the monkey wet himself
36. A menina sujou o garoto

The girl dirtied the boy

37. O macaco mordeu o coelho que derrubou o balde
    The monkey bit the rabbit that tipped the bucket
38. O carrinho foi puxado pelo menino
    The cart was pulled by the boy
39. O pai do elefantinho se molhou
    The father of the little elephant wet himself
40. O macaco que empurrou a vaca comeu a mação
    The monkey that pushed the cow ate the apple
41. Que menina o menino sujou?
    Which girl did the boy get dirty?
42. O macaco e o urso levantaram a pedra
    The monkey and the bear lifted the rock
43. O coelho que o cachorro molhou derrubou a cerca
    The rabbit that the dog wetted tipped the fence
44. O jacaré viu que o macaco se mordeu
    The alligator saw that the monkey bit himself
45. O porco beijou o gato que segurou a flor
    The pig kissed the cat that grabbed the flower
46. Quem espetou o sapo?
    Who poked the frog?
47. O gato foi empurrado do sofá
    The cat was pushed from the sofa
48. O sapo que espetou o cachorro pegou a flor
    The frog that poked the dog grabbed the flower
49. O pai do porquinho se coçou
    The father of the little pig scratched himself
50. Quem o palhaço beijou?
    Who did the clown kiss?
51. O urso abraçou o tigre e derrubou a cerca
    The bear hugged the tiger and tipped the fence
52. O porco que a tartaruga mordeu pegou o balde
    The pig that the turttle bit grabbed the bucket
53. O coelho espetou o sapo
    The rabbit poked the frog
54. Que coelho chutou o porco?
    Which rabbit kicked the pig?

55. O tigre machucou o cavalo que pulou a cerca
    The tiger hurt the horse that jumped over the fence
56. O jacaré foi mordio pelo leão
    The alligator was bit by the lion
57. O pato disse que o gato jogou ele
    The duck said that the cat pushed him
58. Que elefante o urso espetou?
    Which elephant did the bear poke?
59. O urso que coçou o tigre pegou a bola
    The bear that tickled the tiger grabbed the ball
60. O palhaço e a bailarina carregaram a caixa
    The clown and the ballerina carried the box
61. O urso que o tigre empurrou segurou a pedra
    The bear that the tiger pushed held the rock
62. O porco disse que o tigre espetou ele
    The pig said that the tigre poked him
63. Que girafa beijou o elefante?
    Which giraffe kissed the elephant?
64. O chinelo foi calçado pela menina
    The sandal was put on by the girl
65. O tigre pulou o coleho que segurou a bola
    The tiger jumped over the rabbit that grabbed the ball
66. O gato foi lambido no sofá
    The cat was licked on the sofa

EXAMPLES OF PICTURES: