UNIVERSIDADE DE LISBOA

FACULDADE DE LETRAS

THE ACQUISITON OF GENDER AGREEMENT IN L2 PORTUGUESE BY ADULT L1 HUNGARIAN SPEAKERS

Veronika Flóra Lacsán

Trabalho final orientado pela Doutora Nélia Alexandre, especialmente elaborado para a obtenção do grau de mestre em Linguística

Dissertação

2015

To my guardian angels,
Mikey, Alex and Nagyi.
I miss you so much.

Table of Contents

Abstract

This study investigates the knowledge of adult Hungarian L1 speakers of gender agreement in Portuguese L2. According to the Failed Functional Features Hypothesis and other theoretical accounts, difficulty with the acquisition of L2 grammatical gender by an L1 speaker of a language that does not exhibit gender features (Hungarian) is due to a critical period effect on one hand, and to the failure to reset options or set new parameters that are already fixed in the L1 on the other hand. Therefore, these accounts predict that adult Hungarian L1 speakers will not be accurate with L2 Portuguese gender agreement. Thirty-six L1 Hungarians with L2 Portuguese and thirty Portuguese L1 speakers were tested on their gender assignment accuracy with determiners and adjectives via an online cloze test. Based on the existing literature, we formulated six hypotheses to test in this study. We wanted to investigate whether the level of proficiency of the L2 Portuguese participants, the distance between the noun and the adjective, the default/non-default word markers and the left/right position of the target item (determiner/adjective) have an effect on gender assignment accuracy. The L2 Portuguese speakers were also divided into subgroups, based on whether they spoke another Romance language or not, and whether they spent more than 3 months living in an immersion context in a Portuguese-speaking country or not. These subgroups were investigated in terms of transfer from the other Romance L2 to the L3 (Portuguese) and the effects of naturalistic input received in an immersion context.

Keywords: Acquisition of Portuguese L2; Hungarian L1; gender agreement; transfer from L2 to L3; immersion effects.

Resumo

O objetivo principal do presente trabalho é a investigação de uma área específica da aquisição de L2, nomeadamente a aquisição da concordância de género em português L2 por falantes adultos de húngaro L1. Este é provavelmente o primeiro estudo que investiga a aquisição de L2 entre estas duas línguas, pelo menos quanto ao aspeto em questão.

O húngaro pertence às línguas urálicas enquanto o português é uma língua românica, o que significa que possuem raízes completamente diferentes. Esta diferença tipológica e genética conduz a muitos assuntos interessantes para estudar, sendo um deles a aquisição da concordância de género. É um facto bem conhecido que as línguas românicas têm morfologia verbal e nominal rica, incluindo morfemas de género, os quais podem ser de dois tipos: masculino e feminino. Pelo contrário, a língua húngara não exibe morfologia alguma de género, que não significa que não tenha uma morfologia ainda mais rica do que a portuguesa. Segundo a *Failed Functional Features Hypothesis* (Hawkins & Chan 1997), depois de uma criança fixar os parâmetros de género na sua língua materna nunca mais pode alterá-los como adulto durante a aquisição de uma L2. No nosso caso, esta hipótese prediz que falantes cuja língua materna não tem morfologia de género, como o húngaro, não vão ser capazes de adquirir esta categoria gramatical de uma L2 que tem morfologia de género, como o português. A investigação levada a cabo neste trabalho é orientada por esta hipótese.

Para o efeito, em primeiro lugar, no Capítulo 1, fornecemos uma síntese das teorias e da literatura em aquisição de L2 mais importantes para este trabalho e de seguida expomos algumas outras hipóteses existentes na área mais específica da aquisição da condordância de género. Todas estas teorias e conceitos estiveram na base da criação do nosso teste experimental, além disso, vão aparecer ao longo da nossa investigação e constituir o suporte da análise dos nossos dados. Desta maneira, no Capítulo 1 revemos hipóteses teóricas sobre qual é o estado inicial da aquisição de uma L2 (se há ou não há acesso à Gramática Universal, ou se ele é feito apenas parcialmente); quais os processos que conduzem à aquisição da língua materna e em que diferem esses processos daqueles que funcionam durante a aquisição de uma L2. Além disto, fazemos uma síntese do funcionamento da interlíngua dos falantes de uma L2; analisamos hipóteses sobre a existência de um período crítico para a aquisição da linguagem; exploramos opiniões sobre as razões e os fatores da potencial fossilização da L2 e apresentamos os processos psicolinguísticos

que podem conduzir à aquisição da L2. Finalmente, falamos da questão específica do funcionamento e dos processos de aquisição da condordância de género em L2.

Em seguida, o Capítulo 2 trata da literatura e das teorias sobre o género gramatical e condordância do mesmo. Primeiro, exploramos como funciona a classe gramatical do género entre as línguas do mundo e para isto baseamo-nos, principalmente, nos trabalhos do Corbett (1991, 2005, 2006). Para a definição da classe gramatical do género, este autor usa tanto critérios semânticos como formais e declara que para a identificação de tal classe o critério fundamental é a capacidade de desencadear concordância sintática (Corbett 1991, 2006). Também explicamos como evoluiu o género gramatical, começando pelas línguas indo-europeias, passando pelo latim e chegando às línguas românicas. Aqui presentamos uma síntese sobre a inconsistência de uso do género e algumas regras gerais da sua atribuição. Depois, estudamos especificamente o género em português, baseando-nos, sobretudo, nas propostas de Villalva (1994). Segundo esta autora, o género gramatical português não é uma categoria flexional, mas uma propriedade inerente dos nomes. Nesta seção, explicamos quais as regularidades e irregularidades que existem em algumas classes de nomes em português quanto à atribuição do valor de género. Por fim, usamos a classificação da Villalva (1994) para apresentarmos as classes nominais e adjectivais portuguesas e mostramos como funciona a concordância de género em português.

O Capítulo 3 fala sobre a metodologia aplicada na recolha dos dados deste estudo e expõe as hipóteses da investigação. As seis hipóteses do estudo procuram investigar se os resultados dos informantes húngaros melhoram com a subida do nível de proficiência (Hipótese 1); se a distância entre nome e adjetivo tem algum efeito na atribuição correta dos valores de género (Hipótese 2); se os grupos de nomes com índice temático *default* recebem melhores resultados do que os grupos com marcadores *non-default* (Hipótese 3); se os informantes húngaros têm melhores resultados com adjectivos do que com determinantes (Hipótese 4); se os informantes que já falam outras línguas românicas têm melhores resultados do que os que não (Hipótese 5); e se os que viveram em contexto de imersão têm melhores resultados do que aqueles que não (Hipótese 6).

Utilizamos os dados de 36 falantes nativos de húngaro com português L2 e mais 30 informantes de português L1 como grupo de controlo para a análise das nossas hipóteses. Todos os informantes húngaros são estudantes universitários adultos e foram divididos segundo os seus níveis de proficiência (nível A2, B2 e C1). Fizemos uma divisão entre falantes e não falantes de

outra(s) língua(s) românica(s) e entre os informantes que viveram e que não viveram em contexto de imersão num país lusófono. O teste experimental foi um "*cloze test" online* que incluía 70 frases no total, 51 delas frases-alvo e o resto frases distratoras. Com base nos resultados obtidos, criamos uma base de dados, onde as respostas corretas foram marcadas como 1 e as incorretas como 0.

Para o teste experimental, criamos 12 grupos de nomes com diferentes marcadores (2 *default* e 10 *non-default*) com base nas classificações de Villalva (1994) e outras pistas visuais com as quais os estudantes de L2 podem contar. Além disto, estabelecemos três tipos de frases onde o adjetivo é posto cada vez mais longe do nome. No Tipo 1, com a estrutura DET+N+ADJ usamos adjectivos atributivos; no Tipo 2 (DET+N+COP+ADJ) aplicamos estruturas copulativas; e no Tipo 3 (DET+N+REL+ADJ) recorremos a uma oração relativa em que o adjectivo modifica o nome antecedente e entre eles ocorre, na posição do sujeito, um N [+animado, +humano] que possui o valor de género contrário ao N antecedente, criando assim um efeito de intervenção.

No Capítulo 4 apresentamos uma análise pormenorizada dos dados obtidos. Partindo das tendências que identificamos nos resultados, no Capítulo 5, fazemos a exposição das conclusões e a análise das hipóteses.

A Hipótese 1 foi parcialmente confirmada, pois vimos que havia progresso do nível básico para o intermédio, mas o mesmo não se verificou de forma evidente do nível intermédio para o nível avançado.

A Hipótese 2 não foi confirmada, porque obtivemos os melhores resultados com as frases do Tipo 3, que incluíam as estruturas mais complexas (com efeitos de intervenção entre o N e o ADJ) e o adjetivo ficou mais longe do nome. As frases do Tipo 1 (adjetivos atributivos) obtiveram ligeiramente piores resultados, mas os valores mais baixos vieram das frases do Tipo 2 (estrutura copulativa). Assim, concluímos que talvez não seja a distância entre nome e adjetivo condicionar a aquisição da condordância do género, mas a construção frásica e a diferença tipológica entre o húngaro e o português.

A Hipótese 3 foi infirmada; porque os informantes húngaros conseguiram, inclusive, melhores resultados com alguns grupos *non-default* de nomes (sobretudo com o marcador nominal <*-dade*>) do que com os grupos *default*, embora os últimos também tinham resultados próximos do nível nativo. Isto pode-se dever à saliência gráfica do marcador <*-dade*>, que é dissilábico e tem sempre o mesmo valor de género ([+feminino]) em todas as línguas românicas,

razão pela qual se pode falar de um efeito positivo de *transfer* de L2 românica também.

A Hipótese 4 foi infirmada, pois à posição esquerda do N (a do determinante em português e a do adjetivo em húngaro) revelou ser mais proeminente para os húngaros do que a posição à direita (a do adjetivo em português e a das declinações em húngaro).

A Hipótese 5 foi igualmente infirmada, já que o domínio de pelo menos uma outra língua românica (o *transfer* de L2 para L3) apenas se revelou vantajosa nos níveis básicos, mas a partir do nível intermédio, possivelmente por causa do treino escolar específico na área da aquisição do género gramatical, os informantes que não falaram outra(s) língua(s) românica(s) alcançaram melhores resultados do que aqueles que falaram pelo menos uma L2 românica.

Finalmente, a Hipótese 6 também foi infirmada, porque, de um modo semelhante a Hipótese 5, o efeito de imersão apenas se revelou vantajoso no nível básico, mas a partir do nível intermédio os informantes que não viveram em contexto de imersão produziram melhores resultados, talvez por causa do treino formal específico na área da aquisição do género gramatical.

Apesar de, na maioria dos casos, não temos obtido resultados de nível nativo, as percentagens de acerto foram bastante altas, o que indica que os húngaros têm acesso parcial à Gramática Universal, mas não conseguem adquirir (completamente) a concordância de traços de género, o que confirma a *Failed Functional Features Hypothesis*.

Palavras-chave: Aquisição de Português L2; Húngaro L1; concordância de género, *transfer* de L2 para L3; efeitos de imersão.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my thesis advisor, Nélia Alexandre, for her true dedication and patience throughout all these months. I am most thankful for the endless hours she spent with me and our work, revising every single line I wrote, explaining complicated theories, researching literature and passing her wisdom onto me. She truly is an inspiration. If it wasn't for her encouragement and optimism, this work could not have been possible. Thank you for believing in me.

Secondly, I would like to thank my family from the bottom of my heart. I am forever grateful for my father for making my studies in Portugal possible. Even though it meant that I would live on the opposite side of Europe, he was always on my side, helping my dream come true. I truly appreciate the support and loving words of my mother, who, despite working in a completely different research field, always helped me in any way she could. Thank you for the longest phone calls and numerous emails, for everything. Also, I thank my sister for all the times she made me laugh and for all the ways she made me feel better when I was down. Even though I am far away from them, my family is always with me.

I don't even know how I could thank João for all that he has done for me. He helped me with every single detail he could, however minute, and put up with my tantrums throughout all these times. These past two years have not been easy for me, but he always gave me comfort in the stressful times and cheered me up when I felt overwhelmed with the work. Thank you for being there for me, thank you for your support and insight. Thank you for being you. You truly, really are my rock, my pillar. Thank you for all those dishes you washed and all the snacks you picked up for me. I am forever indebted to you.

I also want to express my gratitude to all the people in Hungary who helped me find participants for the study. They were not easy to track down or to reach and it would have been impossible without them. Therefore I am most grateful to Gáspár Kónya, João Miguel Henriques, Eszter Andorka, Evelin Hargitai and Ildikó Szijj.

Furthermore, I would like to thank all my participants who were so kind as to contribute to this study, since their input was the most crucial part of all this work. I owe you one.

And lastly, I would especially like to thank my linguistics professor, Tibor Berta and my Portuguese language professor, Gáspár Kónya from the University of Szeged for having passed their love for linguistics and Portuguese onto me. If it wasn't for them, I would never have chosen this path or ended up where I am now. They inspired me so greatly, for which I will be forever grateful.

List of Tables and Maps

Chapter 5

List of Appendices

Digital Appendices:

List of Abbreviations

ADJ – Adjective

Agr – Agreement

C – Complementizer

CEFR – Common European Framework of Reference

CEM – Cumulative Enhancement Model

COP – Copula

CP – Complementizer Phrase

D – Determiner

DEM – Demonstrative

DET – Determiner

DP – Determiner Phrase

FFFH – Failed Functional Features Hypothesis

FL - Foreign Language

IL – Interlanguage

IND – Indicative

L1 – Native Language

L2 – Second Language

L3 – Third Language

LAD – Language Acquisition Device

L_n – N^{th} Language

MSIH – Missing Surface Inflection Hypothesis

N – Noun

PFFs – Parameterized Functional Features

PL – Plural

PLD – Primary Linguistic Data

POSS – Possessive

PRS – Present Tense

PST – Past Tense

REL – Relative

SG –Singular

SL – Second Language

SLA – Second Language Acquisition

SVO – Subject-Verb-Object

TPM – Typological Primacy Model

UG – Universal Grammar

WALS – World Atlas of Language Structures

Chapter 1
Introduction

In this first chapter, we shall discuss some of the most pertinent theories to the study of Second Language Acquisition (hereafter, SLA) and the acquisition of grammatical gender agreement in a second language (hereafter, L2). First, we shall provide a more generic overview of the field and some of the most important concepts that exist up to day then we move onto the more specific issue of the acquisition of gender features in a second language. The objective of this dissertation is to find out whether native speakers of Hungarian, a language that does not morphologically mark gender (i.e. a language that lacks grammatical gender features and thus gender agreement), are capable of acquiring the grammatical gender agreement of a second language that does possess gender features, such as European Portuguese.

This dissertation is principally inspired by the works of Montrul, Foote & Perpiñan (2008), who investigate the knowledge of gender agreement of Spanish L2 learners and heritage speakers who differ in age and mode of acquisition, concluding that heritage speakers are more accurate with gender assignment than L2 learners; and by Franceschina (2005), who compares the accuracy and tests the native-likeness of gender assignment of L2 Spanish speakers whose native language (hereafter, L1) is a language that exhibits morphological gender marking with that of other L2 speakers of Spanish whose L1 does not exhibit gender morphology, discovering an advantage that the former L2 speakers have over the latter.

Furthermore, we draw on Ferreira (2011), who tries to identify the variable patterns in the acquisition of grammatical gender marking within the noun group based on the analysis of errors produced by foreign students of various proficiency levels when learning Portuguese as an L2, concluding that all levels of proficiency show inaccuracies in gender assignment; and on Mariotto (2014), who examines the sensitivity of native English speakers to violations of gender agreement between [+/-animate] nouns and the predicative adjectives. Her experiment was conducted with learners of Portuguese as an L2 with B1 level proficiency, concluding that the L2 speakers were much less accurate with gender assignment than the L1 Portuguese control group.

Based on the aforementioned studies, our research shall concentrate on the accuracy of gender assignment in an L2 Romance language, in our case Portuguese, by L1 speakers of a

language that does not exhibit gender marking, Hungarian. Since there is a complete lack of research for L2 acquisition between these two languages, we hope to introduce a novel research subject to which we shall apply an already well-established and well-investigated theoretical background, which is to be discussed in Chapter 1 and 2.

For the sake of this investigation, we created a grammatical test that comprises three different types of sentences where the adjective is placed increasingly further away from the noun, in order to examine whether distance has an effect on the acquisition of correct gender assignment; twelve noun groups based on twelve different word markers found in European Portuguese to analyze which word markers pose more or fewer problems to our participants; and two variable items (determiner and adjective) that receive their gender specifications from the noun they are in relation with, to attempt to find out which item causes more difficulty as for correct gender assignment. Our methodology is described in Chapter 3.

Based on these criteria, we shall examine the performances of our participants from three different levels of proficiency (A2, B2 and C1), and later on divide them into four subgroups: in order to compare the results of those participants who learned an additional Romance language to Portuguese with those who did not (Romance L2 vs. Non-Romance L2 Subgroups); and furthermore, to compare the results of those participants who have lived in an immersion context with those who only learned Portuguese in a classroom environment (Immersion vs. Non-Immersion Subgroups). We shall apply the theories exposed in Chapter 1 and Chapter 2 to our findings that will be analyzed and discussed in Chapter 4 and 5.

1.1. Theoretical background

Let us begin this first chapter by establishing that this study is inserted in the framework of the field of research called *second language acquisition*. SLA is defined by Van Patten & Benati (2010:1) as an area that "focuses on learners and learning rather than teachers and teaching". According to Gass & Selinker (2008:1), SLA is "the study of how learners create a new language system". Consequently, SLA as a research field deals with what *is* and what *is not* acquired of an L2. Therefore, the ultimate role of SLA is to address "the fundamental questions of how learners come to internalize the linguistic system of another language and how they make use of that linguistic system during comprehension and speech production" (Van Patten & Benati 2010:2).

As we have mentioned, in this chapter, we are going to explore briefly some basic concepts that we consider paramount in the research field of SLA and that are highly relevant to the present study. In Chapter 4 and 5, we shall use these concepts and hypotheses for the analysis and understanding of our data.

This study shall apply the abbreviation *L2* instead of using the terms *second language* or *target language* to refer to a language that is acquired after one's native language(s). There are some authors, however, who distinguish between *second language* and *foreign language* acquisition. The first term is applied to contexts in which the language is acquired in a region where it is spoken outside of the classroom, like learning Portuguese in Portugal or Hungarian in Hungary, i.e. there is immersion in the target language; whereas the second term is used in situations where a language is learned in a region where it is not spoken outside of the classroom, like learning Portuguese in Hungary and vice versa (Van Patten & Benati, 2010).[1] In our study, the latter case applies; nevertheless, we will not distinguish between these two terms. Thus, "it is common in SLA to place all contexts of learning under the umbrella term [of SLA]" (Van Patten & Benati 2010:2).

Some other authors, like Krashen (1988), may also differentiate between *subconscious language acquisition* (what learners spontaneously "pick up") and *conscious language learning* (grammar rules and strategies that are explicitly taught), stating that these are two independent systems underlying the ability of adults to learn an L2. Our research will focus on conscious learning and shall attempt to deduce the strategies and processes that underlie it. We will not, however, use Krashen's differentiation between learning and acquisition and shall use these terms interchangeably.

[1] In Portuguese literature, the same differentiation exists. According to Leiria (2000), second language (SL) refers to the learning and use of a non-native language within a territory where it does possess a recognized function, whereas foreign language (FL) is a term used for instances where the learning and use of a non-native language does not have any socio-political status in that given territory.
SL is frequently (one of) the official language(s) and is usually learned at school, but not necessarily used at home. This is the case of Portuguese in Portugal's ex-colonies (Cape Verde, Mozambique, Angola, São Tomé and Príncipe, Guinea-Bissau, East-Timor), where Portuguese was adapted as the official language, however a significant part of the population speaks an African or a creole language as their L1.
FL is acquired at school, physically far away from where the given language originates and in many cases it is taught by non-natives, through textbooks, in a formal context and without any sort of immersion. FL is usually based on the/a standardized version of the given language, which is its most prestigious form. This is exactly how our Hungarian participants have acquired standard European Portuguese in Hungary.

By attempting to describe aspects of the study of SLA, we have to part from a huge cluster of theories that have their roots in first language acquisition and were developed during the course of the second half of the past century. The present investigation, however, will exclusively concentrate on and highlight those theories that are directly relevant to our research and, furthermore, as we have stated above, shall be used as an attempt to account for our findings. For a more exhaustive account on theories about the field of SLA, see, for instance, Ellis (1997), Gass & Selinker (2009), Van Patten & Benati (2010).

First, in this dissertation, it shall be assumed that what underlies the linguistic competence, production and comprehension of a native speaker is an unconscious and abstract linguistic system: a grammar. L1 grammars are constrained by an innate biologically endowed language faculty called Universal Grammar (UG) (Chomsky 1986, White 2003). According to White, "UG provides a *genetic blueprint*, determining in advance what grammars can (and cannot) be like. […] UG includes invariant principles, that is, principles that are generally true across languages, as well as parameters which allow for variation from language to language" (White 2003:2).

Secondly, we have to state that despite the existence of numerous theoretical approaches to the study of the processes in SLA, our research is integrated in the line of studies that operate based on Chomsky's (1986, 1995) Principles and Parameters framework and the Minimalist Program (MP). According to an innatist view, language acquisition is guided and driven by an interaction of the following elements in (1.1) (see, for instance, Chomsky 1986, White 2003 and Franceschina 2005):

(1.1) Essential elements in language acquisition:

a. computational resources made available by Universal Grammar (UG) (principles and operations)[2],

b. representational resources made available by UG (feature inventory),[3]

c. exposure to primary linguistic data (PLD)

-(Franceschina 2005:2).

[2] "[..] the term computational resources [is] not used to refer to real-time processing abilities. Instead, it [is] used to refer to the abstract principles and operations underlying the workings of mental grammars" (Franceschina 2005:9).

1.1.1. The initial state of SLA

When children start learning their L1(s)[4], it is assumed that the principles and rules provided by UG is what they begin with. From birth, or starting even earlier, children are subjected to PLD, based on which they build up a language-specific lexicon and set the parameters of UG to values appropriate for the given language. This is a gradual process during which their grammars "may be restructured over the course of time as the child becomes responsive to different properties in the input. In due course, the child arrives at a steady state grammar for the [L1]" (White 2003:2).

Corder (1967, 1981) was the first researcher who hypothesized that L2 learners – as opposed to children learning their L1 – do not part from their L1, but from UG on which they rely while developing their own linguistic system – a system he called 'transitional competence'. To this day, there are two basic points of view on the supposed initial state of SLA:

- There is access to UG (White 1986, Schwartz & Sprouse 1996).

- There is no access to UG (Clahsen & Muysken 1986, Schachter 1988, Bley-Vroman 1989).[5]

All other existing hypotheses are variations of these two basic assumptions. The Partial Access (Partial Transfer) hypothesis, for example, claims that the L2 learner has only partial access to UG (see more on this topic below in section 1.2.1.).

In sum, whatever the supposed initial state is, it is never believed to be a 'blank slate' by any scholar (White 2003, Van Patten & Benati 2010). Furthermore, Corder (1967) also argued that the L1 could act as a positive resource for SLA, because it facilitates the acquisition of features that are similar in the L1 and in the L2 (positive transfer).

---

[3] "[..] [this term] will be used to refer to the grammatical primitives that [the faculty of language] utilizes in the representation of grammatical knowledge", i.e. grammatical features (Franceschina 2005:9).

[4] Children don't necessarily learn only one language as their L1. For example, in multicultural families and multilingual societies it is highly possible that children will learn two (bilingualism) or even more grammars simultaneously as their L1s.

[5] Hardly any researcher in the field still maintains this point of view, however it was popular in the 1980s and early 1990s.

1.1.2. Differences between L1 and L2 acquisition

One of the major differences between L1 and L2 acquisition is that L1 acquisition is based solely on positive evidence, whereas SLA is guided by negative evidence (França 1997)[6]. This means that (adult) L2 learners already have a complete linguistic system available (their L1) and all parameters are fixed for the grammar of the L1. What is believed to happen when they start learning an L2 is that they start adjusting this new linguistic system to the one that is already mastered. They do so through negative evidence (recognition and correction of errors) and other, mainly intentional, cognitive strategies which are not at all effortless or target-like at times (see psycholinguistic hypotheses below in section 1.1.3.3.).

Another important difference between the two types of acquisition is the very fact that L1 acquisition happens spontaneously and effortlessly, and the result is a perfectly target-like grammar.[7] The observation that L2 learners "fail to acquire certain types of knowledge even when the evidence [that] is available to them [in our case the PLD], is very rich" (Franceschina 2005:2) was made by Chomsky (1986) and he called it *Orwell's* or *Freud's problem*. This problem was tested by Reuland (1993) in terms of first language acquisition. He concludes that children do not significantly speed up language acquisition even if they are exposed to enriched PLD. It has to be mentioned here that some researchers working in the weak continuity tradition, such as Clahsen, Eisenbeiss & Vaikikka (1994) and the maturational hypothesis framework, like Radford (1990) believe that "L1 development consists of gradual building up of phrase structure" (Franceschina 2005:2). If this is true then children must not at all be sensitive to certain aspects of the input at some stages of development (Franceschina 2005:2). What is more, they receive a great deal of 'noise' in the input, (such as mistakes, hesitations, accents, various registers and even different dialects, etc.) and they still manage to, eventually, obtain a fully target-like grammar of their L1. Therefore, they must ignore some aspects of the input and "[impose] some sort of structure on the chaotic input" (Franceschina 2005:3).

Franceschina (2005) proposes that Orwell's problem be investigated in terms of the course and outcome of SLA. "If, when exposed to the same PLD, L1 and L2 learners do not

[6] When children make mistakes and get corrected by adults they do not internalize the corrections and continue to make the same mistakes until they have enough PLD that show the correct (target) underlying grammar. Adults, however, are taught explicitly about ungrammaticality and learn through the internalization and correction of their mistakes.

[7] Given that the child is developing typically, i.e. is mentally and physically healthy.

arrive at similar states of knowledge about the language they are learning, then it is possible to assume that the same data may have different triggering status for different types of learners" (Franceschina 2005:3). She also claims that different groups of L2 learners seem to be insensitive to different aspects of the PLD and that they may be persistent in failing to detect certain aspects of the input. "Orwell's problem is about this blinding effect observed in some areas of language learning" (Franceschina 2005:4).

This study, following in the footsteps of Franceschina (2005) and Montrul, Foote & Perpiñán (2008), will attempt to shed some light on the sensitivity (or lack thereof) of adult Hungarians with different levels of proficiency in L2 Portuguese to a morphosyntactic aspect of Portuguese, which is grammatical gender. The acquisition of grammatical gender is a problematic area in SLA (Franceschina 2005, and Montrul, Foote & Perpiñán 2008). Since Portuguese is a language that is filled with grammatical gender cues, then the acquisition of grammatical gender and syntactic agreement is most probably a very pertinent section of grammar within which Orwell's problem can be investigated.

1.1.3. Interlanguage

Even though there are quite a few parallels that can be drawn between L1 and L2 acquisition, there is one notion that only applies to SLA: *interlanguage* (henceforth IL).[8] This term was introduced by Larry Selinker in 1972 "to refer to the linguistic system evidenced when an adult second language learner attempts to express meanings in the language being learned" (Tarone 2006:747).[9] Considered as a separate linguistic system, IL differs both from the L2 and from the learner's L1, but at the same time, it is closely connected to both by interlingual identifications in the perception of the learner (Tarone 2006).

IL is usually considered as pertaining only to adult SLA. Adults are thought to no longer have access to the language acquisition device (LAD) - the "innate language learning structure that was instrumental in their acquisition of their native language" (Tarone 2006:748, see also Chomsky 1986). Children, however, since they are still able to employ this structure when

[8] Selinker's notion of IL was mainly based on what Corder (1967) had previously called 'transitional competence'.

[9] The original definition was only formulated for adult L2 learners but later on it was also extended to child L2 learners. For more information, see Tarone (2006).

learning an L2, can re-engage the LAD, avoiding errors and fossilization that characterize adult IL (Tarone 2006).

1.1.3.1. Critical period

Therefore, we have to discuss another important concept (used in the innatist literature) pertaining to SLA that concerns a period within which language acquisition seems to be more effortless, because of cognitive development. This timeframe is called the *critical period* and it is a highly controversial notion; partly because its existence is not accepted by all researchers and also it is quite difficult to accurately define until exactly when it applies. However, some researchers state that after this period access to UG is still possible but the re-setting of parameters is not (Smith & Tsimpli 1995, see more in section 1.2.1.).

In general, it is assumed that the end of this period coincides with the end of puberty, which biologically entails hemispheric specialization. This assumption was made based on the weakened ability of adults to recover from brain injuries that occurred in the Broca and Wernicke areas that are supposedly responsible for language (Lenneberg 1967). This way, adults are considered to have passed this milestone (Johnson & Newport 1989, White 2003). Penfield & Roberts (1959) argue that the reason why adults are less successful at acquiring an L2 than children is because of the loss of neural plasticity in the brain. Novel research claims that the culprit is the myelination of neuronal pathways over time that makes the connection between neighboring neurons more difficult (see more on this discussion in Hyltenstam & Abrahamsson 2003). Some researchers, such as Felix (1985, 1987), suggest that cognitive maturity may not at all be an advantage in adult SLA. His Competition Hypothesis claims that "there is a certain kind of competition between UG and domain-general cognitive systems which has adverse effects for adult L2 learning" (Felix 1985, ap. Franceschina 2005:44). Around puberty, problem-solving abilities also emerge that interfere with the language learning process. Thus adults seem to be poorer at SLA than prepubescent children (Birdsong 1994).

Be it as it may, we will not attempt to define the precise age that "might be a cut-off point for [a] native-like acquisition of [the L2]" (Franceschina 2005:6). What we will try to find out is whether adult L2 learners are able to acquire (in a native-like manner) one certain aspect of L2 grammar, which is grammatical gender agreement, thus, as this study concentrates only on adult SLA, it could be considered as an investigation into post-critical period L2 grammars.

Some researchers state that SLA can never be complete. Furthermore, it is quite consensual that "age effects may be different in different areas of the grammar. A distinction is usually maintained between critical period claims for morphosyntax, phonology and vocabulary" (Franceschina 2005:37). According to Slabakova (2008) and her Bottleneck Hypothesis, functional morphology (to which the marking of Portuguese grammatical gender also belongs) cannot fully be attained by L2 learners, while syntax and semantics can, in fact, be completely acquired, since she considers the latter universally accessible through UG. It is also generally accepted that the critical period for phonology might end earlier than that for morphosyntax (Seliger 1978, Walsh & Diller 1981). Interestingly, there seems to be no age limit after which L2 vocabulary acquisition becomes problematic (Singleton 1995). Montrul (2011) argues that incomplete acquisition can be the result of transfer errors; "such that the structure of the [L1] impedes acquisition of other aspects of the [L2], which are either not present or exist in a different form in the [L1] of the learners" (Montrul 2011:593) (see more in section 1.2.1.).

1.1.3.2. Views on fossilization

This way, as we have mentioned earlier, another crucial property of IL is that it can *fossilize* if the learner begins the acquisition of the L2 after puberty (Chomsky 1986). Fossilization is a phenomenon that is claimed to occur at the endstate of SLA (see, for instance Franceschina 2005). IL stops developing before it could fully become identical to the L2 (Tarone 2006). This way, adult L2 learners are never able to reach the level of facility in the use of the L2 as a child learner does in the same L2 (Selinker 1972). The fact that (mentally and physically healthy) children always succeed at acquiring their L1 impeccably and that adults not always do at their L2 is another fundamental difference between L1 acquisition and SLA, as we have mentioned above. Selinker (1972) claims that the fossilization of IL is inevitable because of neurolinguistic reasons. Lenneberg (1967) and also Scovel (1988) share Selinker's view and argue "that the causes of […] fossilization are neurolinguistic in nature and related to the process of cerebral lateralization, which is completed at puberty. But there is certainly disagreement among interlanguage researchers as to both the inevitability of fossilization and (relatedly) the causes of fossilization" (Scovel 1988, ap. Tarone 2006:751).

On the other hand, some researchers argue that fossilization is not inevitable and that it is caused by sociolinguistic constraints. Tarone (2006:751) claims that

> "[…] if learners can identify with the L2 social group, or if their need is great enough, they will be able to continue learning the [L2] until their production/perception is indistinguishable from that of native speakers. This issue also is far from settled, since it relates to matters of human potential rather than humans' actual behavior."

Schumann (1978) also seems to think that the reason behind fossilization is learner-internal and sociolinguistic in nature and lies in the L2 learners' attitude towards the L2 culture. Schumann uses the expression 'acculturation' to refer to the phenomenon of integrating into a foreign society and claims that it is the lack thereof that causes fossilization. The L2 learners who do not manage to succeed in 'acculturating' are those who remain socially and psychologically distant (Schumann 1978).[10]

From a purely sociological standpoint, Peirce (1995) argues that success in SLA correlates to the L2 learners' degree of engagement with L1 speakers of the language in question and their willingness to modify their social status and identity. SLA appears to be a lot more successful when the L2 learner makes an 'investment', according to Peirce. This means that they actually commit to acquiring the L2 because they believe that speaking the L2 will increase the value of their 'cultural capital', in other words, they believe that knowing and speaking the L2 will give them certain benefits and knowledge to operate successfully within most contexts of the society.

Krashen (1982, 1985) has also considered social factors as impediments to native-like attainment, namely, his Affective Filter Hypothesis argues that adult L2 learners' negative affections, such as anxiety, fear of embarrassment, low self-confidence or motivation act as obstructions for the LAD, this way making for a kind of filter. Children, however, do not seem to possess these affective filters, therefore Krashen regards the existence of this filter in adults as another basic difference between SLA and L1 acquisition.

The participants of this study all started to learn L2 Portuguese in a classroom environment far from Portugal, therefore (most of them) have not had the need to try to integrate into or function in a foreign society and attempt to 'acculturate' or make an 'investment'. However, some participants did, in fact, spend a considerable amount of time in an immersion

[10] These distant behaviors are "determined by factors such as domination, assimilation, enclosure, group size, congruence, attitude" and, furthermore, "language shock, culture shock, motivation and ego permeability, among other things" (Schumann 1978, ap. Franceschina 2005:46).

context, abroad in Portugal. The time spent 'immersed' in Portuguese culture and society could have triggered a change in their affective filters and their attitudes toward learning the language much more thoroughly than those who never lived in Portugal and this might have happened out of sheer necessity – the necessity to be able to communicate with the L1 speakers. Therefore, this process might have led them to acquire more of the L2 than those who never lived in an immersion context. We shall test this hypothesis (see Hypothesis 6 in Chapter 3) and compare the results of those participants who have lived more than 3 months in an immersion context to those who have not. Based on the assumptions of Tarone, Schumann, Krashen and Peirce mentioned above, we expect to see a difference in the success rates of these two groups of L2 speakers in favor of those who have lived abroad in Portugal.

Some factors external to learners can also contribute to imperfect attainment, such as the environment in which they are learning the L2 and thus the type and amount of received input (Lightbown 1985). For instance, if we compared the number of hours of exposure of an infant to the L1 in a naturalistic environment and those of the adult L2 learner to the L2 in a classroom environment, we would most definitely find significant differences. The exposure of a child to the L1 is basically non-stop, while and L2 learner takes only a limited amount of lessons in the L2 (except for when they are studying in an immersion context) and/or applies a limited amount of other resources (television or radio programs in the L2, etc.). This must also influence the ultimate attainment between L1 and L2 speakers, although Patkowski (1980), Birdsong (1992) and Flege & Liu (2001) question whether the length of naturalistic exposure (for instance residence in the country where the L2 is spoken) is an adequate means to predict attainment. This means that beyond a minimum period of naturalistic exposure to the L2 more input does not appear to have a significant improvement on L2 learners' proficiency. These aforementioned authors made this assumption based on the fact that some immigrants who had lived in the L2 country for a very long period of time still do not possess native-like grammars, in spite of their lengthy exposure to the target naturalistic L2 input. Therefore no matter how much naturalistic L2 input the L2 learners receive, this does not guarantee that they will become perfectly native-like in the L2 (Franceschina 2005).

Now that we have seen a few opinions on the existence and causes of fossilization, we shall have a look at the other side of the debate: Pascual y Cabo & Rothman (2012) argue that there is no such thing as "incomplete acquisition", i.e. fossilization. These researchers formulated

their view for heritage speakers[11] and state that competence outcomes of heritage speakers and adult L2 learners are similar, in spite of the fact that heritage speakers acquired their heritage language in childhood (within the critical period). This suggests that age might not be the most prominent and deterministic factor in explaining adult SLA outcomes. "This means that one cannot simply dismiss adult L2 grammars as being wild grammars – unconstrained by UG from our view or whatever underlying linguistic/cognitive mechanisms one subscribes to – simply because they diverge from the arbitrarily chosen monolingual benchmark" (Pascual y Cabo & Rothman 2012:454). The authors claim that a heritage language (and in a broader view an L2) is never incompletely acquired, but is simply different in its state of ultimate attainment, because the input itself is different (it has either already undergone attrition or, in our case, does not come from a native speaker, but from another L2 speaker – a nonnative language teacher).

In this study, nevertheless, we will not argue for or against a fossilization stage of L2 acquisition for gender agreement nor will we regard our participants as having reached the endstate of SLA. In fact, all our participants are still in the process of acquisition of L2 Portuguese, therefore they have not yet arrived at any final state or state of ultimate attainment of the L2 grammar that could (eventually) fossilize. Even though this study is not based on longitudinally collected data, we shall analyze our data as samples from different states of IL development, since we have groups of L2 speakers with different levels of proficiency (level A2, B2 and C1).

1.1.3.3. Psycholinguistic processes driving SLA

In this section, we shall have a look at Selinker's (1972) hypotheses, according to which adult SLA is guided by five psycholinguistic processes of a latent psychological structure instead of an LAD, which are the following:

A. **Native language transfer:** contrastive analysts proposed L1 transfer to be the only shaper of IL and, though their claim has been disproved, this process still plays a key role,

[11] Heritage speakers are bilinguals who acquired the majority language (the language of the broader society they are inserted in) and the family's language (the heritage language, i.e. a minority language that differs from the broader society's language) in their childhood. These speakers tend to end up as speakers of the majority language in adulthood due to the fact that they do not usually receive education in the heritage language and that the domain of use of the heritage language is very restricted.

however it is not the only one. Selinker (1972) proposed that this process happens through 'interlingual identifications', i.e. the learner, "in approaching the task of learning a second language, [perceives] certain units as *the same* in their [L1], IL and [L2]" (Tarone 2006:748).[12]

B. **Overgeneralization of L2 rules**: this process refers to the extension of a general rule to even the exceptions to said rule (which are yet to be learned). This process shows clear evidence of progress and its extent in SLA is comparable to that of L1 acquisition. A textbook example for this strategy is the overgeneralization of the default English past tense marker <–*ed*> extended to irregular verbs (e.g.: **catched*, **drinked*, intead of *caught*, *drank*) or the overgeneralization of paradigms of regular verbs extended to irregular verbs in Portuguese (**fazi* instead of *fiz* 'did', from the verb *fazer* 'to do', using the analogy to *comer* > *comi,* 'to eat> ate').

C. **Transfer of training**: it applies when the L2 learners use rules that they learned from textbooks or instructors. Sometimes this process is successful, thus the resulting IL rule matches the L2 rule, but sometimes errors occur, called 'induced errors'. For instance, if an L2 student learns the use of the English past perfect tense as being the 'past past', then this can lead to the use of this tense for all events that occurred a long time ago, whether or not this event precedes another past event that occurred afterwards, such as **Columbus had discovered America in 1492* (Tarone 2006).

D. **Strategies of communication**: these strategies are applied by the learner when the IL system proves to be insufficient for the task in communication, thus it is a way to solve a problem. While attempting to communicate meaning, if the learner lacks the needed item, they can resort to various strategies to make themselves understood. This could include describing the item, for instance. The problem is, however, that these linguistic forms and patterns used in such situations could become fairly permanent in the learner's IL.

[12] "For example, they may perceive [L1] *table* as exactly the same as [L2] *mesa*, and develop an interlanguage in which *mesa* can (erroneously in terms of the [L2]) be used in expressions like 'table of contents', 'table the motion' and so on." (Tarone 2006:748-749)

E. **Strategies of learning:** these are mechanisms that a learner uses consciously when trying to master the L2. A good example for this process is the establishment of interlingual identifications, i.e. the conscious comparison of production in the IL with L1 and L2. Another such strategy is the mnemonic memorization of target items (vocabulary, declensions, dialogues, etc.). Unfortunately, these memorized lists can often get confused with L1 (or L3) forms, for example.

Based on systematic evidence, all these psycholinguistic processes were found to be present in the construction of ILs (Tarone 2006). However, according to Selinker (1972), the only relevant data are utterances produced by L2 learners when they attempt to communicate meaning. Therefore, he disregarded grammaticality judgment tests, elicited answers and classroom exercises where the learner was concentrating on grammar rules or the L2 form, dubbing these techniques erroneous because, in his view, these data would not provide information about the IL system, but merely about the learner's perception of the L2 system – and these systems were not the same. Nevertheless, this is solely Selinker's opinion on the matter. Corder (1967), for instance, thinks differently: he claims that researchers are to use all available data sources in exploring IL systems, therefore the techniques rejected by Selinker are promoted by him and many others. According to White (2003:18), "it is important to recognize that there is no one methodology that is appropriate for investigating all aspects of linguistic competence".

We have to ask, however, a fundamental question when analyzing data from different elicitation techniques: which linguistic system do all these data pools provide information about, the IL, the L1 or the L2? Tarone (2006:750) states that

> "in essence, the most basic research design question involved in the study of [ILs] – what data shall one use to study [IL]? – raises very complex issues concerning the relationship between intuitions of grammaticality, language production, and language perception, very similar to issues raised by Labov (1972) in sociolinguistic work. This issue is unresolved in SLA research and in fact is complicated by evidence that interlanguage seems to vary by discourse domain."

With IL being our focus, we aimed to replicate the investigations conducted in Franceschina (2005) and Montrul, Foote & Perpiñán (2008), who, as we mentioned, studied the

acquisition of grammatical gender in L2 Spanish – Spanish being morphosyntactically similar and genetically close to Portuguese.

Thus, we found it fundamental to provide this extended – but still incomplete – review of IL hypotheses and literature, since the main subject of our study is the IL of our participants, especially the errors they might commit on each level of proficiency, which we shall use as indicators about the state of their ILs.

1.2. Acquisition of L2 gender features

In the following sections we shall explore how abstract or uninterpretable morphological features can be inserted into the study of SLA and how they behave with respect to UG constraints.

Let us begin by quoting Franceschina once again, who assumes that "we can only acquire knowledge of linguistic properties for which we possess primitive features capable of capturing them" (2005:4). She believes that "learners 'break down' the linguistic input and process it through the use of the set of features made available by UG" (2005:4). In Franceschina (2005, id.), she attempts to investigate the idea that the major difference between L1 and L2 learners is in "the properties of the linguistic input that they can 'assimilate' into their mental grammars" and also whether "such difficulties in assimilating certain aspects of the L2 data may be the result of adult L2 learners' use of different sets of representational primitives to 'break down' the L2 input".

The data on L2 gender agreement acquisition that we shall provide is an example, just like Franceschina's (2005) work, of "different groups of learners reacting differently to the same PLD, with some learners 'assimilating' some features […] and others apparently ignoring them" (Franceschina 2005:4).

The morphology of Portuguese gender agreement makes the dependencies that exist between nominals very clear. In fact, gender morphology triggers the "establishment of uninterpretable gender features in native speakers and some [L2] speakers" (Franceschina 2005:4), however it does not seem to have the same triggering effect in some other L2 speakers. Franceschina (2005) argues that this is due to the lack of appropriate uninterpretable gender features in the latter case. In our case, this precisely means that (some) Hungarian speakers might

have difficulty acquiring Portuguese morphological gender and, consequently, syntactic gender agreement, therefore we shall attempt to investigate this assumption in our study.

As to attempt to answer why this aforementioned difficulty in the acquisition of uninterpretable gender features happens, we need to see what happens on the morphology/syntax interface.

Lardiere (2000) claims that there is dissociation between the recognition of surface morphology rules and the recognition of abstract syntactic features. This researcher suggests that it is only the knowledge of surface morphology rules that can pose difficulty in SLA whereas abstract syntactic features are non-problematic. This is why L2 learners may experience difficulty when mapping abstract feature specifications onto the adequate morphological form (Lardiere 2000, see more on the issue in section 1.2.1.).

It wasn't, however, only Lardiere who had such an assumption, but Prévost & White (2000) also suggested that L2 learners may have problems when trying to access the correct form from the lexicon, but they are, in fact, able to attain perfectly native-like L2 representations. This defect results in the retrieval of default or stem forms instead of the target form. Their proposal is known as the Missing Surface Inflection Hypothesis (MSIH), which is based on the assumption that "[...] variability in adult L2 performance does not reflect a deeper lack of functional categories or features associated with tense and agreement. Rather, L2 learners have difficulties with the overt realization of morphology" (2000:104). Much like Lardiere's claim, this means that there might be dissociation between abstract syntactic features and surface morphology. Prévost & White assign this mismatch to a mapping problem between the abstract features and the surface forms.

1.2.1. The Failed Functional Features Hypothesis

When children start acquiring their L1, as we have seen, first they are exposed to PLD "which [trigger] the fixing of the values of the parameters, crucially in conjunction with the surface morphophonological paradigms, which instantiate Agr, C and D" (Hawkins & Chan 1997:188). Afterwards, they encode parameter values in the lexical entries for items belonging to the aforementioned functional categories. Therefore "functional lexical items consist of a pairing of the morphophonological form with functional features encoding a specific parametric option (or options)" (Hawkins & Chan 1997:188). According to these two scholars, after the parameter

setting is complete in L1 acquisition, only those functional features that are already encoded in the entries for specific lexical items become accessible to modification. They assume that it becomes impossible for adult L2 learners to reset options or set new parameters that are already fixed in the L1. However, some other principles of UG remain operative to govern the construction of grammar and, what is more, the morphophonological segments of functional lexical items are still free to be modified, since they are non-parametrized. Thus, even if a category like Agr has its value of features fixed for the L1 before the end of the critical period, "it is still possible for an L2 learner to map new morphophonological material onto those features" (Hawkins & Chan 1997:189). This theory is called the Failed Functional Features Hypothesis (FFFH).

Tsimpli & Smith (1991) and Smith & Tsimpli (1995) adopted "a modular view of linguistic ability in which the [LAD] is composed of a number of differentiated (but interacting) modules, they argue that the principles of UG are located in a separate component from the parametric options which give rise to variation between grammars" (Hawkins & Chan 1997:188 based on Tsimpli & Smith 1991). These parametric options are located in a functional component called the *UG lexicon*, which primarily contains the functional categories Agr (agreement), C (complementizer) and D (determiner) (Hawkins & Chan 1997). They claim that the inaccessible subparts of UG, i.e. those that are subject to a critical period, seem to be the features belonging to these functional categories, also called as parameterized functional features (PFFs).

This all comes from the observation that those children whose L1 is a morphologically rich language that marks gender will learn very early on that nouns are classified in different classes according to their gender (Hawkins & Franceschina 2004). Thus, when such adults learn other languages that share a similar morphological structure to their L1, they tend to be consistent in the selection of gender agreement, i.e. tend to be more native-like in terms of ultimate attainment (Franceschina 2005). This cannot be said, though, for children whose L1 does not exhibit grammatical gender and learn a language that does mark gender agreement (i.e. when the features of the L1 do not match those of the L2, speakers find it difficult to set the values of the parameter in L2). These adults will show a persistent inconsistency with respect to gender agreement in the L2, no matter how advanced their level of proficiency is or how very exposed they are to the L2 or how native-like their performance is in other aspects of the grammar of the L2 (Montrul, Foote, & Perpiñán 2008).

For even more evidence, we can draw on two other investigations carried out by Franceschina (2001a, b) that examined the acquisition of L2 Spanish grammatical gender marking by L1 English speakers. Even though the participants of these studies were very highly proficient L2 speakers they still not perform impeccably, occasionally making gender mistakes. Franceschina reports that "The mistakes were not distributed randomly across gender marked categories, but only occurred in those categories that receive gender via syntactic agreement (or concord). By contrast, nouns, which are inherently marked for gender, were never incorrectly inflected for gender" (Franceschina 2005:58). Al-hamad (2002) also noted such asymmetries in the distribution of gender mistakes in the L2 Arabic of Chinese L1 speakers.

This way, according to what has been said above we can suppose that L1 speakers of Portuguese can easily acquire the grammatical gender agreement of L2 Italian, Greek, French, German, etc. (any language that has grammatical gender morphology) and vice versa, because they simply have to map the new morphophonological material onto [+gender].

On the other hand, however, a speaker of L1 Hungarian (or L1 English, or Japanese, or Korean, or any other language that has no grammatical gender morphology) whose default gender feature value is [-gender] do not have the simple option of assigning new material to the gender feature when learning L2 Portuguese, Italian, Greek, French, German, etc. because this feature is non-existent in the L1, therefore there is nothing to map the new material onto. This is an assumption that we are going to investigate in this study.

In sum, if the values of the parameter for gender are not automatically activated during the acquisition of the L1 before the end of the critical period (like in Hungarian, because of the complete lack of grammatical gender marking), adult language learners will generally have, especially in the beginning, difficulty acquiring this aspect of an L2 that does exhibit gender agreement. This happens simply because after the critical period these parameters may no longer be available and thus they will have to resort to, for instance, the psycholinguistic mechanisms introduced in section 1.1.3.3. (Hawkins & Franceschina 2004).

Chapter 2
Grammatical Gender

2.1. Grammatical gender and agreement around the world

After having explored a few theories connected to the research field of SLA, let us continue by establishing why it is worthwhile and interesting to study grammatical gender. Greville Corbett, perhaps the most well-known scholar who has spent decades conducting research in the area of grammatical gender sums this issue up perfectly in the introduction to his book *Gender* (1991:1):

> "Gender is the most puzzling of the grammatical categories. It is a topic which interests non-linguists as well as linguists and it becomes more fascinating the more it is investigated. […] One of its attractions for linguists is that there are interesting aspects of the study of gender in each of the core areas of linguistics. And work on it promises practical benefits, even in the short term, in meeting the problems which gender causes in [SLA]. In the longer term, research into gender will be important for at least two other areas: first, it can shed light on the way in which linguistic information is stored in the brain; and second, it has implications for natural language processing, notably for the elimination of local ambiguities in parsing."

The word *gender* has its origin in the Latin word *genus*, a cognate of the Antique Greek *genos*, meaning "kind" or "sort" and this term "is normally reserved for kinds or classes of nouns" that are, as Hockett (1958: 231) puts it, "reflected in the behavior of associated words" (ap. Corbett 2006:749).

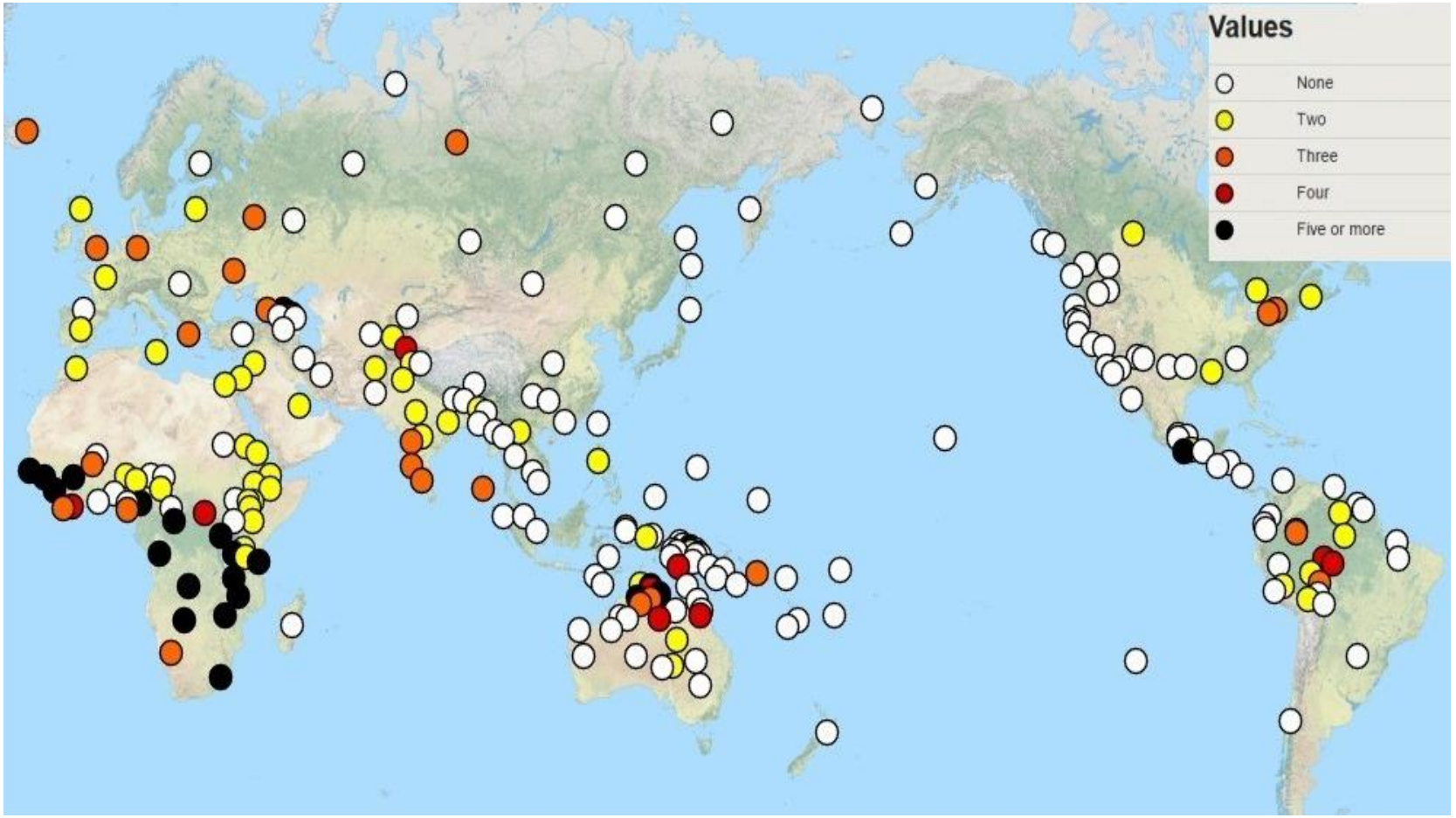
This 'association' is called *agreement*. There can be various types of agreement based on which the noun inventory is divided into different genders or other nominal classes, as we will see (Corbett 2006). Therefore, as Corbett claims, "a language has a gender system only if noun phrases headed by nouns of different types control different agreements. No amount of marking

on a noun can prove that it has gender; the evidence that nouns have gender in a given language lies in the agreement targets that show gender" (2006:749).

In Portuguese, gender marking, and thus agreement, affects determiners, quantifiers, adjectives, some pronouns and, naturally, nouns. Gender, however, is not the only form of nominal classification that exists in the languages around the world, in fact, there are numerous other possibilities.

However widespread this grammatical category may seem in Indo-European languages, it is not even nearly a universal phenomenon. Nichols (1992), for instance, studied a total number of 174 languages from different language families and she found that only 47 languages of this sample (1/4) contained some variety of nominal classifications, including gender. Corbett's (2005) sample, on the other hand, consisted of 256 languages from various language families, of which 112 had gender systems.

As Corbett (1991) noted, while gender is paramount for some languages (e.g. Niger-Congo languages as observed in Map 1 below), it is completely absent in others (e.g. Hungarian, and some Sino-Tibetan, most Austronesian, Eskimo-Aleut, Miwok-Costanoan, Quechuan languages, etc., more than 140 languages, according to the World Atlas of Language Structures, see Map 1 below).



Map 1: Number of gender classes in languages around the world[13]

[13] Map 1 was copied from http://wals.info/feature/30A#2/25.5/148.2 and is based on Corbett (2013).

2.1.1. The concept of grammatical gender, gender agreement and classifications

Corbett (1991, 2006) assumes that the marking of a noun's gender depends on two types of information:

(i) Semantic: the meaning of the nominal item itself, in which the value of the gender marking is directly associated to the referential content of the noun, i.e. it is not arbitrary.

(ii) Formal: the respective form of the nominal item, including arbitrary morphophonological rules that can be associated to gender values.

Furthermore, these semantic and formal properties that define the concept of grammatical gender can only be applied to a given language if, and only if they trigger the phenomenon of syntactic agreement. This important fact allows us to distinguish between systems with nominal classes and systems with gender classes.

Based on these criteria and on data from samples from 200 languages, Corbett (1991) proposed the existence of the following classifications:

a. Systems with purely lexical gender marking, where the marking of gender values is attained through the application of semantic criteria
b. Systems with simultaneously lexical and grammatical/formal gender marking
c. And the possibly exclusively grammatical/formal gender systems[14]

This study relies on the application of (ii), as Portuguese belongs to this classification. Corbett (2005) concludes that, out of his 256-language sample, 59 languages belong to this assignment system, slightly more than those that belong to (i) (53 languages).

Therefore we shall assume that the gender specification of Portuguese nouns is lexical, which means that this information ([+feminine] or [+masculine]) comes from the lexicon and is encoded in the lexical items. Nevertheless, in the composition of a sentence structure determiners and adjectives that are in relation with the noun have to verify their gender values through the grammatical process of gender agreement with the noun, since the gender value of determiners

[14] Corbett (2006) excluded the last type of system-classification later on due to lack of evidence is his subsequent sample data. This demonstrates how gender marking is indeed linked to semantic criteria, in a lesser or greater extent.

and adjectives is encoded grammatically unlike that of nouns, which is encoded lexically (see more on this issue in section 2.2.).

2.1.2. On the origins and evolution of gender marking

In primitive Indo-European, gender marking was based on the animacy of the noun's referent. Masculine and feminine nouns were [+animate] and neutral nouns were semantically [-animate]. In Antique Greek and Latin these three gender classifications survived, although their assignment no longer depended closely on the animacy of the referent of the noun. Later on, as Classical Latin developed into Vulgar Latin, the neutral gender gradually disappeared, which could be attributed to the many coincidences between morphologically inflected forms of neutral and masculine nouns (Villalva 1994, Gouveia 2004). The next step in the evolution of Vulgar Latin was the total elimination of the neutral class, and this is how most Romance languages ended up having only the masculine and feminine classes.[15]

Additionally, some languages, such as English, have lost the grammatical category of gender during their evolution. This is visible, for instance, in the masculine/feminine nouns 'prince/princess' and in the mere personal pronouns 'he/she'. Also, the noun 'ship' or names of countries in English can appear with the covert or connotative [+feminine] pronoun 'she'.

2.1.3. Gender value vs. biological sex

It is paramount that we note that biological sex is not to be confused with grammatical gender. Biological sex only applies in the case of [+animate] nouns, where it represents the sex of the entity in question. Thus the grammatical gender of the nouns can coincide with the biological sex of its referent, if, and only if the referent is [+animate]. As for [-animate] nouns, there is likely to be no connection with biological sex, whatsoever.

Throughout the languages that exhibit grammatical gender, a given referent might have different nominal gender values. Even languages that are genetically closely linked can vary in this sense. For example, the word for *nose* is [+feminine] in Spanish (***la*** *nariz*), but [+masculine] in Portuguese (***o*** *nariz*), French (***le*** *nez*) and Italian (***il*** *naso*). If we take two genetically different languages, like German and Portuguese, the same can be observed. For instance, the word for *sun* in Portuguese (***o*** *sol*) is [+masculine], but it is [+feminine] in German (***die*** *Sonne*). However, the

[15] Romanian still has the neutral class in addition to masculine and feminine.

gender values can also be the same even crosslinguistically; in the case of *oven* in German (***der*** *Ofen*) and Portuguese (***o*** *forno*) both are [+masculine]. This illustrates how the distribution of gender values is completely inconsistent (Ferreira 2011).

The number of gender categories can also differ from one language to the other, as we have seen on Map 1. When two nouns belong to the same agreement class, i.e. they have the same gender value, they "must take the same agreements under all conditions – that is, if we hold constant other features as case and number" (Corbett 2006:750). In many languages there is no doubt as for the number of genders, as is the case of Portuguese, which is a two-gender system. In Corbett's (2005) sample, out of the 256 languages, of which 112 exhibited nominal classifications, 50 had an at least two-gender system, therefore he dubs this kind of system fairly common. However, another Romance language, Romanian, is quite polemic in this sense (Corbett 2006). Furthermore, according to Corbett (1991), some Indo-European languages have up to three genders (for comparison, 26 languages out of his 256-language worldwide sample), but the Slavonic group, for example, is introducing additional novel sub-genders. Corbett (2005) found 12 four-gender systems and 24 languages with five- or more gender systems in his sample (see Map 1.)

These details above all support the claim that biological gender does not necessarily correspond to grammatical gender.

2.1.4. Gender agreement and rules for gender value assignment

As we have mentioned before, syntactic agreement is the most important factor for the establishment of a gender classification system (Foley & Van Valin 1984, Spencer 1999). Thus, there can be as many genders in a language as many possibilities for syntactic agreement the noun can trigger. Additionally, Vilela (1973) noted that article-noun agreement and the article itself play a key role in the identification of the gender of the noun.

Franceschina (2005:71-72) distinguishes between 'triggers' (or controllers) and 'targets' of gender agreement. "Triggers are defined as the lexical items containing intrinsic gender values that can be copied onto other lexical items, namely targets, which are not inherently marked for gender and receive this via syntactic agreement (or concord)". This author also claims that "targets are the lexical items that agree with nouns either present in the linguistic context or

tacitly implied by it. What counts as a gender target varies crosslinguistically, and it can include almost any word category (Franceschina 2005:72).

Corbett (1991) proposed some morphophonological rules for cases when the semantic information in itself is insufficient. These rules can be applied to Portuguese, for there are some word markers and derivational suffixes (see more below in section 2.2.2.) that carry gender information and determine if the nominal products will be masculine or feminine. This reveals the existence of correlation between these morphological items and the gender value of the noun integrating them (Ferreira 2011).

These rules can be synthesized in (2.1):

(2.1) Those nouns whose final segment is α belong to the gender value β.

Corbett (1991) states that in such systems there is only one morphological form of the noun that determines the differentiation of the gender and the alteration of the phonological segments of the nominal item. Furthermore, syllable stress can also be a factor in nominal gender value assignment. Corbett uses this rule to describe Afro-Asian languages, but he also mentions French as belonging to this group of languages.[16]

We applied such morphophonological rules in this study for the elaboration of our experiment for the analysis of the acquisition of L2 gender agreement, since these are clear visual cues that L2 learners can rely on when semantic information on the gender value is unavailable (see Chapter 3 for further details).

As a summary of this section, we shall accept the composite definition in (2.2) of grammatical gender developed by Franceschina (2005:78), as we find it comprehensive and also applicable to our study.

(2.2) *Composite definition of grammatical gender*

a. genders are classes of nouns that result from the partitioning of the lexicon into nominal classes;

[16] Nouns like *mutualité* 'mutuality', *activité* 'activity', *singularité* 'singularity' are always feminine and *voyage* 'journey', *courage* 'courage', *âge* 'age', *message* 'message', *massage* 'massage' are masculine (Corbett 1991: 51-61). The Portuguese equivalents, on the contrary, have the same gender values in both *mutualidade*, *atividade, singularidade* and *viagem, coragem, idade, mensagem, massagem* ([+feminine]).

b. nouns are gender triggers, and other categories marked for gender are targets;

c. gender triggers and targets are structurally related;

d. nouns in gender systems are exhaustively classified (in most cases this means inherently classified, but there are some exceptions);

e. the following categories can be gender targets: determiners, pronouns, quantifiers, numerals, possessives, adjectives, past and passive participles, verbs, adverbs, complementizers, adverbs, adpositions;

f. gender assignment rules vary crosslinguistically;

g. the domain of gender agreement shows some crosslinguistic variation.

After this general overview on the concept of grammatical gender and agreement across the languages of the world, we shall now move onto the exploration of literature specifically on Portuguese grammatical gender and agreement.

2.2. Grammatical gender in Portuguese

According to Villalva (1994), Portuguese nominal gender assignment is lexically conditioned and affects all nouns. If we closely examine the morphophonetic and semantic characteristics of Portuguese nouns, we will see how dubious gender value assignment is in numerous cases.

In Portuguese, [masculine] is the default value, thus [feminine] is the non-default form of nouns (Mattoso Câmara 1994:88, Villalva 1994). When taking a look at Portuguese nouns, it is obvious that not all of them end in unstressed <–*a*> or <–*o*>, which are the default theme indices[17] for feminine and masculine values, respectively. However, Portuguese noun endings are, in fact, a bit more complex; there are nouns ending in: unstressed <–*e*>, stressed vowels (<-*á, -é, -ó*>), consonants (<-*z, -s, -l*>), and (phonetically) nasal vowels or diphthongs (<-*ã, -ão*>). The gender assignment of these nouns is more complicated, since all of these endings can be both [+feminine] and/or [+masculine].

[17] We adapt the denomination 'theme index' based on Villalva (1994). In comparison, later on we shall use the term 'word marker' when describing our methodology (see Chapter 3). The difference between the terms is that our 'word marker' refers to the orthographic ending of nouns, therefore it is a concept based on visual cues that L2 learners rely on. Some of the 'word markers' used in this study, nevertheless, do coincide with some of Villalva's 'theme indices', in those cases we maintained the latter denomination.

To this day, it is a controversial topic whether gender can be classified as a nominal inflection category or not: Carvalho (1967, 1984:601) states that Portuguese nouns do not exhibit inflection in gender, thus sharing Villalva's (1994) point of view, while Mattoso Câmara (1994:88) claims that gender is a formal category and not a semantic one.[18]

2.2.1. Singularities of some [+animate] Portuguese nouns

Let us now have a look at some [+animate] nouns that exhibit unique behaviors as to the assignment of gender values and gender contrast. Numerous nouns belong to each group below, however, not nearly every [+animate] noun reflects these qualities.

A. [+animate] nouns that admit variation (contrasts) in gender

In the case of nouns whose referent is [+human], the opposition between [feminine/masculine] values is expressed through either lexical processes as in *mulher/homem* 'woman/man'; through the contrast in theme indices as in *menino/menina* 'boy/girl'; or through morphological processes as in *o imperador/a imperatriz* 'emperor/empress', *o galo/a galinha* "rooster/hen" (derivational suffixation of the [+feminine] form) (for more information, see Villalva 1994).

B. [+animate] nouns that do not admit variation (contrasts) in gender

Among the group of [+animate, +human] nouns we can find some which refer to entities having only one gender value (but referring to entities with either biological sex) that do not admit gender contrasts, neither syntactically nor morphologically. This subgroup is called 'sobrecomuns' in Portuguese and consists of nouns like: *a pessoa* 'person', *o indivíduo* 'individual', *a testemunha* 'witness', *a criança* 'child' (Villalva 1994, Ferreira 2011, Mariotto 2014).

There is another such group of nouns called epicenes. This is a term used for nouns whose referents are animals and they present only one gender value independently of the biological

[18] Although, according to Mattoso Câmara (1984), inflection does exist in word pairs of the type *menino/menina* 'boy/girl', *professor/professora* 'teacher' , *mestre/mestra* 'master', because in these pairs there is a masculine form in opposition to which there exists a feminine form, i.e. these pairs exhibit gender contrast. His view is not shared by Villalva (1994). Furthermore, in the case of [+animate] nouns having both gender values, (like *o/a dentista* 'dentist') and nouns with a single gender value, Mattoso Câmara (1966) states that inflection does not apply, because these nouns lack a masculine counterpart.

gender of said animal, such as *corvo* ‘crow’, *baleia* ‘whale’, etc. In such cases, the contrast between the two genders is established through the process of nominal composition where the radical does not change but it is associated with ‘male/female’, as in *corvo-macho/corvo-fêmea* ‘male/female crow’ (Villalva 1994, Ferreira 2011, Mariotto 2014).

Another classification of nouns that belongs to this group is that of nouns whose genders are underspecified in their strict nominal form. In these cases, gender values are established syntactically, i. e. through other syntactic constituents that specify or modify them, like determiners or adjectives. Morphologically, these nouns do not exhibit variation and can be semantically ambiguous, e.g.: *o/a dentista* ’dentist’, *o/a imigrante* ‘immigrant’, *o/a estudante* ‘student’ (Villalva 1994).

This demonstration strengthens the view that Portuguese grammatical gender cannot be thought of as an inflectional category because gender variation is not obligatory and can be carried out via morphological processes, like derivation and composition or via lexical contrasts. It is not restricted to [+animate] nouns, since, as we mentioned, it does not affect the totality of these nouns, and, in addition, it is far from being systematic. Therefore it is not coherent or cohesive which would, precisely, be the definition of an inflectional system (Villalva 1994, Ferreira 2011, Mariotto 2014).

Ultimately, what *is* Portuguese nominal gender and how does one attempt to define its essence? This study shall apply Villalva’s definition, which claims that it is a morphosyntactic category whose specification is either lexically determined or is the result of the intervention of a non-inflectional morphological process (1994:225).

2.2.2. Portuguese nouns and gender value assignment

As we have seen, Portuguese has its roots in Latin, a language which originally had three genders: feminine, masculine and neutral. The gradual loss of the neutral gender value came with some disturbance in the maintenance of direct relations between the remaining gender values (feminine/masculine) and semantic oppositions (animacy and biological sex) (Villalva 1994:227-228). The original five Latin declensions gave birth to four Portuguese thematic classes (theme indices <–*a*>, <-*o*>, null index, and athematic nouns).

The reduced number of gender values meant that those nouns that originally belonged to the neutral class had to be redistributed between the two remaining gender values

(feminine/masculine). This redistribution seems to have worked in the following way: those Latin nouns that ended in <–*a*> would become [+feminine] in Portuguese and those that ended in <–*o*> would become [+masculine]. This is proven by the fact that even originally [+feminine] Latin nouns with the theme index <–*o*> became [+masculine] in Portuguese (like *figo* 'fig' or *pinho* 'pine') and it also affected nouns with null theme indices (for more information see Nunes 1919, 1975:217 and Villalva 1994:228).

However, as we have mentioned above, the assignment of gender values is not at all systematic in Portuguese or, as a matter of fact, in any other Romance language. Thus, we must ask the obvious question: how can one identify the factors that are responsible for the attribution of a given gender value to a given Portuguese noun?

This question, apparently, has no solution to this day. Semantic criteria do not explain the assignment of gender values in inanimate nouns. Formal criteria are also insufficient since there are various counter-examples.

Corbett (1991) might have already provided us with a solution, though. As we have seen in (i) (see section 2.1.1), Corbett claims that gender contrasts usually have a semantic basis (distinction based on biological sex, animacy, and whether the noun refers to a human or not), but then again, semantic criteria on their own cannot account for the gender assignment of all nouns. The rest is called 'semantic residue' by Corbett (1991) and is often subject to formal criteria. Portuguese is such a mixed system, therefore, to the residue of nouns the morphophonological rule in (2.1) (see section 2.1.4.) can be applied.

The application of this rule to the Portuguese nominal system brings about four criteria for the identification of gender values in Portuguese nouns, according to Villalva (1994: 230-231), presented in (2.3):

(2.3) a. Those nouns that refer to beings belonging to the feminine sex have feminine as their gender value (e.g. *filha* 'daughter')
Those nouns that refer to beings belonging to the masculine sex have masculine as their gender value (e. g. *filho* 'son')

b. Those nouns that end in <-*a*> have feminine gender value (e.g. *filha* 'daughter')
Those nouns that end in <-*o*> have masculine gender value (e.g. *filho* 'son')

There is one ultimate problem, though: there exist nouns that end in <-*a*> and are masculine, like *poeta* 'poet'. This way, there still are situations of conflict. However, the predominance of semantic criteria proposed by Corbett (1991) makes these cases non-exceptional. The semantic criteria in [+human] nouns overrule the formal morphophonological criteria.

Other than the criteria in (2.3) above, there are further formal ones that make it possible to predict the gender value of complex ([-animate]) nouns. On the one hand, compound nouns made up of a verb and another element are always masculine in Portuguese (e.g. *abre-latas* 'can opener', *faz-tudo* 'handyman'). Furthermore, Portuguese derivational suffixes responsible for noun formation also carry gender information and are the following (Villalva 1994: 231):

Feminine derivational suffixes:

-*eza* as in *cert+eza* → *certeza* 'certainty';
-*aria* as in *drog+aria* → *drogaria* 'drugstore';
-*agem* as in *jardin+agem* → *jardinagem* 'gardening'
-*ção* as in *organiza+ção* → *organização* 'organization'
-*idade* as in *ambigu+idade* → *ambiguidade* 'ambiguity'

Masculine derivational suffixes:

-*mento* as in *envelheci+mento* → *envelhecimento* 'aging'
-*ume* as in *queix+ume* → *queixume* 'complaint'
-*ismo* as in *simbol+ismo* → *simbolismo* 'simbolism'

Mentioning some of these suffixes here might seem redundant, though, as we have already established in (2.3) that nouns that end in <-*o*> or <-*a*> belong to the feminine and masculine classes, no matter the suffix. Therefore, the true 'residue' that we have to account for are those nouns that refer to entities object to a distinction in biological sex, those that are feminine but do not end in <-a> and those that are masculine but do not end in <-o>.

According to Corbett (1991:66), the nouns whose gender values cannot be predicted by rules are exceptions that are tolerated by the system (Villalva 1994:232). The behavior of such nouns can be generally explained by diachronic processes in Portuguese, as we have seen.

However, the maximum percentage of these exceptions cannot exceed 15%, which was the limit detected by Corbett (1991) in his sample. It is unclear whether this is true or not for Portuguese (Villalva 1994:233), because there have been no quantitative studies on this particular topic so far.

All in all, Villalva (1994:233) claims that Portuguese "nouns require a specification as for their associated gender value and for the possibility to participate in gender contrasts. This specification is registered in the lexicon, associated to the radicals and to the suffixes of nominalization; and, in morphology, when the gender value is determined by a process of syntactic composition".[19]

This information, nevertheless, can be generalized to some extent (see, for instance, the aforementioned rules (i) and (ii) for semantic and formal criteria in section 2.1.1.) and organized into thematic classes. What are these classes, then? We will address this issue in the next section.

2.2.2.1. Portuguese nominal thematic classes

Portuguese nouns admit variations in gender and they also do in number, the latter being in fact an inflectional category, unlike the former, as we have established before. As for nouns, gender is an inherent property and might have referential value. The attribution of a gender value to a noun, once again, results from the combination of predominantly semantic and some formal criteria. These criteria have the capacity to predict the gender value but there is a considerable amount of exceptions to these rules (Villalva 1994).

What we can establish as a general rule is that (with few exceptions) all nouns with the theme index <*–o*> are [+masculine]. The rest can be organized in the following manner according to Villalva (1994:234), where 'variable' essentially means [+animate] and 'uniform' means [-animate]:

- variable [+masculine] nouns with the theme index *–a* — e.g. *poeta* 'poet'
- variable [+feminine] nouns with the theme index *–a* — e.g. *aluna* 'student'
- variable [+masculine] nouns with the theme index *–o* — e.g. *aluno* 'student'
- variable [+masculine] nouns with null theme index — e.g. *apresentador* 'presenter'
- variable [+feminine] nouns with null theme index — e.g. *imperatriz* 'empress'
- variable [+masculine] athematic nouns — e.g. *avô* 'grandfather'

[19] My translation from the Portuguese original.

- variable [+feminine] athematic nouns — e.g. *avó* ‘grandmother’
- uniform [+masculine] nouns with the theme index *–a* — e.g. *mapa* ‘map’
- uniform [+feminine] nouns with the theme index *–a* — e.g. *casa* ‘house’
- uniform [+masculine] nouns with the theme index *–o* — e.g. *livro* ‘book’
- uniform [+feminine] nouns with the theme index *–o* — e.g. *tribo* ‘tribe’
- uniform [+masculine] nouns with null theme index — e.g. *mar* ‘sea’
- uniform [+feminine] nouns with null theme index — e.g. *paz* ‘peace’
- uniform [+masculine] athematic nouns — e.g. *pé* ‘foot’
- uniform [+feminine] athematic nouns — e.g. *pá* ‘dustpan’

2.2.2.2. Portuguese adjectival thematic classes

Adjectives and nouns have a very similar morphosyntactic behavior, since both admit variation in number and in gender. The difference is that while gender is an inherent property of nouns, it is not of adjectives. On the other hand, unlike nominal gender, that of adjectives is strictly grammatical and is obtained contextually or via agreement with the noun they belong to (Villalva 1994:226).

The existence of contrasts in gender is an idiosyncratic property of adjectives. It is not semantically determined, but formally two generalizations were formulated by Villalva (1994:226):

a. Feminine is formulated via the addition of the theme index <*–a*> to the adjectival radical
b. All adjectives with the theme index <*– o*> are masculine (and this is the default value)

If we use these generalizations, it is enough to lexically register the radical *clar* and specify its gender value through the theme indices <*–o*> for [+masculine] or <*–a*> for [+feminine] if we want to obtain *claro* or *clara* ‘clear, light’, respectively.

However, this is only the case of variable adjectives, thus not all adjectives are this easy to specify for gender. To define the lexical specification of the “residue”, it is necessary to know whether they are variable or uniform[20] adjectives and to which thematic class they belong.

According to Villalva (1994) Portuguese has six thematic classes of adjectives:

-adjectives with the theme index *–o* — e.g. *claro* ‘clear’

[20] I.e. whether they exhibit a contrast between the two gender values or not.

-variable adjectives with the null thematic index	e.g. *falador* 'talkative'
-variable, athematic adjectives	e.g. *bom* 'good'
-uniform adjectives with the theme index *–a*	e.g. *careca* 'bald'
-uniform adjectives with the null thematic index	e.g. *leve* 'light, easy'
-uniform, athematic adjectives	e.g. *ruim* 'awful'

Villalva (1994) states that uniform adjectives are not specified for gender, since this is not required for the verification of the agreement with the noun (see also Chapter 3 for more on this topic).

2.2.2.3. Gender agreement in Portuguese

Romance languages are generally referred to as morphologically rich systems because they allow for the identification of grammatical categories via formal features that carry morphological information. Nevertheless, as we have seen, they exhibit mixed gender marking systems, relying both on semantic and formal information (see (i) and (ii) in section 2.1.1. for Corbett's classification).

In the case of systems with formal gender marking, when the radical of a noun is selected to trigger the emergence of other syntactic derivations it appears with other categories, resulting in syntactic agreement. Thus, agreement consists of the relation between two elements where the form of one determines the form of the other (Foucart 2008).

In Portuguese, we can encounter agreement in gender, number and person, which could be nominal (gender and number) or verbal (number and person). In this study we will only consider nominal gender agreement. This kind of agreement affects other word classes, such as possessive and demonstrative pronouns, determiners, quantifiers and adjectives. These have to obligatorily agree with the gender value (and number) of the noun. Gender, therefore, is an inherent feature of the nouns which is transmitted to other constituents or word classes (Mariotto 2014).

Chapter 3
Methodology

The main purpose of the present chapter is twofold: first, we shall provide a detailed overview of the empirical study itself that we employed in this research to be able to confirm or invalidate our assumptions detailed in our hypotheses below in section 3.3.; and second, based on the theories and literature that have been discussed in Chapter 1 and 2 with regard to SLA and grammatical gender agreement, we shall lay down the basic assumptions of this investigation that are going to be examined in Chapter 4.

3.1. The participants

The data reported in this study originate from a total number of 36 participants, all of whom are adult L1 speakers of Hungarian (between the ages of 19-29), speak Portuguese as an L2 which they started acquiring as adults (i.e. the age of first exposure was after puberty, therefore after the end of the critical period – see Chomsky 1986 and Chapter 1 of this dissertation), in a classroom setting, in Hungary.

The research was conducted on students from three Hungarian universities.[21] In addition to these participants, we also used the results of 30 adult native speakers of Portuguese as a control group.[22]

The experiment was not controlled for the sex of the participants, because it was not considered relevant for the study of the acquisition of L2 Portuguese gender agreement.[23] This study was inspired by the experiments provided by Montrul, Foote & Perpiñán (2008) and Franceschina (2005), where the sex of the participants was also considered irrelevant.

[21] In the University of Szeged and the University of Pécs Portuguese is taught as a Minor course as part of a Bachelor's Degree and in the Eötvös Lóránd University Portuguese exists as a Master's Course and also, as a Minor course.

[22] "[…]control groups are necessary simply to ensure: (i) that the tasks devised by the experimenter in fact are successful in testing what they are supposed to test; and (ii) that the facts in question are indeed as the experimenter supposes them to be (or as claimed in the theoretical linguistics literature)" (White 2003:55).

[23] However, this factor is very much pertinent in, for example, sociolinguistics and dialectology.

The levels of proficiency of the participants range from A2 through B2 to C1[24]. For each level of proficiency we employed 12 participants in our overall results.[25] Later on, the overall population was divided into subgroups of participants who spoke another Romance language and to those who did not; furthermore another division of the overall population into subgroups was made between those participants who lived abroad in an immersion context for longer than 3 months and those who did not.[26]

We did not implement a proficiency level placement exam, since we adopted the respective levels that our participants were frequenting at university. The duration of their studies ranges from 4 months (one semester) to 7 years.

Even though our sample consists of classroom-instructed students, some (10 individuals) had the opportunity to spend a longer duration of time (a minimum of 3 months) in a Portuguese-speaking country[27]. Thus, these students most probably have further improved their skills in an immersion context, which will be taken into consideration in our data analysis (see Hypothesis 6).

It is also worth noting that none of our participants speak Portuguese as a (chronologically) second language, but as an n^{th} language (L3, L4, L5, etc.)[28]. This means that all of them speak at least two other foreign languages in addition to their mother tongue and Portuguese, furthermore Portuguese was not the first foreign language to be learned (therefore it is at least an L3). In many cases (for 26 participants), these other L2s are Romance languages which have similar gender morphology to Portuguese (e.g. Spanish or Italian). In this study, we

[24] These levels correspond to the proficiency levels described by the Common European Framework of Reference for Languages (CEFR) developed by the Council of Europe.

[25] We did originally have 3 participants on level B1 and 9 participants on level A2 but, since the B1 students were very few in number and performed very much like the A2 students, we merged the two groups into one, thus we have 12 participants in our level A2 group.

[26] The basis of this division was a linguistic questionnaire that all L2 Portuguese participants filled out (see Appendix 2).

[27] In the case of this study: Portugal.

[28] The reason behind this fact that is that, in order to even apply to a university in Hungary, students are required to take at least one intermediate or advanced level foreign language exam, therefore they have to acquire an L2 before entering university. Since all of our participants are university students, they speak at least one other foreign language apart from Portuguese, however their level of proficiency in Portuguese might be higher than that in the other L2(s).

shall analyze the possible effects of transfer from these other Romance L2s during the acquisition of Portuguese grammatical gender.

In our data collection of all of our participants signed an informed consent (see Appendix 1).

3.2. The database and the experimental test

Both the target participants and the control group had to fill out an online cloze test (see Appendix 4). Additionally, the target participants were presented a paper-based language questionnaire which was a linguistic profile consisting of questions about their language studies in general. It also investigated which languages they spoke and whether they had spent a significant amount of time in a Portuguese-speaking country (see Appendix 2).

As for the online cloze test, we used the platform of *learnclick.com*[29] (see Appendix 4). This website is easy to use, very logical, and, furthermore, the spaces of the cloze test are easy to fill in. Thus the participants did not have the chance to get tired while filling out the test, since they could quickly click on and select the desired answer. This way, the data collected could be thought of as rather reliable and might actually reflect the real knowledge of the participants.

The platform of *learnclick.com* also provides an online database where all results of our participants were registered. The number and percentages of right/wrong answers were immediately given. Thus, *learnclick.com* proved to be ideal for the creation of this type of exercise. However, this initial online database still contained the results for the distractor sentences as well, therefore for the creation of the database we used for our analyses these data had to be extracted.

The test itself was a grammatical exercise of 70 sentences in total, 51 of them being target sentences and 19 of them being distractors. In each sentence, the participants had to choose the correct gender form of the determiner and of the adjective related to a target noun from the two provided drop-down options (feminine or masculine form).

As for the distractors, we used sentences containing the plural of nouns with the noun ending <*–ão*>. Nouns that end in this nasal diphthong have irregular plural morphology, which is

[29] Learnclick.com was created by Philip Perry and is available at http://www.learnclick.com/.

typically a complex issue for foreign learners of Portuguese.[30] This noun ending has three possible plural allomorphs: <–*ãos*>, <–*ães*> and <–*ões*>. In these sentences, all of these three options were provided to choose from.

The sentences appeared in a randomized order, differently for each participant. The time spent while filling out the test was not measured.[31] To facilitate the comprehension of the sentences for those participants with lower levels of proficiency, the Hungarian translation was provided after each sentence.[32]

3.2.1. Noun groups

The target objects were distributed in 12 different groups of nouns and 3 different sentence structures. These groups were created based on 12 word markers available in Portuguese[33], 10 of which were non-default and 2 were the default ones (see Chapter 2). Our assumption was that Hungarian learners were going to have more difficulty acquiring the gender values of the nouns in the non-default word marker groups than in the default ones (see Hypothesis 3).

In the selection of these word markers we relied on the salient orthographic cues that learners find in L2 Portuguese grammars and, consequently, they are taught to use in a classroom setting. In a way, all nouns in Portuguese are orthographically salient, since all of them can be divided into a relative small number of groups based on their orthographic ending, which could not be done for Hungarian nouns. However, we are going to use the term 'orthographic saliency' to refer to specific nouns or word markers from the noun groups established below.

Our target nouns were not controlled for the number of syllables or for stress, because

[30] The plural form of such nouns poses a problem not only for L2 learners, but even for Portuguese children acquiring their L1 (for more on this topic, see Freitas, Gonçalves & Gonçalves 2010 and Castro 2010).

[31] Since our study does not belong to the field of psycholinguistics, we did not specifically measure the reaction time or our participants for each sentence. However, we did urge them to take as little time to answer as possible, in order to obtain spontaneous answers.

[32] In hindsight, this could have facilitated the task with respect to Type 3 sentences (see more on this issue in Chapter 4 and 5.)

[33] These groups of word markers only partly coincide with Villalva's (1994) classifications, even though her system served as a basis for this study. The rest of the groups were created based on visual cues learners possibly rely on and that are taught at L2 Portuguese classes.

these factors are considered to be irrelevant in Portuguese gender assignment.[34] In each group we employed three [-animate]/uniform nouns for each gender value (if applicable)[35] to formulate the three types of target sentences with. These groups of nouns are the following:

- **Group 1** consists of uniform nouns with the non-default null theme index <-*e*>. This theme index can equally appear with [+feminine] and [+masculine] nouns, this way the gender assignment of such nouns is not transparent. This group excludes nouns ending with <-*dade*>, which are in a separate group (see Group 3).
  [+masculine]: *leite* 'milk', *filme* 'movie', *tomate* 'tomato'
  [+feminine]: *noite* 'night', *chave* 'key', *árvore* 'tree'

- **Group 2** comprises nouns with the non-default uniform word marker exclusive to the [+feminine] class, <-*gem*>,[36] thus, in this case, the gender assignment is transparent.
  [+feminine]: *viagem* 'journey', *imagem* 'image', *origem* 'origin'

- **Group 3** includes nouns with the non-default uniform word marker <-*dade*>[37]. This word marker is exclusively assigned to the [+feminine] gender value.
  [+feminine] : *sociedade* 'society', *cidade* 'city', *liberdade* 'freedom'

[34] However, in hindsight, stress might have had an effect on the performance of our participants, since both the control group and level C1 of the L2 Portuguese participants scored very low with Group 7, which included nouns ending in stressed vowels. Nevertheless, these results could also be due to the low frequency of some nouns used in this groups therefore more future research is needed in order to identify the causes (see Chapter 4 and 5).

[35] Some word markers are exclusively assigned to one gender value, such as Group 2, Group 3, Group 5, Group 8, Group 10, Group 11 and Group 12.

[36] Phonetically, these nouns end in [ẽʒ̃], therefore we need to emphasize once again that we relied on orthographic cues when creating our groups. Furthermore, we have to mention that Villalva (1994) refers to these nouns as products of derivational suffixes that carry [+feminine] gender information (see Chapter 2), but we chose to formulate a distinct group and treat is as a word marker based on its orthographic saliency.

[37] This word marker is also considered as a product of feminine derivational suffixation by Villalva (1994), but because of its orthographic saliency, we decided to form a separate group for it.

- **Group 4** comprises nouns with the non-default uniform word marker <-*r*>. This word marker can be assigned to both [+masculine] and [+feminine] nouns in Portuguese, thus nouns belonging to this group are opaque concerning their gender assignment.
  [+masculine]: *açúcar* 'sugar', *lugar* 'place', *jantar* 'dinner'
  [+feminine]: *flor* 'flower', *cor* 'color', *dor* 'pain'

- **Group 5** was created for the non-default uniform word marker <-*l*>, which only appears in [+masculine] nouns.
  [+masculine]: *papel* 'paper', *pastel* 'pastry', *jornal* 'newspaper'

- **Group 6** contains nouns with the non-default uniform word marker <-*s/z*>, which can be assigned to [+masculine] and [+feminine] nouns alike. We grouped these two orthographic cues (graphemes) together because phonetically they are realized as [ʃ] and the same sandhi rules apply to both of them (Mateus & d'Andrade 2000).
  [+masculine]: *lápis* 'pencil', *nariz* 'nose', *país* 'country'
  [+feminine]: *paz* 'peace', *luz* 'light', *cruz* 'cross'

- **Group 7** includes the non-default uniform word marker of words ending with the stressed vowels <-*é/á/ó*>. This group of word markers can be assigned to both [+masculine] and [+feminine] nouns, hence its opacity.
  [+masculine]: *café* 'coffee', *chá* 'tea', *pó* 'dust'
  [+feminine]: *chaminé* 'chimney', *fé* 'faith', *pá* 'dustpan'

- **Group 8** consists of nouns containing the non-default uniform [+masculine] word marker <-*a*>.[38] Their opacity comes from the fact that this very word marker coincides with the default theme index for feminine nouns, which is also <-*a*>.
  [+masculine]: *mapa* 'map', *problema* 'problem', *dia* 'day'

[38] These nouns have their origin in Ancient Greek.

- **Group 9** comprises nouns with the non-default uniform word marker <-*ão*>. This nasal diphthong can appear in [+feminine] and [+masculine] nouns alike.[39]
  [+masculine]: *coração* 'heart', *limão* 'lemon', *avião* 'airplane'
  [+feminine]: *mão* 'hand', *televisão* 'television', *tradição* 'tradition'

- **Group 10** was created for nouns with the non-default uniform word marker <-*im*>.[40] This word marker is solely assigned to [+masculine] nouns.
  [+masculine]: *pudim* 'pudding', *jardim* 'garden', *amendoim* 'peanut'

- **Group 11** includes the default uniform [+feminine] theme index <-*a*>, which, as opposed to the non-default uniform [+masculine] word marker <-*a*>, is only assigned to [+feminine] nouns.
  [+feminine]: *receita* 'recipe', *janela* 'window', *festa* 'party'

- **Group 12** consists of nouns with the default uniform [+masculine] theme index <-*o*>. All such nouns are [+masculine], very few exceptions aside.[41]
  [+masculine]: *vinho* 'wine', *sapato* 'shoe', *almoço* 'lunch'

All target nouns chosen for the experimental task had to meet the following criteria:

(i) All nouns had to be known by all participants, even by those with low levels of proficiency. Thus, we mainly used nouns from the vocabulary of *Português XXI Nível A1* (Tavares 2008) and *Grão a Grão…* (Csaba, Fodor & Szijj 2006), which are, perhaps, the most well known and most used textbooks for teaching Portuguese in Hungary.

[39] As for the gender assignment of nouns with this word marker, there is a very vague rule that L2 Portuguese learners are sometimes taught: those nouns whose referent is an abstract concept are usually [+feminine], whereas those whose referent is a tangible object are [+masculine]. This rule, however, has many exceptions, such as *televisão* 'television' which, when referring to the television set, is a tangible object, but it is [+feminine], thus this rule can be rather misleading.

[40] Once again, we have to state that we relied on orthographic cues (the written forms of the nouns) when creating our word groups, since phonetically this word marker is realized as [ĩ].

[41] Perhaps the only [+feminine] Portuguese word ending with *–o* is *tribo* "tribe".

(ii) No nouns could be [+animate]/variable, thus forcing our participants to rely on grammatical criteria instead of any semantic ones, i.e. our target nouns could not carry any information on the biological sex of their referents which could facilitate their gender assignment.

(iii) No nouns could be recent loanwords (e.g. *táxi* 'taxi', *gin* 'gin', *internet* 'internet').

(iv) The nouns could not be (hyphenated) compound words (e.g. *amor-perfeito* 'pansy', *couve-flor* 'cauliflower', *guarda-chuva* 'umbrella'), because the gender assignment of such nouns is subject to complex rules depending on the constituents of the composition.

(v) The nouns could not be abbreviated words used in colloquial speech such as *foto* for *fotografia* 'photography' or *mota* for *motocicleta/motociclo* 'motorbike', because the word ending in the abbreviated form might be misleading as to the gender value of the full word.

(vi) They could not be nouns with the non-default uniform [+feminine] theme index <–o>, because we only found one such word, *tribo* 'tribe'.

(vii) The nouns could not be non-default athematic nouns ending in <–*u/-eu*>. There are very few nouns with these endings and they are all [+masculine]. All such nouns are stressed on the last syllable. We did not include such nouns in a group, because of their low frequency and also because they might not have been known by our participants[42], aside from *museu* 'museum', *liceu* 'high school', *bacalhau* 'codfish', *chapéu* 'hat'. Furthermore, such nouns would not have constituted a homogenous group, since some of them end in a vowel *(<-u>)* and some end in a diphthong *(<-eu>).*

(viii) No noun could be a non-default athematic uniform noun ending in <-*i*>, since there are very few such nouns and most of them are loanwords or form the diphthong <-*ei*>, thus not constituting a homogenous group (e.g. *táxi* 'taxi', *rei* 'king', *trólei* 'trolleybus', *rubi* 'ruby',

[42] Some such examples are: *grau* 'grade', *fogaréu* 'an iron container elevated with rods used as a torch', *tatu* 'armadillo', *pau* 'stick', *tabu* 'taboo', *troféu* 'trophy', *pneu* 'tire', etc.

râguebi ‘rugby’, *pónei* ‘pony’, *piripiri* ‘chili pepper’, *penalti* ‘penalty’, *júri* ‘jury’, *javali* ‘wildhog’, *lei* ‘law’, *hóquei* ‘hockey’, *colibri* ‘hummingbird’, *boi* ‘ox’, *álibi* ‘alibi’, *abacaxi* ‘pineapple’[43]).

3.2.2. Determiners and adjectives

Portuguese determiners always precede the noun. Articles can be definite (*o* [+masculine, +singular], *os* [+masculine, +plural], *a*, [+feminine, +singular] *as* [+feminine, +plural], “the”) or indefinite (*um* [+masculine, +singular], *uns* [+masculine, +plural], *uma* [+feminine, +singular], *umas* [+feminine, +plural], “a/an”). Determiners always match the gender and number of the noun, constituting the DP together.

In our study, we only employed the articles mentioned above, since their definiteness or number did not interfere with gender assignment. The participants did not have to choose between singular or plural forms, therefore, if the target noun was in plural, we provided two articles in their plural forms in the drop-down options that still differed in their gender value (*os/as* or *uns/umas*).

In one of our sentences the definite article is also accompanied by a possessive pronoun located between the determiner and the noun, see in (3.1):

(3.1) Apanhei uma gripe, por isso o meu nariz está entupido.
Catch.1SG.PST DET flu for that DET POSS nose COP blocked
‘I caught the flu, therefore my nose is blocked.’

The possessives also agree in gender and in number with the noun. The following table contains all European Portuguese possessive forms (determiners and pronouns). The columns ‘Masculine/Feminine’ and ‘Singular/Plural’ refer to the properties of the noun which the possessives match in gender and number. Therefore, for example, if the noun is in its plural form, the possessive will take its plural form as well.

[43] Note that (with the exception of *lei* ‘law’) all nouns are [+masculine].

	Masculine	Feminine	Masculine	Feminine
	Singular		Plural	
1SG	meu	minha	meus	minhas
2SG	teu	tua	teus	tuas
3SG	seu	sua	seus	suas
1PL	nosso	nossa	nossos	nossas
2PL	vosso	vossa	vossos	vossas
3PL	seu	sua	seus	suas

Table 1: Portuguese possessives

When the target noun was preceded by a definite article and a possessive both were provided together in the dropdown options, matching the number of the noun, e.g. *o meu/a minha*.

Also, in one of our sentences (3.2) we had to include the preposition *de* 'of' which, if it precedes the article (or any other determiner, for that matter), fuses to it, creating *do(s), da(s), duma(s), dum/duns*.[44]

(3.2) O Pedro gosta das cidades antigas da tua terra (…)
DET Peter like.3SG of.DET cities old of.DET POSS country
'Peter likes the old cities of your country (…)'

For this sentence, we provided the already fused options agreeing in number with the noun, but giving both gender value options to choose from (*dos* or *das*).

In Hungarian, the rules for determiner selection are a little different. Hungarian definite articles occur based on a phonetic rule: if the noun starts with a consonant, *a* is selected (as in *a nap* 'the day/the sun'), whereas if the noun starts with a vowel, *az* is selected (as in *az élet* 'the life'). In the case of the Hungarian indefinite article, there are no phonetic rules, since there is only one, invariant numeral *egy* 'one, a, an' (*egy nap* 'a day/a sun', *egy élet* 'a life'). Moreover, Hungarian articles do not vary in number and are genderless, as in (3.3)

[44] This phenomenon also happens with the preposition *em* "in/on". It fuses with the articles (or other determiners) and creates *no(s), na(s), numa(s), num/nun(s), etc.*

(3.3) a. *a kutya* b. *a kutyák*
DET dog.SG DET dog.PL
'the dog' 'the dogs'

Since all of our target sentences are declarative, we did not employ interrogative pronouns. Demonstrative and indefinite pronouns were not used either, nor were numerals.

As for Portuguese adjectives, as we have seen in Chapter 2, they can be divided into two groups: there are some that admit variation according to gender value and there are some that do not (see more in section 2.2.2.2. and Villalva 1994).

The latter group consists of uniform adjectives which are either simple (non-suffixed) adjectives that end in <*–e, -z, –r*>, or athematic <*–s*> or <*–m*> or were created through derivational processes with suffixes such as <*–vel*> (e.g. *inteligente* 'intelligent', *eficaz* 'efficient', *anterior* 'anterior', *simples* 'simple', *ruim* 'awful', *imaginável* 'imaginable') or are invariant otherwise (*careca* 'bald') (Villalva 1994). This group of adjectives was not employed in our study, since it would have been impossible to deduce if the participants chose the correct gender value or not.

Similarly, we did not use adjectives with a null theme index that were created via the derivational suffixes <*-dor*> and <*–ês*> or adjectives that end with <*–ão/ã*>, because we aimed for the most transparent adjectives concerning gender assignment. Therefore we only used variable adjectives with the theme index <*–o*> that admit the <*–a/-o*> gender contrast like *claro/clara* 'clear', *direito/direita* 'right', etc. (Villalva 1994).

We also applied, in many cases and mainly in Type 2 sentences[45], past participles in the position of the adjective. These past participles are created through the suffixes <*–ado(s)/ -ada(s)*> from verbs with the thematic vowel <*–a*>, and <*-ido(s)/-ida(s)*> from verbs with the thematic vowels <*–e*> and <*–i*>, e.g. *molhado/molhada* 'wet', *escolhido/escolhida* 'chosen', *partido/partida* 'broken'. These past participles also agree in gender value and in number with the noun they modify, however, they occur in copulative constructions after copulative verbs like *estar* 'to be', *ser* 'to be', *ficar* 'to be, to stay, to become', *tornar* 'to become', etc.

[45] Type 2 sentences are sentences with the format [DP+COP+ADJ]. See more in section 3.2.3.2. below.

In Portuguese, the default position for adjectives is to the right of the noun, but a restricted class of them (namely, predicative adjectives with a valorative meaning) can also appear before the noun, as in (3.4):

(3.4) a. Um amigo **fiel**.
DET friend loyal
'a loyal friend'

b. Um **fiel** amigo
DET loyal friend
'a loyal friend'

In the present study, for ease of exposition and because word order was not at stake, all adjectives are located to the right of the noun in all target sentences. In Hungarian, adjectives always precede the nouns, unless there is a copula in the sentence, in which case they will follow the noun, as in (3.5). Another important attribute of those Hungarian adjectives that precede the noun is that they are invariable, meaning that they do not agree with the noun in number and they are also genderless, similarly to nouns.

(3.5) a. A kék egek
DET blue.SG sky.PL
'the blue skies'

b. Az ég kék volt.[46]
DET sky blue COP.PST.IND
'The sky was blue.'

We will address the topic of adjectives further in section 3.2.3.2. when considering Type 2 sentences.

3.2.3. Types of sentence structures

One of the aims of this study was to investigate whether the distance between the noun and the adjective has an effect on the performance of L2 Portuguese speakers with regard to the correct gender assignment of the adjective and whether it has any such effect on the determiner as well.

This question has already been raised and investigated in the research field of psycholinguistics. According to Rourke & Van Petten 2011:3 "[the] successful resolution of long-distance dependencies has long been identified as troublesome for readers and listeners due

[46] In such copulative sentences, the copulas do not occur in present tense (indicative) in Hungarian.

to the burden placed on working memory." Brysbaert & Mitchell (2000) claim that during initial sentence parsing, most readers do not succeed at using gender agreement information. Thus we shall investigate whether distance is a factor in L2 Portuguese grammatical gender agreement in language productions from L1 speakers of a genderless language, Hungarian

In order to establish a relation between the selection of gender value and distance from the noun (thus, to be able to investigate Hypothesis 2) this study employed three types of sentence structures where, from Type 1 to Type 3, the adjective is placed increasingly further away from the noun. In Type 1 sentences, the adjective is located immediately after the noun, thus these sentences contain attributive adjectives. In Type 2 sentences, the adjective is separated from the noun by a copula, forming copulative constructions. In Type 3 sentences, there is a relative clause between the antecedent noun and the adjective where the subject of this relative clause is a noun that has the opposite gender value to the antecedent of the relative.

3.2.3.1. Type 1 sentences

Type 1 sentences contain the following clausal word order: [DET+N+ADJ], irrespectively of the function of the phrase itself (subject or object). In all Type 1 sentences the adjective immediately follows the nouns, as in (3.6):

(3.6) [O açúcar branco] não é muito saudável.
DET sugar.N white. ADJ not is very healthy
'White sugar is not very healthy.'

In many languages, like in English or Portuguese, adjectives can appear in the attribute (a.) or in the predicate (b.) position, see (3.7) and (3.8) below:

(3.7) a. The **blue** sky (English)
b. The sky is **blue**.

(3.8) a. O céu **azul** (Portuguese)
b. O céu é **azul**.

Portuguese is a Subject-Verb-Object (SVO) language, such as Hungarian, however the latter is a topic-prominent language, therefore the word order is more fluid and can change according to which constituent is emphasized (Mateus *et al*. 2003; É. Kiss, Kiefer & Siptár 2003).

In Type 1 sentences, adjectives always appear as attributes, thus are attributive adjectives and not predicative expressions, since there is no copula involved in these constructions (see more on this topic in section 3.2.3.2.). Table 2 below lists all the used attributive adjectives in our experiment. The words in bold in the left columns are the target nouns – to their left are located the determiners and to their right are the adjectives. The English translations are given in the right column.

uma **noite** fria	a cold night
um **filme** italiano	an Italian movie
uma **viagem** longa	a long journey
(d)as **cidades** antigas	the old cities
a **cor** favorita	the favorite color
o **açúcar** branco	the white sugar
uns **pastéis** deliciosos	some delicious pastries
um **lápis** vermelho	a red pencil
a **luz** branca	the white light
uma **chaminé** suja	a dirty chimney
um **café** fraco	a weak coffee
um **mapa** moderníssimo	a very modern map
um **coração** pequeno	a small heart
a **mão** esquerda	the left hand
o **jardim** botânico	the botanic garden
uma **receita** fabulosa	a fabulous recipe
o **sapato** direito	the right shoe

Table 2: Type 1 target attributive adjectives

3.2.3.2. Type 2 sentences

Type 2 sentences consist of a [DP+COP+ADJ] structure, which is also known as a predicative clause. In these constructions all verbs are copulas and they are located to the immediate right of the noun. Therefore, compared to Type 1 sentences, the adjective is pushed a

little further away from the noun (the noun being preceded by the determiner) with the copula occurring between the DP and the adjective.

(3.9) [O limão está maduro] demais.
DET lemon. N COP ripe.ADJ too much
'The lemon is too ripe.'

As it is well known, verbs and nouns are the two major lexical categories present in all languages. Adjectives, on the other hand, are not that common in languages for they can semantically behave as either nouns or verbs in certain languages. Therefore they may be hybrid in nature and can be perceived as an intermediate class between nouns and verbs (Croft 1991, Pustet 2003).[47] Nevertheless, in languages that have all three lexical categories (like Portuguese and Hungarian), nouns, verbs and adjectives are all bound to prefer to fulfill a specific syntactic function in clausal context. Nouns usually act as arguments, verbs usually behave as predicates and adjectives usually function as attributes (Pustet 2003).

In our Type 2 sentences we used a special type of verbs called copulas. Copulas are considered to be meaningless, or in other words, semantically empty intransitive verbs (Pustet 2003). Copulas can only be found in predicate constructions therefore they serve a syntactically relevant purpose. According to Hengeveld (1992) and Stassen (1997), these functions could be the following:

(i) the copula acting as a linker between subject and predicate (which is exactly our case);
(ii) the copula acting as a syntactic 'hitching post' or 'dummy' to which verbal inflectional categories can be attached;
(iii) the copula acting as a predicator which is added to lexemes that do not form predicates on their own (Pustet 2003:3).

[47] Dyirbal, for instance, is a language where lexemes that express adjectival concepts can hardly be distinguished from nouns in their morphosyntactic behavior. Also, Mandarin treats adjectives as a verbal subclass (Pustet 2003).

Portuguese, such as other Romance languages, has a multi-copula system. *Estar* (‘to be’) is used with adjectival items that carry a change or transition in meaning. Furthermore, there is set of lexicalized participles and adjectives derived from participles that are restricted to this copula (Pustet 2003).

(3.10). a. Ele está cheio/preso/vazio/etc.
3SG COP.PRS.IND.3SG full/captive/empty/etc.
‘He is full/captive/empty/etc.’

b. A porta está aberta.
DET door COP.PRS.IND.3SG open.
‘The door is open.’

Ser (‘to be’) is used with adjectival items that do not carry a meaning of change or transition.

(3.11) O gato é branco.
DET cat COP.PRS.IND.3SG white.
‘The cat is white.’

This ‘permanent vs. non-permanent’ distinction is to be regarded as the general principle underlying copula alternation (Pustet 2003).

In Hungarian, the copula *van* ‘to be’ (present, indicative, 3SG) “is obligatorily deleted in third person [singular and plural] present indicative both with nominals and adjectivals in predicate position” but always occurs in past indicative (Pustet 2003:67).

(3.12) a. Ez egy szék.
DEM DET chair
‘This is a chair.’

b. Ez egy szék **volt.**
DEM DET chair COP.PST.IND
‘This was a chair.’

(3.13) a. Ez a történet hosszú.
DEM DET story long
'This story is long.'

b. Ez a történet hosszú **volt.**
DEM DET story long COP.PST.IND
'This story was long.'

In Portuguese, the copula can never be dropped (as in (3.16)).

(3.14) a. Esta **é** uma cadeira
DEM COP.PRS.IND DET chair
'This is a chair.'

b. Esta **foi** uma cadeira.
DEM COP.PST.IND DET chair
'This was a chair.'

(3.15) a. *Esta uma história longa
DEM DET story long
'This story is long.'

b. *Esta uma história longa.
DEM DET story long
'This story was long.'

The Portuguese copulas we most often used in the target sentences were *estar* and *ser* 'to be'. Other than these, we also used the verb *ficar* 'to become, to remain', defined by Hengeveld (1992) and Pustet (2003) as a semi-copula, which "[occupies] an intermediate position between copulas and full verbs in that they show similarities with both lexeme types" (Pustet 2003:6). The main difference between a copula and a semi-copula is that the latter does convey meaning, whereas the copula does not.

Copulas, as we have mentioned, only appear in predicative constructions, being followed by predicate phrases, which can be of different categories, such as adjectives, nouns, prepositions, adverbs and sentences. The purpose of the predicative elements is to express a property that is assigned to the clause subject. Predicative expressions can be of two types, generally speaking: predicative nominals or predicative adjectives (Burton-Roberts 1997, Radford 2004).[48]

Table 3 below lists of all the predicative expressions we used in our research, where the underlined words in the left column are the target nouns and in bold are the copulas. Situated to

[48] Furthermore, predicative expressions are not attributive adjectives or arguments or adjuncts. Nevertheless, overlaps can occur, but this phenomenon will not be discussed in this study.

the immediate left of the nouns are the determiners and to the right of the copulas are the adjectives agreeing in gender with the nouns they predicate on. The right column contains the English translation of the target predicative expressions.

uma árvore **ficou** partida	a tree became/got broken
o leite **estava** estragado	the milk was spoiled
a imagem **ficou** tremida	the image became/got blurry
a liberdade **foi** garantida	the freedom was guaranteed
a dor **era** demasiada[49]	the pain was too much
o lugar **foi** escolhido	the place was chosen
o papel **ficou** molhado	the paper became/got wet
o meu nariz **está** entupido	my nose is blocked
a paz **ficou** estabelecida	the peace was established
a pá **ficar** cheia	the dustpan became/got full
o chá **ficou** frio	the tea became/got cold
o problema **era** sério	the problem was serious
o limão **está** maduro	the lemon is ripe
a televisão **ficou** avariada	the television became/got broken
o pudim **está** feito	the pudding is done
a janela **está** embaciada	the window is foggy
o vinho **está** aberto	the wine is open

Table 3: Type 2 target predicative expressions

3.2.3.3. Type 3 sentences

Type 3 sentences consist of a [DET+N+REL+ADJ] structure. Of all our research sentences, in a linear order, the adjectives are the furthest away from the nouns in this type. Here, the relative clause includes a [+human] subject[50] with the opposite gender value to the antecedent

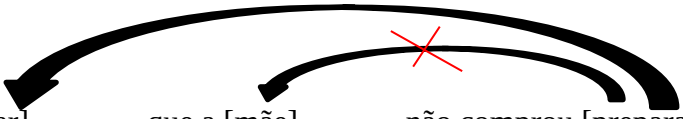
[49] Even though our Portuguese native speaker control group found this construction to be grammatical, we have to mention that in principle *demasiado/demasiada* 'too much' is not an adjective but an intensifier that may require an attribute, like *demasiado intenso* 'too intense'.

[50] The reason why we used [+human] nouns in this case is because we suspected that a foreign language learner may feel that the semantic value of [+human] nouns is stronger/more salient than the gender values of [-human] nouns. Therefore the contrast between the gender values of the subject of the relative clause and the target noun of the main clause was highlighted. See more on this topic in Mariotto (2011).

of the relative clause, which the adjective has to agree with in gender, thus creating an intervention effect (Friedmann, Belletti & Rizzi 2009).[51] The rationale for this type of sentence is that, at least on the surface, the adjective is far away from the noun it modifies and, on top of that, there is a noun closer to it that can serve as its potential modifee, as in configuration (3.16):

(3.16) $N_{1[\alpha\ \text{masculine}]}$ … $N_{2[\beta\ \text{masculine}]}$… $ADJ_{[\alpha\ \text{masculine}]}$ …

Considering (3.16) above, N_1 and ADJ cannot meet a local relation and the intervener N_2 is a potential candidate for it. See, specifically, the example in (3.17), where *jantar* 'dinner' is [+masculine] and *mãe* 'mother' is [+feminine]:



(3.17) [um [jantar]$_{[+\text{masculine}]}$ que a [mãe]$_{[+\text{feminine}]}$ não comprou [preparado]$_{[+\text{masculine}]}$

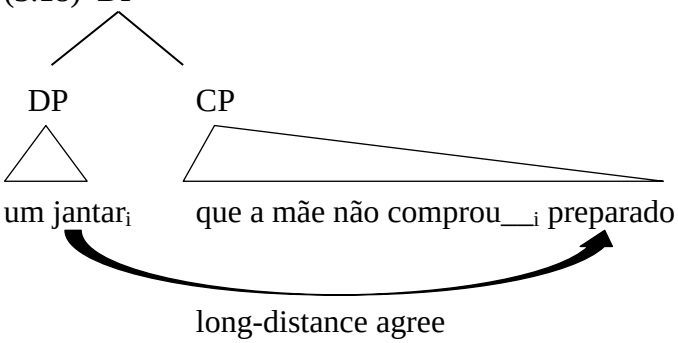
DET dinner REL DET mom not buy.PST prepared.ADJ

'a dinner that mom did not buy prepared'

This sentence type relies on the long-distance agreement between the head of the relative clause (e.g. *um jantar* [+masculine] 'a dinner' in (3.17)) and the adjective (*preparado* [+masculine] 'prepared') inside the relative. Specifically, if we assume that the relative clause is adjunct to the right of the head noun (e.g. Brito 1991, for European Portuguese), as in (3.18), the agreement relation between this noun, which is external to the relative clause, and the adjective (internal to the relative and occurring after the syntactic variable – a gap) cannot be local, but a long-distance one, as (3.18) illustrates.

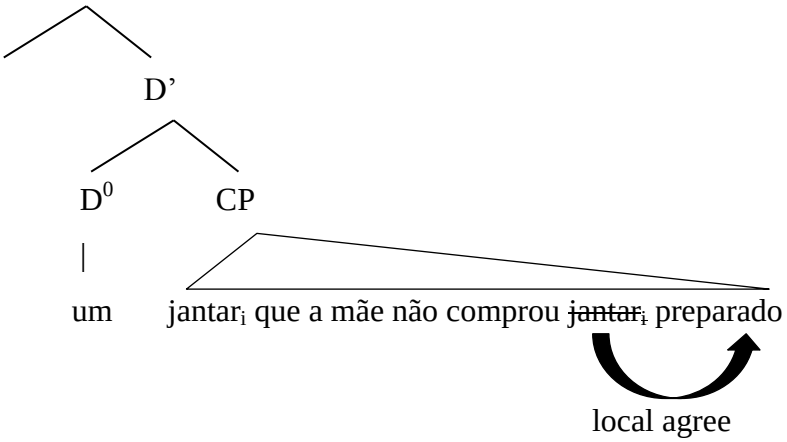
[51] With this sentence type, we aim to investigate not only the distance between the noun and the adjective but also to check whether a noun with a different gender value occurring between the target noun and the adjective has any influence on the selection of the gender value of the adjective.

(3.18)



However, if we take the derivation of relative clauses to be that of Kayne's (1994) – a head raising analysis – then the relation between the noun *jantar* and the adjective *preparado*, in (3.19), may be local before the head noun moves to the specifier position of the relative clause, the adjective checking first its gender against the noun before it raises, as in (3.19).

(3.19)



This will be further developed in Chapter 4 when discussing the results of the data analysis.

The table below contains all target Type 3 structures. The underlined words in the left column are our target nouns which are, consequently, the relativized nouns that are found before the relative pronoun and with which the adjective has to agree in gender value. The words in bold are the intransitive verbs we used, all of them being transitive. The italicized items in the left column are the subjects of the relative clause which intervene between the head/antecedent of the relative clause and the adjective. The target adjectives are located to the right of the verbs

because they modify the object nouns that were relativized. The right column contains the (rather literal) translation of these constructions.

a chave que o *advogado* **achou** __ pequena demais	the key that the lawyer found too small
um tomate que a *menina* **comprou __** estragado	a tomato that the girl bought rotten
uma origem que o *aluno* **não achou__** clara	an origin that the student did not find clear
a sociedade que os *políticos* **julgam__** desenvolvida	the society that the politicians consider developed
uma flor que o *Vítor* **não acha** __ linda	a flower that Victor does not find pretty
um jantar que a *mãe* **não comprou** __ preparado	a dinner that mom did not buy prepared
um jornal que a *Catarina* **não achou __** chato	a newspaper that Catherine did not find boring
um país que a *tia* **acha __** belíssimo	a country that (my) aunt finds gorgeous
a cruz que o meu *tio* **tinha encontrado** __ partida	the cross that my uncle had found broken
uma fé que o *professor* **julgou** __ antiquíssima	a religion that the teacher considers ancient
um pó que a *professora* **achou __** espesso	a dust that the teacher found thick
um dia que a *Luísa* **declarou __** péssimo	a day that Luisa declared horrible
uma tradição que o *avô* **não acha** __ adequada	a tradition that (my) grandfather does not find adequate
o avião que as *notícias* **anunciaram** __ perdido	the airplane that the news had announced lost
o amendoim que a *Sílvia* **tornou** __ triturado	the peanut that Silvia turned/ had ground
uma festa que o meu *patrão* **achou** __ fantástica	a party that my boss found fantastic
um almoço que a minha *namorada* **achou** __ delicioso	a lunch that my girlfriend found delicious

Table 4: Type 3 target constructions with relative clauses

3.3. The hypotheses of the study

Hypothesis 1:

The increase in the rates of success of our L2 Portuguese participants is directly proportional to the increase in their levels of proficiency.

In other words, the more training these participants have received, the better their results

are going to be (a progression effect). Therefore we expect level A2 to have the lowest scores, level B2 to have higher scores and level C1 to have the highest scores.

Hypothesis 2:

The increase of the distance between the noun and the determiner is inversely proportional to the rates of success in gender assignment of our L2 Portuguese participants.

In other words, the further the adjective is away from the noun it must agree with, the lower the success rates in gender assignment are going to be. Thus we expect Type 1 sentences (local agree construction, attributive adjectives, see in section 3.2.3.1.) to have the highest scores, Type 2 sentences (copulative constructions, see in section 3.2.3.2.) to have lower scores and Type 3 sentences (surface/linear long-distance agree construction with an intervener between the head noun and the adjective, see in section 3.2.3.3.) to have the lowest scores.

Hypothesis 3:

The noun groups comprising the default theme indices (Group 11-12, see section 3.2.1) will have higher rates of success in gender assignment than the noun groups consisting of non-default word markers (Group 1-10, see section 3.2.1.).

We expect the default theme indices to be acquired and mastered before the non-default word markers. Since the default theme indices *<-o> ([+masculine])* and *<-a> ([+feminine])* possess the most transparent cues for gender value assignment, these are going to be the best controlled noun groups compared to the non-default word markers that are characterized by more opaque and ambiguous gender value assignments.

Hypothesis 4:

Assuming transfer from L1 to L2, we predict that our L2 Portuguese participants will have higher success rates in gender assignment with adjectives compared to determiners.

We expect that, based on the characteristics of the L1, our participants will assign more correct gender values to adjectives than to determiners. Since in Hungarian the case marking is in the end of the noun, we assume that the post-nominal position (the position of the adjective in Portuguese, (N+ADJ)) is more salient in the acquisition of L2 Portuguese gender agreement.

Hypothesis 5:

Assuming transfer from L2 to L3, we predict that those participants who, apart from Portuguese, learned at least one other Romance language will have higher rates of success compared to those participants who did not learn any other Romance language.

We expect that even if the L1 (Hungarian) does not contain any gender cues and has its value of gender feature fixed for [-gender] (as the FFFH claims, see Chapter 1), previous training in other Romance languages, which are morphosyntactically similar and genetically linked to Portuguese, facilitates the acquisition of Portuguese gender agreement, be it through transfer from this/these other Romance language(s) or via the application of other psycholinguistic processes (Selinker 1972, see Chapter 1) already acquired for the other Romance language(s).

Hypothesis 6:

Those L2 Portuguese participants who not only studied Portuguese in a classroom context but were also immersed in Portuguese language and culture for at least 3 months will have higher success rates than those L2 Portuguese participants who exclusively studied Portuguese in a classroom context.

We expect that those participants who lived in an immersion context received more PLD (i.e. naturalistic input) from L1 speakers and also made an 'investment' according to Peirce (1995) (see Chapter 1), therefore they were more successful at language acquisition than those participants who had no access to input from L1 Portuguese speakers and who did not have the need to increase their 'cultural capitals' (Peirce 1995, see Chapter 1)

Chapter 4
Data analysis

As we have mentioned in the previous chapter, our study comprises 36 adult Hungarian native speakers who are L2 speakers of Portuguese (all of them university students) and 30 adult native speakers of European Portuguese for the control group.

Our analysis of the collected data is based on the rates of success, i.e. the correct answers of our participants. We established a binary system in our database where every correct answer was valued as 1 and every incorrect answer was valued as 0 (see Appendix 5, for example). These values are the focus of the analysis put forward in this chapter.

Before proceeding to the investigation of the results of our L2 Portuguese speaker participants, let us begin by presenting our findings with respect to the results of the control group formed by Portuguese L1 speakers. The results of the control group shall be considered as the native speaker level. Thus the data of the control group shall serve as a basis of comparison for our study, to which we will compare the performances of the L2 Portuguese speakers.

4.1. Results of the control group

4.1.1. Overall Scores

Table 5 below presents the results of the control group with regard to the overall achieved correct answers, i.e. correct answers for determiners and adjectives combined:

Control Group Overall Scores	(%)	Maximum	(%)
3049	99,64	3060	100
Control Group Average Overall Scores	**(%)**	**Maximum**	**(%)**
101,63	99,64	102	100

Table 5: Control Group: overall scores

The numbers in Table 5 show us that Portuguese L1 speakers make very few mistakes when it comes to gender agreement. In total, the control group had 11 (0,36%) incorrect answers,

which strengthens Villalva's (1994) view, according to whom Portuguese nominal gender assignment is lexically conditioned, therefore is a formal category (see Chapter 2).

If we separate the data on determiners and adjectives from the combined results we find the following scores, presented in Table 6 and 7, respectively:

Control Group DET Score	(%)	Maximum	(%)
1528	99,87	1530	100
Control Group Average DET Score	**(%)**	**Maximum**	**(%)**
50,93	99,87	51	100

Table 6: Control Group: scores on determiners

Control Group ADJ Score	(%)	Maximum	(%)
1521	99,41	1530	100
Control Group Average ADJ Score	**(%)**	**Maximum**	**(%)**
50,70	99,41	51	100

Table 7: Control Group: scores on adjectives

If we compare Table 6 to Table 7, we can see that the control group scored slightly better (by 7 more correct answers, which is 0,46%) with determiners than with adjectives. This phenomenon might be due to several factors. Firstly, in the configuration of the target sentences of this study the distance between the determiner and the noun is always constant, which means that the determiners always immediately precede the nouns (see Chapter 3) and the (surface/linear) distance between the noun and the adjective varies by the three sentence types investigated in this study. Secondly, according to Carroll (1989) and MacWhinney (1992), native speakers treat determiners as part of the noun (i.e. as a syntactic unit) in the first stages of acquisition of their L1[52] and only later do they realize that they are two separate items which then triggers the establishment of the gender feature. This is called 'top-down constructive processing', where the L1 learners use 'protodeterminers' (Liceras, Díaz & Mongeon 2000). Furthermore, as Franceschina (2005) established for the acquisition of L1 Spanish "gender

[52] For more information, see Freitas, Alves & Costa 2007; Gonçalves, Guerreiro, Freitas & Sousa 2011.

agreement between the article and the noun is mastered before the agreement between the adjective and the noun (e.g., Dato 1975, González 1970, Hernández-Pina 1984, López Ornat 1997, Schnell de Acedo 1994, Soler 1984, Tolbert 1978)" (2005:113).

Therefore these configurational and developmental factors could explain why our control group performed better, however minimally, with determiners than with adjectives.

4.1.2. Overall scores on types of sentences

The following table (Table 8) details the performance of the control group with respect to the three types of sentences we used in our study, described in Chapter 3.

Control Group Scores (%)		
Type 1 Overall	**Type 2 Overall**	**Type 3 Overall**
99,80	99,71	99,41
Type 1 DET	**Type 2 DET**	**Type 3 DET**
99,80	99,80	100
Type 1 ADJ	**Type 2 ADJ**	**Type 3 ADJ**
99,80	99,61	98,82

Table 8: Control Group: scores by sentence types

If we take a look at the overall results for each sentence type, we can notice a decreasing tendency in success rates from Type 1 to Type 3 sentences. This might indicate that L1 speakers find it increasingly difficult to assign the correct gender value to the adjective as the distance between it and the noun increases.

If we compare the performances on determiners and adjectives broken down to the three sentence types, we acquire a more detailed picture of the processing of the sentence types. As we can see, determiners received higher scores with Type 2 and Type 3 sentences than adjectives, and the scores were equal with both items with Type 1 sentences. Both Type 1 and Type 2 determiners had a 99,80% success rate and Type 3 determiners received a perfect score. However, the control group found the assignment of correct gender values to adjectives most difficult with Type 3 sentences (98,82% success rate), where the adjectives were the furthest away from the noun and there was an intervener noun in the relative clause with the opposite

gender value to the head noun. The second lowest scores with adjectives were achieved with Type 2 sentences (copulative constructions) and the highest scores with Type 1 sentences (attributive adjectives).

These data indicate that the assignment of gender values tends to be minimally more accurate with determiners and that the distance between the noun and the adjective and the type of syntactic construction seem to have a minimal effect on the performance of native speakers, based on which we formulated Hypothesis 4 and 2, respectively. We shall test these hypotheses on our L2 Portuguese speaker participants.

4.1.3. Scores on noun groups

In this section we shall examine the results of the control group with respect to the twelve established noun groups (see more in Chapter 3). Table 9 below contains the combined data for determiners and adjectives broken down to noun groups:

Overall Scores (%)											
Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
100	99,44	98,89	99,44	100	99,72	99,17	100	99,44	100	100	100

100%
>99,5-100%
>99-99,5%
98-99%

Table 9: Control Group: overall scores broken down to noun groups

To be able to analyze the data on nouns groups, we established a color-coding system where a different color is assigned to the results of each degree of a four-degree scale.

The lowest scores that range from 98-99% and are assigned to the light grey color, only appear with Group 3 nouns, within which *sociedade* 'society' achieved the least correct answers (96,67% success rate, see Appendix 10). This might not have happened because of the difficulty of the noun itself, but because of that fact that it appeared in a Type 3 sentence, which, as we have seen previously, proved to be the most problematic sentence type, see (4.1) below:

(4.1) Esta é a **sociedade** que os políticos julgam **desenvolvida**, mas na realidade não é bem assim.
'This is the society that the politicians deem developed, but in reality, this is not exactly true.'

The second lowest degree of the scale that ranges from >99-99,5% is assigned to the blue color. It appears with Group 7 which comprises the word marker of stressed vowels and the most problematic word was *chaminé* 'chimney' with a 96,67% success rate (see Appendix 10), perhaps because of its low frequency[53]. Furthermore, Group 2 and 9 also received results within this range, all of them scoring the lowest results with the noun that appeared in a Type 3 sentence (*origem* 'origin' 98,33% and *tradição* 'tradition' 96,67%, see Appendix 10). Finally, the results for Group 4 also fall within this range, where *lugar* 'place', the noun that appeared in a Type 2 sentence, had the lowest score (98,33%, see Appendix 10). Therefore, once again, the type of sentence seems to pose the real problem, not the word marker or type of noun.

We obtained one result within the >99,5-<100% range, appearing with the purple color, and it was with Group 6 (word marker <-*s/z*>), where, again, a noun that was used in a Type 3 sentence received the fewest correct answers (see Appendix 10).

The results that appear with the red wine color received perfect (100%) scores and include four groups with non-default word markers: Group 1 (word marker <-*e*>), 5 (word marker <-*l*>), 8 (the [+masculine] word marker <-*a*>), 10 (word marker <-*im*>), and also the groups of the two default theme indices (Group 11 and 12). This might indicate that nouns with these word markers are the best dominated and their gender values are still easily transmitted to its modifees, even in a long-distance agreement construction.

If we take a look at Table 10 and Table 11 below, we can find the comparison of performances between determiners and adjectives broken down to nouns groups:

[53] To confirm this claim, we carried out a search in the CRPC (Corpus de Referência do Português Contemporâneo) online database created by the CLUL (Centro de Linguística da Universidade de Lisboa) (http://alfclul.clul.ul.pt/CQPweb/portugalonlyv23/).
The query for *chaminé* returned 442 matches in 297 different texts, in 289,840,619 words (332,332 texts) which indicates a frequency of **1.52** instances per million words.
On the other hand, if we compare these data with those on a noun that has a default theme index, like *casa* 'house', the results are the following: 89,498 matches in 44,522 different texts, in 289,840,619 words (332,332 texts); indicating a frequency of **308.78** instances per million words.
Furthermore, if we compare the previous results to that of a noun that also belongs to the group of non-default word marker of stressed vowels, like *pé* 'foot', we find 17,708 matches in 12,031 different texts (in 289,840,619 words [332,332 texts]); which indicates a frequency of **61.1** instances per million words. Therefore we can see that *chaminé*, in fact, is a very infrequently used noun in European Portuguese.

DET Scores (%)											
Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
100	100	100	99,44	100	100	99,44	100	100	100	100	100

Higher than ADJ Score
Equal to ADJ Score
Lower than ADJ Score

Table 10: Control Group: scores on determiners broken down to noun groups

ADJ Scores (%)											
Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
100	98,89	97,78	99,44	100	99,44	98,89	100	98,89	100	100	100

Higher than DET Score
Equal to DET Score
Lower than DET Score

Table 11: Control Group: scores on adjectives broken down to noun groups

The tendency that determiners are assigned slightly more accurate gender values can be observed in the overall results (Table 5) and it is also visible if we compare Table 10 with Table 11. As we have seen, the nouns that received the lowest scores are those that were used in Type 3 sentences (*sociedade* 'society', *origem* 'origin', *tradição* 'tradition', see Appendix 10), therefore their long distance from the noun and the presence of the intervener noun with the opposite gender value to the head noun in the relative clause might be the culprits behind this particular issue.

4.2. Data analysis of the L2 Portuguese participants

In this section, we shall analyze the data obtained from our L2 Portuguese speaker participants and based on these data, we shall attempt to prove or disprove our hypotheses, established in Chapter 3.

4.2.1. Overall data

Table 12 below contains the overall scores (correct answers for determiners and adjectives combined) of our Hungarian participants. The data are broken down to each investigated level of proficiency.

Level	Overall Score	(%)	Maximum	(%)
A2	969	79,17	1224	100
B2	1118	91,34	1224	100
C1	1108	90,52	1224	100
Total	3195	87,01	3672	100
Level	**Average Overall Score**	**(%)**	**Maximum**	**(%)**
A2	80,75	79,17	102	100
B2	93,16	91,34	102	100
C1	92,33	90,52	102	100
Total	88,75	87,01	102	100

Table 12: L2 Portuguese participants: overall obtained scores

First of all, after examining Table 12, we can conclude that the overall results show lower scores than 99,64% which was the overall result of the control group (see Table 5), meaning that

our L2 Portuguese participants did not attain the native-speaker level. However, both level B2 and C1 scored over 90% which renders their performance similar and still highly accurate on this issue. Even level A2 had results that are very close to 80% accuracy, therefore the overall results show that, in general, the L2 speakers have most probably mastered the gender values of most nouns in question.[54]

When we compare the results between the levels of proficiency, there is an observable 12,17% increase in success rates from level A2 to level B2, however our subjects on level B2 performed slightly better than those on level C1. This means a difference of 10 (0,82%) more correct answers on level B2 than on level C1. This regress might be due to a training effect, namely, that level C1 learners are already being trained in other areas of grammar, while level B2 learners might still be concentrating on the issue of gender values and gender agreement.

Table 13 shows that, if we separate the scores of all correctly chosen target-determiners from the overall scores, we obtain the same tendency of initial progress than slight regress:

Level	Overall DET Score	(%)	Maximum	(%)
A2	489	79,90	612	100
B2	563	91,99	612	100
C1	555	90,69	612	100
Total	1607	87,53	1836	100
Level	**Average DET Score**	**(%)**	**Maximum**	**(%)**
A2	40,75	79,90	51	100
B2	46,92	91,99	51	100
C1	46,25	90,69	51	100
Total	44,64	87,53	51	100

Table 13: L2 Portuguese participants: overall scores on determiners

There is a 12,09% increase in success rates from level A2 to level B2. Similarly to the combined data findings, level B2 participants outscored level C1 participants in the correct

[54] This reinforces the idea that some areas of grammar are easier to acquire in an L2 than others, gender being easier to master (perhaps because of its lexical nature, see Villalva 1994 and Chapter 2 of this dissertation). Also, the lack of native-likeness indicates the validity of Slabakova's (2008) Bottleneck Hypothesis, according to which morphology cannot be fully acquired in an L2 (see Chapter 1).

assignment of determiners by 1,3%.[55] However, the performance of both level B2 and C1 is over 90% and very similar. Again, this might be due to the aforementioned training effect.

Also, if we examine the rates of success of the target adjectives separated from the overall data, the tendencies remain the same, however the difference of success rates from level B2 to C1 becomes lesser, as Table 14 shows.

Level	Overall ADJ Score	(%)	Maximum	(%)
A2	480	78,43	612	100
B2	555	90,69	612	100
C1	553	90,36	612	100
Total	1588	86,49	1836	100
Level	**Average ADJ Score**	**(%)**	**Maximum**	**(%)**
A2	40,00	78,43	51	100
B2	46,25	90,69	51	100
C1	46,08	90,36	51	100
Total	44,11	86,49	51	100

Table 14: L2 Portuguese participants: overall scores on adjectives

From level A2 to level B2 there is a 12,26% increase in correct assignment of gender values, however level B2 only outperformed level C1 by 0,33%, which, again, means that the performance of these two levels was rather similar, pointing to a very high competence.

When comparing the scores on determiners with those on adjectives, it is visible that level A2 assigned correct gender values to 1,47% more determiners than adjectives. The same tendency shows on the other two levels as well: level B2 was 1,3% more successful with determiners than with adjectives and level C1 had a 0,33% better score on determiners than on adjectives. The total difference on scores between determiners and adjectives on the three levels combined is 1,04%. In sum, this means that our participants managed to assign slightly more correct gender values to the determiners than to the adjectives.

[55] Although the difference between level B2 and C1 does not seem relevant, one possible explanation for this result is that the IL of B2 learners of L2 Portuguese is more like the target grammar because of a training effect. Note that at this stage, learners are still reviewing and reinforcing some of these grammar competences, whilst C1 learners are usually focusing on complex structures and other grammar topics. This may lead to a regression effect on some more basic topics of structural competence; such as gender agreement (for more information see Tarone 2006).

This small difference may be explained by the fact that determiners have a scope over the nouns, favoring the agree relation, while adjectives do not. However, this can also be due to a training effect: L2 learners often times have to memorize lists of nouns, where the nouns appear together with a determiner. Since the determiner always immediately precedes the noun in Portuguese, it might pose less of a difficulty to L2 learners. Furthermore, as Franceschina (2005) put it, studies of "adult SLA (e.g., Bartning 2000, Bruhn de Garavito & White 2000, Chini 1995, Dewaele & Véronique 2001, Fernández-García 1999, Finnemann 1992) show that [DET+N] agreement is overall more accurate than [N+ADJ] or [ADJ+N] agreement, as in L1 development" (2005:114).

4.2.2. Overall data on types of sentences

We shall examine in this section how our participants performed with respect to the investigated types of sentences. As it was described in the previous chapter, we employed 3 different types of sentence structures and established 12 different groups of nominal word markers for which we chose 51 target nouns in total (17 for each sentence type).

Let us begin with our data collected on Type 1 sentences where the noun and the adjective are adjacent and there is no intervener between the noun and the adjective (the distance between the determiner and the noun is constant in all sentence types, as we have seen in Chapter 3). Table 15 below sums up the overall performance of our participants for this sentence type:

Level	Type 1 Overall Score	(%)	Maximum	(%)
A2	332	81,37	408	100
B2	365	89,46	408	100
C1	372	91,18	408	100
Total	1069	87,34	1224	100
Level	**Type 1 Average Overall Score**	**(%)**	**Maximum**	**(%)**
A2	19,53	81,37	24	100
B2	21,47	89,46	24	100
C1	21,88	91,18	24	100
Total	62,88	87,34	72	100

Table 15: L2 Portuguese participants: overall scores on Type 1 sentences

If we examine the scores for Type 1 sentences, it is obvious that there is growth in rates of success from level A2 to C1, showing a gradual learning progress from level A2 through B2 to C1. Specifically, from level A2 to level B2 the increase is by 8,09%, but from B2 to C1 it drops to only 1,72%; thus rendering the difference between the two higher levels of proficiency quite minimal, probably because the great leap in learning happens around level B1/B2, where a significant amount of vocabulary is introduced and, more importantly, trained in some real texts, thus consolidating the knowledge about these words.

When comparing the results of the L2 Portuguese participants to those of the control group, it is apparent that the L2 Portuguese participants did not reach native-like results, which would be 99,80% or more. However, level C1, with scores over 90%, does seem to perform in a highly proficient manner.

Let us now see what happens if we separately examine the results on Type 1 sentence determiners and adjectives.

Level	Type 1 DET Score	(%)	Maximum	(%)
A2	167	81,86	204	100
B2	183	89,71	204	100
C1	186	91,18	204	100
Total	536	87,58	612	100
Level	Type 1 Average DET Score	(%)	Maximum	(%)
A2	9,82	81,86	12	100
B2	10,76	89,71	12	100
C1	10,94	91,18	12	100
Total	31,53	87,58	36	100

Table 16: L2 Portuguese participants: scores on Type 1 sentence determiners

Level	Type 1 ADJ Score	(%)	Maximum	(%)
A2	165	80,88	204	100
B2	182	89,22	204	100
C1	186	91,18	204	100
Total	533	87,09	612	100
Level	Type 1 Average ADJ Score	(%)	Maximum	(%)
A2	9,71	80,88	12	100
B2	10,71	89,22	12	100
C1	10,94	91,18	12	100
Total	31,35	87,09	36	100

Table 17: L2 Portuguese participants: scores on Type 1 sentence adjectives

The tendencies of growing numbers of correct answers from the lower to the higher levels of proficiency are visible, which were expected from the combined data that we saw above (Table 15): the increase in success rates from level A2 to B2 are much higher than from B2 to C1. If we compare the performance on determiners to adjectives of the different levels of proficiency, what can be stated is that the participants on level A2 assigned 0,98% more correct gender values to determiners than to adjectives; on level B2 0,49% more determiners were correctly chosen than adjectives; and on level C1 there is no difference between the attained scores. These last figures may illustrate a tendency that gender assignment of determiners is more easily acquired than in adjectives. These minimal differences between the obtained values on determiners and adjectives might be due to the fact that both items are in a local agreement relation with the noun.

Again, if we compare these results with those of the control group (99,80% for both Type 1 determiners and adjectives), we can conclude that the performance of the L2 Portuguese participants is not native-like, but highly accurate on level C1.

We move on to Type 2 sentences, where the space between the noun and the adjective is increased by an intervening copula. Table 18 demonstrates the data on determiners and adjectives combined:

Level	Type 2 Overall Score	(%)	Maximum	(%)
A2	306	75,00	408	100
B2	367	89,95	408	100
C1	347	85,05	408	100
Total	1020	83,33	1224	100
Level	**Type 2 Average Overall Score**	**(%)**	**Maximum**	**(%)**
A2	18,00	75,00	24	100
B2	21,59	89,95	24	100
C1	20,41	85,05	24	100
Total	60,00	83,33	72	100

Table 18: L2 Portuguese participants: overall scores on Type 2 sentences

If we examine the results on Type 2 sentences and compare them with the corresponding results of the control group (99,71%), we can conclude, yet again, that the performance of the L2 Portuguese participants is not native-like and is lower than that with Type 1 sentences.

On level A2, our participants obtained 14,95% worse results than those on level B2, however those on level C1 underperformed our level B2 participants by 4,9%, which, as we have mentioned before, might be due to the fact that level C1 participants are already being trained in different areas of grammar, therefore they might experience regression with regard to the specific issue of gender agreement.

If we take a look at Type 2 sentences with respect to results on determiners[56] and adjectives separately, we can draw very similar conclusions:

[56] Even though Type 2 sentences specifically focus on the adjectives in the predicative position, the results on the determiners might not be as important in this case, however, to continue with the logic and structure that we apply throughout the analysis of our data we included the separate results on Type 2 determiners as well.

Level	Type 2 DET Score	(%)	Maximum	(%)
A2	154	75,49	204	100
B2	184	90,20	204	100
C1	174	85,29	204	100
Total	512	83,66	612	100
Level	**Type 2 Average DET Score**	**(%)**	**Maximum**	**(%)**
A2	9,06	75,49	12	100
B2	10,82	90,20	12	100
C1	10,24	85,29	12	100
Total	30,12	83,66	36	100

Table 19: L2 Portuguese participants: scores on Type 2 sentence determiners

Level	Type 2 ADJ Score	(%)	Maximum	(%)
A2	152	74,51	204	100
B2	183	89,71	204	100
C1	173	84,80	204	100
Total	508	83,01	612	100
Level	**Type 2 Average ADJ Score**	**(%)**	**Maximum**	**(%)**
A2	8,94	74,51	12	100
B2	10,76	89,71	12	100
C1	10,18	84,80	12	100
Total	29,88	83,01	36	100

Table 20: L2 Portuguese participants: scores on Type 2 sentence adjectives

With respect to determiners, level B2 acquired the highest scores. Level C1 obtained 4,91% less correct answers than B2, while A2 scored the lowest, exactly 14,71% lower than level B2. The results on adjectives for Type 2 sentences show the same rather sharp increase then a small decline in the success rates: A2 scored 15,2% lower than B2 and C1 scored 4,91% lower than B2, which might be due to a training effect as we have mentioned before.

If we compare the results on determiners with those on adjectives, we see that all levels performed slightly better with determiners than with adjectives, just as we have seen for Type 1 sentences and might be either because of the determiner having a scope over the noun, and

therefore the gender assignment of determiners being more easily acquired, or simply because in the phrasal structure the determiner is hierarchically above the noun.

However, the results of our L2 Portuguese participants do not reach the native-like levels established by our control group (99,80% for Type 2 determiners and 99,61% for Type 2 adjectives) and do not even reach 90%. This means that our participants performed better, in general, with Type 1 sentences.

Type 3 sentences, as we have seen, consist of a linear order formed by noun, relative clause and adjective, i. e. the relative clause is (apparently) separating the noun from its modifying adjective and, moreover, the subject of the relative clause is another noun (with the opposite gender marking to the head noun) intervening between the target noun and its modifying adjective. Let us examine the collected data on this sentence type, starting with the combined results on determiners and adjectives:

Level	Type 3 Overall Score	(%)	Maximum	(%)
A2	331	81,13	408	100
B2	386	94,61	408	100
C1	389	95,34	408	100
Total	1106	90,36	1224	100
Level	**Type 3 Average Overall Score**	**(%)**	**Maximum**	**(%)**
A2	19,47	81,13	24	100
B2	22,71	94,61	24	100
C1	22,88	95,34	24	100
Total	65,06	90,36	72	100

Table 21: L2 Portuguese participants: overall scores on Type 3 sentences

As we can see in Table 21, the results are the highest with Type 3 sentences. The native-like level would start at 99,41% which was not attained, although level B2 and C1 had scores close to 95% and higher, which indicates a near native-like[57] knowledge of these levels of proficiency with respect this sentence structure.

[57] We define near native-like performance as maximum 5% lower than the corresponding results of the control group.

These findings contradict Hypothesis 2 (see Chapter 3), according to which we expected this type of sentence to have the lowest performance, however these results might be due to an error in the construction of the experimental test. When we provided the Hungarian translations of the Portuguese target sentences in the cloze test, we did not account for the fact that the case marking and the animacy features of the Hungarian relative pronouns could provide important cues based on which our participants could more easily have related the head noun with its modifying adjective and perhaps disregarded the [+human] intervening noun (the subject of the relative clause) (see more on this issue in section 5.2.).

The participants who had the highest scores with Type 3 sentences were those on level C1, thus establishing a tendency of increase in success rates from level A2 through B2 and to C1. The difference between level A2 and level B2 scores is, again, much higher (13,48%) than the difference between B2 and C1 level scores (0,73%), as expected, since in formal language learning contexts, level B is where speakers (of L2 Portuguese in our case) train some already known competences and add a layer of complexity to the structures they train, consolidating some issues (such as gender agreement).

The same pattern shows if we break down the combined results to only determiners and adjectives, the only difference being that B2 participants minimally (by 0,49%) outperformed those on level C1 with regard to determiner score.

Level	Type 3 DET Score	(%)	Maximum	(%)
A2	168	82,35	204	100
B2	196	96,08	204	100
C1	195	95,59	204	100
Total	559	91,34	612	100
Level	**Type 3 Average DET Score**	**(%)**	**Maximum**	**(%)**
A2	9,88	82,35	12	100
B2	11,53	96,08	12	100
C1	11,47	95,59	12	100
Total	32,88	91,34	36	100

Table 22: L2 Portuguese participants: scores on Type 3 sentence determiners

Level	Type 3 ADJ Score	(%)	Maximum	(%)
A2	163	79,90	204	100
B2	190	93,14	204	100
C1	194	95,10	204	100
Total	547	89,38	612	100
Level	**Type 3 Average ADJ Score**	**(%)**	**Maximum**	**(%)**
A2	9,59	79,90	12	100
B2	11,18	93,14	12	100
C1	11,41	95,10	12	100
Total	32,18	89,38	36	100

Table 23: L2 Portuguese participants: scores on Type 3 sentence adjectives

As for the results on adjectives, the tendency matches that of the combined data on Type 3 sentences, meaning an increase from A2 to B2 and from B2 to C1, however the increase reduces between the two latter levels, as we have seen for Type 2 sentences.

The comparison between the performances on determiners and adjectives leads us to a similar conclusion that we already drew for Type 1 and Type 2 sentences: our participants assigned slightly more correct gender values to determiners than to adjectives, probably because of the more intimate relation between determiner and noun in Portuguese and because of the 'noun-adjective' order is different in Hungarian (see Chapter 3).

Once again, the exact native-like levels established by our control group (100% for Type 3 determiners and 98,82% for Type 3 adjectives) were not reached, however our participants did produce very approximate results on level B2 and C1 with regard to determiners and on level C1 with regard to adjectives. Also, as we have seen, the control group scored the lowest results with Type 3 adjectives (98,82%, see Table 8).

When we observe the progress of our participants from Type 1 through Type 2 to Type 3 sentences, we can see a bit of a fluctuation, as is presented in the synthesis below:

Level	Type 1 Overall Score (%)	Type 2 Overall Score (%)	Type 3 Overall Score (%)
A2	81,37	75,00	81,13
B2	89,46	89,95	94,61
C1	91,18	85,05	95,34
Total	87,34	83,33	90,36

Table 24: Comparison of overall scores between sentence types

For level A2, predicative constructions seem to be the most problematic, since their results for Type 2 sentences were the lowest (75%). This level of proficiency mastered Type 1 sentences (attributive adjectives) the best with a success rate of 81,37%, however Type 3 sentences appear to be only the second most problematic with 81,13% of correct answers, which does not confirm Hypothesis 2 (see Chapter 3), since this type was expected to be the hardest to master because of the (superficial) long distance between the noun and the adjective.

Level B2 shows a different pattern, which still contradicts Hypothesis 2, according to which Type 1 sentences would receive the most correct answers and Type 3 the fewest. Our level B2 participants seem to have the most problems with Type 1 sentences with an 89,46% success rate, having scored slightly higher with Type 2 sentences (89,95%) and having mastered Type 3 sentences at 94,61%.

Turning to level C1: like level A2, level C1 also had their lowest scores with Type 2 sentences with an 85,05% success rate. Their second best performance can be assigned to Type 1 sentences (91,18%), which were expected to be the easiest constructions. Lastly, Type 3 sentences earned the best scores by level C1: 95,34%, which can be considered near native-like.

Level C1 participants were expected to outperform all other levels of proficiency, however, this apparently was not the case. Level B2 did reach very similar results to that of level C1 and it also performed better on Type 2 sentences by 4,9%. Level A2, in general terms, attained the expected lower results than level B2, because the knowledge of level A2 participants might not yet fully encompass Type 3 sentence structures. Nevertheless, all results are highly accurate, which means that this particular area might not be very difficult to dominate in the L2.

Therefore, in total, Type 3 sentences received the most correct gender values (90,36%), which is the exact opposite of what we had hypothesized based on the results of the control group, where it was the least successful sentence type (see Table 8). These findings support

Kayne's (1994) 'raising' analysis of this type of sentences: he claims that, in the beginning of the derivation process, the head noun and the adjective are adjacent, i.e. are actually in local agreement and not in a long distance one.

Type 1 sentences were supposed to receive the most correct answers according to our hypothesis, however, they only obtained a 87,34% success rate, which was the second lowest. The lowest scores were earned by our Type 2 sentences (with 83,33%), which were expected to be only the second most difficult. These results might indicate that this is the most difficult construction to acquire for Hungarian native speakers, perhaps because Hungarian copulative structures are different from Portuguese ones (see Chapter 3).

Hungarian native speakers therefore might find Type 2 structures the most difficult to acquire and Type 3 sentences the least difficult, which might mean that the distance between the adjective and the noun plays a lesser role in the acquisition of gender agreement and the assignment of gender values than we previously hypothesized, however we cannot be sure because of the previously mentioned error in the construction of the experimental test with respect to the possible facilitative effect of the Hungarian translation of Type 3 sentences (see more in section 5.2.). What seems to have more significance with regard to this issue is the structure itself, in this case the copulative structure. Also, as Kayne (1994) claims, Type 3 sentences may not actually present long distance agreement (because the head noun and its modifying adjective are only apparently distance), but a relation of local agreement.

4.2.3. Overall data on noun groups

As we established in Chapter 3, we set up 12 different groups of nouns based on 10 non-default word markers and the 2 default uniform [+masculine] and [+feminine] theme indices found in Portuguese.

We shall now examine the overall performance of our participants on each level of proficiency with respect to the established noun groups. Table 25 below is a summary of the overall results (scores on determiners and adjectives combined) of our participants with respect to noun groups and levels of proficiency.

Overall Score (%)												
Level	Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
A2	77,08	63,89	100	65,28	86,11	70,83	82,64	66,67	75,00	90,28	100	97,22
B2	90,28	98,61	98,61	95,14	95,83	95,83	81,94	91,67	81,94	83,33	95,83	98,61
C1	86,11	94,44	100	93,06	100	95,83	68,06	97,22	84,03	93,06	100	100
Total	84,49	85,65	99,54	84,49	93,98	87,50	77,55	85,19	80,32	88,89	98,61	98,61

>90-100%
>80-90%
>70-80%
60-70%

Table 25: L2 Portuguese participants: overall scores broken down to noun groups

To be able to most effectively represent the fluctuations in performance between the groups of nouns and the levels of proficiency, we used a color-coding system. We established a four-degree scale from 60% to 100% (63,89% being the lowest performance and 100% being the highest in our data), assigning different colors to each degree. Since we did not obtain results lower than 63,89%, we can conclude that no scores were around chance level.

Thus, as it is visible, the lowest scores (60-70%, light-grey color) were encountered mostly on level A2 as we had expected, because this level still has the lowest proficiency. On this level, the nouns in Group 2 containing the word marker <-*gem*> seemed to be the most problematic as for gender assignment (63,89% success rate), followed by Group 4 (word marker <-*r*>) and Group 8 ([+masculine] word marker <-*a*>), where the most troublesome noun was *mapa* 'map' with only a 25% success rate (see Appendix 7).

However, level C1 also made a lot of errors with respect to Group 7 (nouns ending in stressed vowels) (68,06% correct answers), especially with two nouns: *chaminé* 'chimney' (16,66%) and *pá* 'dustpan' (25%, see Appendix 7).[58] The results with these two particular target-nouns turn even more interesting if we compare the results of the other levels of proficiency to those of C1: level A2 scored 91,66% and 75% with these two nouns, respectively and level B2 scored 58,33% and 41,66%, respectively (see Appendix 7). This means that level A2 highly outperformed both level C1 and B2 with respect to the aforementioned nouns from Group 7. This might be due to the fact that these vocabulary items are present in level A2 training, but might disappear later on or appear less frequently.

The results from the second lowest degree (>70-80% success rate, blue color) were all obtained by level A2. The affected groups on this level were the following: Group 6 (word marker <-*s*/-*z*>, 70,83%) the lowest performing noun being *nariz* 'nose' (45,83%, see Appendix 7); Group 9 (word marker <-*ão*>, 75%), the lowest performing noun being *coração* 'heart' (58,33%, see Appendix 7); and Group 1 (word marker <–*e*>, 77,08%), the lowest performing noun being *árvore* 'tree' (54,16%, see Appendix 7). *Nariz* 'nose' performing so low was quite surprising, since the parts of the body are usually part of the basic vocabulary that L2 learners acquire already on level A1. However, very many participants had previously learned Spanish, where the same noun (written exactly the same way) has the opposite gender value, i.e. it is feminine, as opposed to Portuguese, where it is masculine. This could easily be the effect of negative transfer from the L2 (Spanish) to the L3 (Portuguese). Additionally, the total

[58] These two words might belong to a section of vocabulary that was taught on lower levels on proficiency, but since they might not have been very frequently used afterwards, they were gradually forgotten. It could be because of this training effect that level A2 performed so much better with these nouns than level B2 or C1. Also, as we have seen before, *chaminé* is a very infrequent noun in contemporary Portuguese, and so is *pá* 'dustpan': a research in the CRPC database (http://alfclul.clul.ul.pt/CQPweb/portugalonlyv23/) for *pá* returned 1,062 matches in 650 different texts (in 289,840,619 words [332,332 texts]) indicating a frequency of **3.66** instances per million words. This research, however, included *pá* as an interjection as well, therefore we ran another research for *a pá* to filter the results, which were the following: 58 matches in 50 different texts (in 289,840,619 words [332,332 texts]) indicating a frequency of **0.2** instances per million words. Afterwards, we did a final research for a different noun with the same word marker, *chá* 'tea', and obtained the following results: 3,535 matches in 1,341 different texts (in 289,840,619 words [332,332 texts]) indicating a frequency of **12.2** instances per million words.
Thus, as we can see, *pá* as a noun has a very low frequency compared to *chá*, which has the same word marker. This could explain the low rates of success with this particular noun.

performance of level C1 also falls into this degree of mastery because of the participants' very poor performance with Group 7 nouns, as we discussed above.

All levels of proficiency managed to attain results that classify as our third degree (>80-90%, purple color). Most of the total scores fall into this category, more precisely Groups 1, 2, 4, 6, 8, 9 and 10.

Nevertheless, it is more remarkable that most of our results range between >90-100%, which is our highest degree (red wine color). Even level A2 had two perfect scores (100% with Group 3 and Group 11)[59]. Level B2 did not reach any 100% scores, however most of their results are within the highest range. Level C1 scored 100% four times: with Groups 3, 5, 11 and 12, as one might expect, since at this stage students should perform more like native speakers, attaining scores similar to those of the control group (see Appendix 7).

The highest scores, in total, were reached with Group 3 (word marker <-*dade*>, 99,54%), Group 11 and 12 (the default [+feminine] and [+masculine] theme indices, respectively, 98,61% both) and Group 5 (word marker <-*l*>, 93,98%).[60]

Now, if we compare the results of our L2 Portuguese participants with the corresponding results of the control group, we can see some interesting details. Group 3, for instance, obtained even higher results than the native-like level in total (the control group scored 98,89% and the target-participants 99,54% in total, and 100% on level A2 and C1). With Group 5, level C1 performed on the native-like level (100%), and so did level A2 and C1 with Group 11 and level C1 with Group 12 (both reaching 100%, just like the control group), the latter two being the default theme indices in Portuguese, therefore the most frequent and most transparent, which might have facilitated their acquisition.

The following tables contain the data on the analysis of the performances on our noun groups broken down to the scores on determiners and adjectives. We also employed a color-coding system here, where, comparing the data on determiners and adjectives, the green colored

[59] The fact that our participants obtained perfect scores with these noun groups on level A2 may be due to <-*dade*> (Group 3) being disyllabic and having a similar form and the same gender value ([+feminine]) across Romance languages (e.g. *liberdade* in Spanish, *liberté* in French, *libertà* in Italian and *llibertat* in Catalan); and Group 11 comprising the default ([+feminine]) theme index <-*a*>.

[60] The gender assignment of nouns that end with <-*l*> in Romance languages is uniform, these words having the [+masculine] gender value, e.g. 'paper': *papel* in Spanish and Portuguese, *paper* in Catalan, and *papier* in French, therefore transfer from these languages could have facilitated the acquisition of the gender value of such nouns, since many of our participants speak at least one of these languages in addition to Portuguese.

cells show the higher results, the yellow colored cells show equal results and the red colored cells show lower scores.

DET Score (%)												
Level	Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
A2	76,39	63,89	100	65,28	86,11	73,61	83,33	66,67	76,39	94,44	100	97,22
B2	90,28	100	100	95,83	97,22	97,22	81,94	94,44	81,94	83,33	97,22	97,22
C1	86,11	94,44	100	93,06	100	95,83	68,06	97,22	84,72	94,44	100	100
Total	84,26	86,11	100	84,72	94,44	88,89	77,78	86,11	81,02	90,74	99,07	98,15

Higher than ADJ Score
Equal to ADJ Score
Lower than ADJ Score

Table 26: L2 Portuguese participants: scores on determiners broken down to noun groups

ADJ Score (%)												
Level	Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
A2	77,78	63,89	100	65,28	86,11	68,06	81,94	66,67	73,61	86,11	100	97,22
B2	90,28	97,22	97,22	94,44	94,44	94,44	81,94	88,89	81,94	83,33	94,44	100
C1	86,11	94,44	100	93,06	100	95,83	68,06	97,22	83,33	91,67	100	100
Total	84,72	85,19	99,07	84,26	93,52	86,11	77,31	84,26	79,63	87,04	98,15	99,07

Higher than DET Score
Equal to DET Score
Lower than DET Score

Table 27: L2 Portuguese participants: scores on adjectives broken down to noun groups

If we examine the yellow colored cells first, it is visible that most of the results on determiners and adjectives show absolutely no difference. However, if we compare all red and green cells, we can conclude, much like previously with other determiner-adjective comparisons, that our participants attained better results with determiners than with adjectives, although the differences are minimal. This, as we have mentioned before, can be due to the fact that in Portuguese the determiner has a scope over the noun, thus facilitating its acquisition. Furthermore, when L2 students acquire vocabulary in the target language, they tend to be trained to memorize lists of nouns that appear with their respective determiners, thus not only learning the noun but the determiner and the noun at the same time. The errors they do make could also be due to the effect of negative transfer from other Romance L2(s), where the same nouns may have the opposite gender value, as we have seen with *nariz* ‘nose’.

4.3. Data analysis of the Romance L2 vs. Non-Romance L2 Subgroups

In order to investigate Hypothesis 5, in the following section we are going to divide our 36 Portuguese L2 speaker participants into two subgroups; one of the subgroups shall constitute those individuals who - Portuguese excluded - had acquired at least one other Romance language (which means that Portuguese is, at least, an L3 for these speakers); and to the other shall be assigned all the individuals who had not acquired any other Romance language than Portuguese.[61] The denomination used for the former subgroup will be 'Romance L2 Subgroup' and the latter subgroup shall be called 'Non-Romance L2 Subgroup'. Therefore, we are going to attempt to examine whether the similar gender morphology of another acquired Romance language (or languages) has any effect on the performance of our participants, thus proving or disproving Hypothesis 5.

Firstly, let us have a look at the data on the Romance L2 Subgroup. In total, 26 individuals of our 36 Hungarian participants qualify to be placed in this category; 8 subjects are from level A2, 6 are from level B2, and 12 are from C1. The following table contains the summary of how many participants speak which other Romance languages:

Language	Pax
Spanish	19
Italian	11
Catalan	6
French	4
Galician	2
Romanian	1
Total	**43**

Table 28: other Romance languages spoken by our participants (i.e. Portuguese Ln, n>2)

As Table 28 exhibits, most of our participants speak Spanish and/or Italian as an L2, which are languages that have the most similar gender morphology to Portuguese. We can also deduce from the data in Table 28 that many participants speak more than one other Romance

[61] The division of the overall population into these subgroups was based on the responses to the linguistic questionnaire they filled out (see Appendix 2).

language, since the total number of other spoken Romance languages (43) is more than the number of participants in this subgroup (26).

Secondly, we shall examine our Non-Romance L2 Subgroup. In total, 10 individuals of the 36 Hungarian participants were placed in this subgroup; 4 participants are on level A2 and 6 of them are on level B2. Since all of our 12 participants on level C1 speak another Romance language, the Non-Romance L2 subgroup lacks subjects from this level of proficiency.[62]

4.3.1 Overall results of the Romance L2 vs. Non-Romance L2 Subgroups

In this section we shall examine the overall results of the two subgroups. The table below shows the results on determiner and adjective success rates combined, for each subgroup in question, side by side.

Level	Overall Romance L2 Score	(%)	Maximum	(%)	Level	Overall Non-Romance L2 Score	(%)	Maximum	(%)
A2	679	83,21	816	100	**A2**	290	71,08	408	100
B2	544	88,89	612	100	**B2**	574	93,79	612	100
C1	1108	90,52	1224	100	**C1**	-	-	-	-
Total	2331	87,90	2652	100	**Total**	783	76,76	1020	100

Table 29: Romance L2 and Non-Romance L2 Subgroups: overall scores

If we compare the results of these subgroups to those of the control group (which was 99,64%, see Table 5), we can conclude that the performance of either subgroup is not native-like, however, as level B2 of the Non-Romance L2 Subgroup scored over 93%, being the most accurate of all levels of proficiency of both subgroups.

The biggest difference in performance that we can see between the Romance L2 and Non-Romance L2 Subgroups is on the lowest level of proficiency: those level A2 participants who belong to the Romance L2 Subgroup scored 12,13% more correct answers than the Non-

[62] When we applied our cloze test to the L2 Portuguese participants we did not consider the possibility of not having any participants on any level of proficiency, but after having revised the literature together with the obtained results the problem was noticed. Unfortunately, we did not have the means to collect more data mainly due to the very limited number of L2 Portuguese students in Hungary.

Romance L2 Subgroup. This could mean that in the initial stages of L2 acquisition (like on level A2) there is positive transfer from the L2 (a Romance language) to the L3 (Portuguese), but in later stages when the knowledge of the L2 is strengthened, there seems to be neutral or negative transfer (compare the results of level A2 with level B2, where the Romance L2 Subgroup scored at 88,89% and the Non-Romance L2 Subgroup at 93,79%) (see White 2003). Therefore, at least on level A2, a previously acquired Romance language does seem to facilitate the acquisition of gender agreement.

On level B2, as we have mentioned, the opposite can be observed: those who do not speak another Romance language obtained better results, however the difference between the performances decreases to only 4,9%. This might be due to the negative transfer effect that the Romance L2 speakers experience on this level but the Non-Romance L2 speakers do not, therefore the latter perform better. Also, among the speakers of the Non-Romance L2 Subgroup there is a rather significant leap from level A2 to level B2: level B2 performed 22,71% better than level A2. This might mean that with more training and without the effect of negative transfer from another Romance L2, gender agreement could be more easily acquired for the Non-Romance L2 Subgroup after having passed the initial stages.

Interestingly, C1 participants, all of whom belong to the Romance L2 Subgroup, did not perform better than level B2 Non-Romance L2 participants, although the scores are quite similar (3,27% difference in favor of the Non-Romance L2 Subgroup). This could be due to the negative transfer from other Romance languages where the target-nouns might have a similar form but the opposite gender value (as we have seen, *nariz* ‘nose’ has the same form in Spanish and Portuguese but is [+feminine] in Spanish and [+masculine] in Portuguese).

In total, however, the Romance L2 Subgroup did outperform the Non-Romance L2 subgroup by 11,14%, which might indicate that in spite of the negative transfer from the Romance L2 the typological and lexical proximity between the L2 and the L3 (Portuguese) does facilitate gender acquisition.

Now we shall separate the obtained values on determiners from those on adjectives to find out which subgroup scored better and with which item.

Level	Romance L2 DET Score	(%)	Maximum	(%)	Level	Non-Romance L2 DET Score	(%)	Maximum	(%)
A2	342	83,82	408	100	**A2**	147	72,06	204	100
B2	273	89,22	306	100	**B2**	290	94,77	306	100
C1	555	90,69	612	100	**C1**	-	-	-	-
Total	1170	88,24	1326	100	**Total**	437	85,69	510	100

Table 30: Romance L2 and Non-Romance L2 Subgroups: scores on determiners

Level	Romance L2 ADJ Score	(%)	Maximum	(%)	Level	Non-Romance L2 ADJ Score	(%)	Maximum	(%)
A2	337	82,60	408	100	**A2**	143	70,10	204	100
B2	271	88,56	306	100	**B2**	284	92,81	306	100
C1	553	90,36	612	100	**C1**	-	-	-	-
Total	1161	87,56	1326	100	**Total**	427	83,73	510	100

Table 31: Romance L2 and Non-Romance L2 Subgroups: scores on adjectives

When we compare the results of the two subgroups with regard to determiners and adjectives, the general tendency prevails: our participants in each subgroup performed slightly better with determiners than with adjectives. The biggest difference in performances, however, occurs on level B2, where the Non-Romance L2 Subgroup outperformed the Romance L2 Subgroup by 5,55% with respect to determiners and by 4,25% with respect to adjectives, which is a tendency that we have observed before in the overall results (see Table 29 above). Also, as we have seen with the overall results, Table 30 and 31 also show a notable leap from level A2 to B2 with the Non-Romance L2 Subgroup, which, again, might be due to training effects and to the fact that there is no negative transfer from any other previously acquired Romance language.

However, as we have seen for the overall results, the performance of either subgroup does not seem native-like, which would mean at least a 99,87% success rate for determiners and a 99,41% success rate for adjectives (see Table 6 and 7). It was only level B2 of the Non-Romance L2 Subgroup that scored the closest to these aforementioned control group results, this way we

can dub their performance highly accurate with this issue (having scored 94,77% with determiners and 92,81% with adjectives).

4.3.2. Data analysis of the Romance L2 vs. Non-Romance L2 Subgroups by types of sentences

In this section we shall examine the different performances of the two subgroups in question with respect to the three types of sentences of our study. The table below shows the comparison of overall scores (determiners and adjectives combined).

Overall Score (%)						
	Type 1		Type 2		Type 3	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	80,51	80,15	79,78	65,44	87,87	67,65
B2	89,22	91,67	86,76	93,14	92,65	96,57
C1	91,18	-	85,05	-	95,34	-
Total	87,44	87,06	83,82	82,06	92,42	85,00

Table 32: Romance L2 and Non-Romance L2 Subgroups: overall scores broken down to sentence types

When comparing these results of the subgroups in question with the corresponding results of the control group (which were 99,80% for Type 1 sentences, 99,71% for Type 2 sentences and 99,41% for Type 3 sentences, see Table 8), we can see that the only results that even approximate the native-like level (thus being near native-like) are from level B2 of the Non-Romance L2 Subgroup for Type 3 sentences (96,57%) and from level C1 of the Romance L2 Subgroup, also for Type 3 sentences (95,34%).

Table 32 shows very little difference between the scores for Type 1 sentences of each subgroup on all levels of proficiency; the total difference being 0,38% in favor of the Romance L2 Subgroup. However, level B2 participants of the Non-Romance L2 Subgroup did perform minimally better than those subjects on level B2 and C1 from the Romance L2 Subgroup in all sentence types, as we have already seen with the overall results for Romance L2 vs. Non-

Romance L2 speakers. This, once again, might be due to lack of negative transfer from another Romance language.

The data of these subgroups on Type 2 sentences shows more divergence than that on Type 1 sentences. The biggest difference in scores occurs on level A2 between the two subgroups: the Romance L2 subgroup attained 14,34% better results. This might be due to the initial benefit of positive transfer from other Romance L2s, which works in favor of the Romance L2 Subgroup, as we have already seen with the overall results for the two subgroups (see Table 29). On level B2, the performances shift in favor of the Non-Romance L2 Subgroup, who obtained 6,38% better scores than the Romance L2 Subgroup on the same level of proficiency, but also having outperformed the latter subgroup on level C1 by even more, 8,09%. This might happen because of negative transfer from other Romance languages. The total score between the two subgroups for Type 2 sentences is quite similar (1,8% difference), however the Romance L2 Subgroup did perform minimally better, which might indicate that for the acquisition of gender agreement in this type of sentence it is beneficiary to speak another Romance language.

As for Type 3 sentences, the biggest difference, once again, shows between the two subgroups on level A2: 20,22% in favor of the Romance L2 Subgroup, which could be explained by the initial advantage of positive transfer from other Romance languages that the Non-Romance L2 Subgroup cannot rely on. On level B2, the Non-Romance L2 Subgroup obtained slightly better scores (by 3,92%), similarly to Type 1 and Type 2 sentences; and also outscored the performance of level C1 Romance L2 Subgroup participants by 1,23%. This might indicate, once again, that training in one specific Romance language might make the acquisition of gender agreement easier than having received training in various Romance languages where, crosslinguistically, the gender values of same noun may vary.

The difference between the total scores of the two subgroups seems to be the most notable with Type 3 sentences: the Romance L2 Subgroup scored 7,42% more correct answers than the Non-Romance L2 Subgroup. This also means that the Romance L2 Subgroup was the most successful with Type 3 sentences, followed by Type 1 sentences and, lastly, by Type 2 sentences, always outscoring, even if minimally, the Non-Romance L2 Subgroup with respect to the total score. The Non-Romance L2 Subgroup obtained the highest scores with Type 1 sentences, followed by Type 3 sentences. The latter subgroup also scored the lowest results with Type 2 sentences, similarly to the Romance L2 Subgroup. The fact that Type 3 sentences proved to be

the least difficult to acquire for either subgroup might mean that the superficial long-distance agreement between the noun and the adjective could actually be a local agreement, which favors Kayne's (1994) raising hypothesis. Also, since Type 2 sentences obtained the lowest scores with both subgroups, where there is only a copula between the noun and the adjective, this might point to a difficulty in the acquisition of the predicative structure itself, i.e. the distance between the noun and the adjective might not be as important of a factor in the acquisition of gender agreement as is the phrase structure.

The following tables detail the results of the two subgroups in question with regard to determiners and adjectives attained with the three types of sentences.

DET Score (%)						
	Type 1		Type 2		Type 3	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	82,35	80,88	80,88	64,71	89,71	70,59
B2	87,25	92,16	87,25	93,14	94,12	99,02
C1	91,18	-	85,29	-	95,59	-
Total	87,56	87,65	84,39	81,76	93,44	87,65

Table 33: Romance L2 and Non-Romance L2 Subgroups: scores on determiners broken down to sentence types

ADJ Score (%)						
	Type 1		Type 2		Type 3	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	81,62	79,41	78,68	66,18	87,50	64,71
B2	87,25	91,18	86,27	93,14	92,16	94,12
C1	91,18	-	84,80	-	95,10	-
Total	87,33	86,47	83,26	82,35	92,08	82,35

Table 34: Romance L2 and Non-Romance L2 Subgroups: scores on adjectives broken down to sentence types

As we can see, the tendencies described in Table 32 for the overall scores on the sentence types of the subgroups in question remain unchanged. As we have concluded previously in connection with our overall population of participants and with regard to the overall scores of the Romance L2 and Non-Romance L2 Subgroups as well, our participants obtained minimally higher scores with determiners than with adjectives – this conclusion also remains with respect to the results demonstrated above in Tables 33 and 34. This, again, might be due to the training effect where the determiner is taught together with the noun and to the distance between the determiner and the nouns being constant, i.e. the determiners always immediately preceding the noun, thus making the gender acquisition between determiner and noun less difficult.

If we compare these results to the corresponding results from the control group (which were 99,80% for Type 1 determiners and 99,80% for Type 1 adjectives; 99,80% for Type 2 determiners and 99,61% for Type 2 adjectives; and 100% for Type 2 determiners and 98,82 for Type 3 adjectives, see Table 8), we can see that only Type 3 sentences received near native-like results. As for determiners, the Romance L2 Subgroup scored 95,59% on level C1 and the Non-Romance L2 Subgroup scored 99,02% on level B2 with Type 3 sentences, which were the highest results of both subgroups, respectively. As for adjectives, Type 3 sentences received near native-like, i.e. the highest scores, too; which were 95,10% for level C1 from the Romance L2 Subgroup and 94,12% on level B2 from the Non-Romance L2 Subgroup. These very high results with Type 3 sentences might also corroborate Kayne's (1994) raising hypothesis, since, if the acquisition of gender agreement is the most successful with this sentence type, it can only be a long-distance agreement on the surface.

4.3.3. Data analysis of the Romance L2 vs. Non-Romance L2 Subgroups by noun groups

To be able to represent and compare the results of the two subgroups in question with regard to the noun groups of our research, we opted for creating a color-coded table. We established a six-degree scale ranging from 40-100% success rates (the lowest result being 41,67% and the highest 100%), assigning a different color to each degree. Let us now examine the performances on each degree of our scale, presented in Table 35 below:

Overall Score (%) (Part 1: Noun Groups 1-6)												
	Group 1 <-e>		Group 2 <-gem>		Group 3 <-dade>		Group 4 <-r>		Group 5 <-l>		Group 6 <-s/z>	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	82,29	66,67	62,50	66,67	100	100	75,00	45,83	87,50	83,33	77,08	50,00
B2	83,33	97,22	100	97,22	100	97,22	95,83	94,44	94,44	97,22	98,61	98,61
C1	86,11	-	94,44	-	100	-	93,06	-	100	-	95,83	-
Total	84,29	85,00	85,90	85,00	100	98,33	88,14	75,00	94,87	91,67	90,71	79,17

Overall Score (%) (Part 2: Noun Groups 7-12)												
	Group 7 <-é/á/ó>		Group 8 <-a> [+masc]		Group 9 <-ão>		Group 10 <-im>		Group 11 <-a> [+fem]		Group 12 <-o>	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	79,17	89,58	79,17	41,67	78,13	68,75	93,75	83,33	100	100	100	91,67
B2	86,11	77,78	88,89	94,44	75,00	88,89	72,22	94,44	91,67	100	97,22	100
C1	68,06	-	97,22	-	84,03	-	93,06	-	100	-	100	-
Total	75,64	82,50	89,74	73,33	80,13	80,83	88,46	90,00	98,08	100	99,36	96,67

>90-100%	
>80-90%	
>70-80%	
>60-70%	
>50-60%	
40-50%	

Table 35: Romance L2 and Non-Romance L2 Subgroups: overall scores broken down to noun groups

The lowest results, ranging between 40-50%, assigned to the light grey color, were solely attained by the Non-Romance L2 Subgroup on level A2 with Groups 4, 6 and 8, all of which are groups of nouns with non-default word markers, respectively, <*-r*>, <*-s/z*> and <*-a[+masculine]*>.

These are results that indicate that the knowledge of the level A2 participants of the Non-Romance L2 Subgroup is at chance level when it comes to these word markers. Furthermore, these are the lowest results we have encountered so far when compared to the corresponding results of the whole population (see Table 25). If we take a look at the results of the Romance L2 Subgroup with regard to Group 4, 6 and 8, we can see how much of an advantage it means to speak at least another Romance language at this initial stage.

There were no scores ranging between >50-60%, therefore the light peach color assigned to this degree does not appear on Table 35. These would have been results still close to the chance level.

The turquoise color, for results ranging between >60-70%, appears mainly with scores obtained on level A2 by the Non-Romance L2 Subgroup (for Groups 1, 2, 9, all of them non-default word markers); and one time, on level A2, for the Romance L2 Subgroup as well, with Group 2 nouns (word marker <*-gem*>). This, again, indicates the beneficial effects of positive transfer from another Romance language in the initial stages of SLA. Interestingly and exactly as we have seen on Table 25, Romance L2 speakers on level C1 also scored within the turquoise color range with Group 7, i.e. the group of the word marker of stressed vowels. This could mean that some items, like *chaminé* 'chimney' (performing at 16,67%, see Appendix 13) and *pá* 'dustpan' (performing at 25%, see Appendix 13) might be very infrequent nouns, as we have seen, that are introduced at lower levels of proficiency but not trained later on, hence the lower results on a high level of proficiency, such as C1.

As for the results ranging between >70-80%, the blue color assigned to this degree indicates that Romance L2 speakers scored within this range with Groups 4, 6, 7, 8 and 9 on level A2, and with Groups 9 (<*-ão*>) and 10 (<*-im*>) on level B2 (all of them being non-default word markers), and that within this subgroup level C1 did not have any results within this range. The total performance of Romance L2 speakers is also within this range with Group 7 nouns because of the poor performance of level C1 speakers with the aforementioned lexical items in the turquoise-colored range. The Non-Romance L2 Subgroup obtained one result on level B2 within

this range, more precisely with Group 7, thus strengthening the claim that some vocabulary items that are introduced on the elementary level (level A1 or A2) is not trained anymore from the intermediate level (level B1 or B2) forward. The Non-Romance L2 Subgroup had no more such results on level A2 nor on C1, however, its total performance falls within this range with Groups 4, 6 and 8 (which were precisely the noun groups that included the lowest results).

The purple color assigned to the degree for results ranging between >80-90% appears on all levels of proficiency with both subgroups. The Romance L2 Subgroup had their total results within this range with Groups 1, 2, 4, 8, 9 and 10, and the Non-Romance L2 Subgroup with Groups 1, 2, 7, 9 and 10 (all of them being non-default word markers).

The highest degree, with results ranging between >90-100%, was assigned to the red wine color. Group 3 (word marker <-*dade*>), 11 (default theme index <-*a*>) and 12 (default theme index <-*o*>) only show this color, which means that all results, from both subgroups and from all levels of proficiency, were highly proficient and could even be considered near native-like or at least highly accurate (to compare with the corresponding results of the control group, see Table 9). These high results with these groups can be explained by Group 11 and 12 being the default theme indices in Portuguese, thus very transparent with regard to gender assignment and agreement; and also, as for Group 3, <-*dade*> is a disyllabic, thus a very salient word marker, which is always assigned to the [+feminine] value in other Romance languages (in its corresponding forms).

Other than these aforementioned groups, the Romance L2 Subgroup obtained results within this range on level A2 for Group 10 (word marker <-*im*>), on level B2 for Groups 2, 4, 5, 6, and on level C1 for 2, 4, 5, 6, 8, and 10. The Non-Romance L2 Subgroup, other than the aforementioned noun groups, attained scores within this range with Groups 2, 4, 5, 6, 8 and 10 on level B2, however on level A2 it was only with the highly successful groups of Group 3, 11 and 12.

There were even a few instances when the subgroups in question reached the native-like level or even outperformed it: with Group 2, level B2 of the Romance L2 Subgroup scored 100%, which is higher than the control group's score (99,44%); with Group 3 the Romance L2 Subgroup scored 100% on all levels of proficiency and so did the Non-Romance L2 Subgroup on level A2, outperforming the control group (98,89%); with Group 5, where the Romance L2 Subgroup performed at the native-like level on level C1, precisely at 100%; with Group 11,

where the Non-Romance L2 Subgroup scored 100% on all levels of proficiency and the Romance L2 Subgroup on level A2 and C1, thus performing as native speakers; and lastly with Group 12, where the Romance L2 Subgroup attained 100% on level A2 and C1 and so did the Non-Romance L2 Subgroup on level B2, thus performing on native-like level.

The following tables present the data on determiners and adjectives by each noun group obtained by the two subgroups in question. As we have attempted previously, our main goal by presenting these data in tables was to compare the performances on determiners and adjectives. We employed the color-coding system that was used for the data on noun groups beforehand: the results in the yellow cells are the same both with determiners and adjectives; the green cells mean higher scores and the red ones represent lower scores in comparison.

DET Score (%) (Part 1: Groups 1-6)												
	Group 1 <-e>		Group 2 <-gem>		Group 3 <-dade>		Group 4 <-r>		Group 5 <-l>		Group 6 <-s/z>	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	85,42	66,67	62,50	66,67	100	100	75,00	45,83	87,50	83,33	85,42	50,00
B2	86,11	97,22	100	100	100	100	97,22	94,44	94,44	100	94,44	100
C1	86,11	-	94,44	-	100	-	93,06	-	100	-	95,83	-
Total	85,90	85,00	85,90	86,67	100	100	88,46	75,00	94,87	93,33	92,31	80,00

DET Score (%) (Part 2: Groups 7-12)												
	Group 7 <-é/á/ó>		Group 8 <-a> [+masc]		Group 9 <-ão>		Group 10 <-im>		Group 11 <-a> [+fem]		Group 12 <-o>	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	79,17	91,67	79,17	41,67	79,17	70,83	95,83	91,67	100	100	100	91,67
B2	86,11	77,78	88,89	100	75,00	88,89	72,22	94,44	94,44	100	94,44	100
C1	68,06	-	97,22	-	84,72	-	94,44	-	100	-	100	-
Total	75,64	83,33	89,74	76,67	80,77	81,67	89,74	93,33	98,72	100	98,72	96,67

Higher than ADJ Score
Equal to ADJ Score
Lower than ADJ Score

Table 36: Romance L2 and Non-Romance L2 Subgroups broken down to noun groups: scores on determiners

ADJ Score (%) (Part 1: Group 1-6)												
	Group 1 <-e>		Group 2 <-gem>		Group 3 <-dade>		Group 4 <-r>		Group 5 <-l>		Group 6 <-s/z>	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	83,33	66,67	62,50	66,67	100	100	75,00	45,83	87,50	83,33	77,08	50,00
B2	83,33	97,22	100	94,44	100	94,44	94,44	94,44	94,44	94,44	91,67	97,22
C1	86,11	-	94,44	-	100	-	93,06	-	100	-	95,83	-
Total	84,62	85,00	85,90	83,33	100	96,67	87,82	75,00	94,87	90,00	89,10	78,33

ADJ Score (%) (Part 2: Groups 7-12)												
	Group 7 <-é/á/ó>		Group 8 <-a> [+masc]		Group 9 <-ão>		Group 10 <-im>		Group 11 <-a> [+fem]		Group 12 <-o>	
Level	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2	Romance L2	Non-Romance L2
A2	79,17	87,50	79,17	41,67	77,08	66,67	91,67	75,00	100	100	100	91,67
B2	86,11	77,78	88,89	88,89	75,00	88,89	72,22	94,44	88,89	100	100	100
C1	68,06	-	97,22	-	83,33	-	91,67	-	100	-	100	-
Total	75,64	81,67	89,74	70,00	79,49	80,00	87,18	86,67	97,44	100	100	96,67

Higher than DET Score
Equal to DET Score
Lower than DET Score

Table 37: Romance L2 and Non-Romance L2 Subgroups broken down to noun groups: scores on adjectives

As we can see when comparing Table 36 with Table 37, our two subgroups in question attained mainly equal scores on both determiners and adjectives, as is demonstrated by the predominance of the yellow-colored cells. There are also many instances of higher scores on determiners, however minimally. The only two instances where a higher score on adjectives can be found is with Group 12, with the Romance L2 Subgroup on level B2 and in the total results of the same subgroup with the same group of nouns. Thus we can conclude once again that determiners were, in general, assigned more correctly than adjectives, which might be due to a training effect (nouns being learned together with the determiner), to the effect of positive transfer from the Romance L2s, to the zero distance between determiner and noun, or to the fact that the determiner has a scope over the noun.

4.4. Data analysis of the Immersion vs. Non-Immersion Subgroups

In order to investigate Hypothesis 6, according to which those participants, who lived in an immersion context would perform better those who did not, we divided our 36 L2 Portuguese speakers into two different subgroups: the Immersion Subgroup includes those individuals who have spent at least 3 months living abroad in a Portuguese-speaking country, i.e. in an immersion context.[63] All of the Immersion Subgroup participants have lived in Portugal, therefore they were immersed in a context where European Portuguese was spoken and they were exposed to a distinct input, both in quantity and quality. Those participants who have never lived in an immersion context (or only for a very short time, less than 3 months) were assigned to the Non-Immersion Subgroup.

The Immersion Subgroup encompasses 10 participants: 1 individual from level A2, 4 individuals from level B2 and 5 participants from level C1. The Non-Immersion Subgroup is formed by 26 participants: 11 from level A2, 8 from level B2 and 7 from level C1.

In the following sections, we shall analyze the results of these two subgroups and attempt to find out whether there is a difference between the performance of the Immersion Subgroup and the Non-Immersion Subgroup or not.

[63] The division of the overall population into these subgroups was based on the responses given to the linguistic questionnaire our Hungarian participants had filled out (see Appendix 2).

4.4.1. Overall results of the Immersion vs. Non-Immersion Subgroups

Table 38 below contains the overall results (for determiners and adjectives combined) of these two subgroups.

Level	Immersion Overall Score	(%)	Maximum	(%)	Level	Non-Immersion Overall Score	(%)	Maximum	(%)
A2	87	85,29	102	100	**A2**	882	78,61	1122	100
B2	361	88,48	408	100	**B2**	757	92,77	816	100
C1	458	89,80	510	100	**C1**	650	91,04	714	100
Total	906	88,82	1020	100	**Total**	2289	86,31	2652	100

Table 38: Immersion and Non-Immersion Subgroups: overall scores

If we compare the results of the two subgroups in question to the corresponding results of the control group (99,64%, see Table 5), we cannot find any scores that would even qualify as near native-like. However, the Non-Immersion Subgroup performed over 91% on level B2 and C1, which were the highest results in this case.

As we can see, the Non-Immersion Subgroup was outperformed by the Immersion Subgroup on level A2 by 6,68%, however the exact opposite can be said for level B2 and C1, where the Non-Immersion Subgroup had better results by 4,29% and 1,24%, respectively. The difference between the total scores, however is still in favor of the Immersion Subgroup, achieving 2,51% better results. The better initial results of the Immersion Subgroup could be explained by the very effect of immersion where the acquisition of the L2 seems to occur faster while living among native speakers. However, this initial advantage disappears on higher levels of proficiency in our data, i.e. the effect of classroom training seems to produce better results than the effect of immersion.

As we have previously seen for the Non-Romance L2 Subgroup, the Non-Immersion Subgroup's results also become notably higher from level A2 to level B2 (by 14,16%). This might be due to the environment they are acquiring the L2 in: a classroom context provides explicit grammar rules and specific training in the problematic areas of grammar (such as gender agreement for Hungarian students). Furthermore, the errors that language learners commit are pointed out to them and get corrected, therefore they learn by negative evidence. Meanwhile,

those L2 learners who are acquiring the L2 in an immersion context might not receive such training. Also, regardless of their previous classroom training, they might still be insensitive to some aspects of the input, therefore without explicit corrections they might acquire and go on to use erroneous constructions until they will have received enough input in the L2 to be able to correct their mistakes based on positive evidence (like L1 learners). Since the time that our Immersion Subgroup participants spent in an immersion context was between 3 months and 1 year, they could not have received, apparently, the sufficient amount of input to be able to produce more native-like gender agreement, or perhaps, for the acquisition of this aspect of grammar, naturalistic input might not play an important role at all, since incorrect gender assignment most probably does not have a negative effect on mutual intelligibility between the L2 and L1 speakers.

As for level C1, the Immersion Subgroup shows a slight progress from level B2 to C1 (1,32%), while the Non-Immersion Subgroup shows a slight regress from level B2 to C1 (1,73%). This regress might be due to a classroom training effect, where level C1 no longer focuses on gender agreement (which was introduced and trained on lower levels of proficiency) and has moved onto different aspects of grammar. The slow progress of the Immersion Subgroup might indicate the slow rate at which these participants are becoming more and more sensitive to more aspects of grammar in an immersion context.

We shall continue with the comparison of the results of the two subgroups in question with respect to determiners (Table 39) and adjectives (Table 40).

Level	Immersion DET Score	(%)	Maximum	(%)	Level	Non-Immersion DET Score	(%)	Maximum	(%)
A2	43	84,31	51	100	**A2**	446	79,50	561	100
B2	181	88,73	204	100	**B2**	382	93,63	408	100
C1	230	90,20	255	100	**C1**	325	91,04	357	100
Total	454	89,02	510	100	**Total**	1153	86,95	1326	100

Table 39: Immersion and Non-Immersion Subgroups: scores on determiners

Level	Immersion ADJ Score	(%)	Maximum	(%)	Level	Non-Immersion ADJ Score	(%)	Maximum	(%)
A2	44	86,27	51	100	**A2**	436	77,72	561	100
B2	180	88,24	204	100	**B2**	375	91,91	408	100
C1	228	89,41	255	100	**C1**	325	91,04	357	100
Total	452	88,63	510	100	**Total**	1136	85,67	1326	100

Table 40: Immersion and Non-Immersion Subgroups: scores on adjectives

Again, if we compare the results in Table 39 and 40 with the corresponding results of the control group, we can conclude that the performance of both subgroups is not native-like or even near native-like. Although, level B2 of the Non-Immersion Subgroup did score over 93% with determiners, which is the highest score out of the two subgroups.

As for the success rates for determiners and adjectives on each level of proficiency, the same applies as for the combined results (Table 38). Similarly to the data we have analyzed so far for determiners and adjectives regarding the control group, the overall population and the Romance L2 vs. Non-Romance L2 Subgroups, both the Immersion and the Non-Immersion Subgroups performed minimally better with adjectives in total. The only instance when this is not true is with the Immersion Subgroup on level A2, where they performed slightly better with adjectives than with determiners. Once again, all differences are minimal, however this might indicate that the gender agreement with determiners is more easily acquired, perhaps because of a training effect (the noun is learned together with the determiner), or because of the adjacency of the determiner and the noun or perhaps because the determiner has a scope over the noun.

Also, as we have seen for the combined results, the Non-Immersion Subgroup outperforms the Immersion Subgroup from level B2 forward, however only minimally. As we have explained before, this might be due to the specific training this subgroup receives in a classroom context that makes them become more conscious of their errors and correct them[64]. Meanwhile the Immersion Subgroup might not receive such feedback from the native speakers

[64] Schwartz (1986) stands for a different view, since according to this author negative evidence does not have an effect on the acquisition of an L2. Van Patten & Lee (1990:39) claim, nevertheless that "if there were a lack of negative evidence in adult second language learning, one would need to assume that adults have access to the same innate universal constraints or properties as children. However, there is a crucial difference: it is not clear that the assumption of lack of negative evidence in second language acquisition is warranted".

they interact with in the immersion context. In the latter context, communication is a key factor, where the transmission of information comes before the correctness of the transmitted input.[65] Therefore if native speakers understand what pieces of information the L2 speaker transmits to them (however incorrect the grammar of the L2 speaker may be), the communication process is already successful without any need for perfection with regard to grammaticality.

4.4.2. Data analysis of the Immersion vs. Non-Immersion Subgroups by types of sentences

In this section, we shall analyze the performance of the two subgroups in question with regard to the three sentence types under investigation in this study. Table 41 below contains our overall findings:

Overall Score (%)						
	Type 1		Type 2		Type 3	
Level	Immersion	Non-Immersion	Immersion	Non-Immersion	Immersion	Non-Immersion
A2	79,41	81,55	82,35	74,33	94,12	79,95
B2	88,24	90,07	83,82	93,01	93,38	95,22
C1	89,41	92,44	84,12	85,71	95,88	94,96
Total	87,94	87,10	83,82	83,14	94,71	88,69

Table 41: Immersion and Non-Immersion Subgroups: overall scores broken down to sentence types

When comparing the results in Table 41 with the corresponding results of the control group (see Table 8), we can see that none reach the native-like level. However the results on Type 3 sentences by level C1 Immersion Subgroup participants and level B2 Non-Immersion Subgroup participants are over 95%, which can be considered as near native-like.

Type 1 sentences received the lowest scores among the Immersion Subgroup on level A2, followed by Type 2 sentences and the best-performing Type 3 sentences. This order is the reverse of what we presupposed it would be in Hypothesis 2, according to which we expected Type 1 sentences to have the highest scores, Type 2 sentences to have lower scores and Type 3 sentences to have the lowest scores.

[65] Sometimes they do receive negative evidence from native speakers, but it is on a sporadic basis, not giving consistent input to the L2 learner.

This also strengthens Kayne's (1994) relative clauses' 'raising' hypothesis, because it assumes that sentences like our Type 3 sentences do not actually involve a long distance agreement, but a local agreement, and therefore are not the most complicated to acquire. The Non-Immersion Subgroup obtained slightly different results on level A2: Type 2 sentences show the lowest results, followed by Type 3 sentences and Type 1 sentences received the most correct answers. Again, Type 3 sentences do not pose much of a difficulty, because there might be a local agreement relation after all. The copulative construction (Type 2) seems to be the most difficult to acquire, due to the difficulty of the copulative structure itself. The Non-Immersion Subgroup outperformed the Immersion Subgroup on level A2 by 2,14% with Type 1 sentences, which might be due to their specific classroom training, but all other sentence types seem to be better-dominated by the Immersion Subgroup on this level, notably Type 3 sentences, where the Immersion Subgroup performed 14,17% better, which might be due to the effect of immersion, i.e. input from native speakers.

On level B2 the Non-Immersion Subgroup outperformed the Immersion Subgroup with regard to all three sentence types, the maximum difference between the two subgroups being 9,19% with Type 2 sentences. Also, there is a notable leap from level A2 to B2 among the Non-Immersion Subgroup, where they perform 8,52% better with Type 1 sentences, 11,46% better with Type 2 sentences and 15,27% better with Type 3 sentences. This quick progress might be due to the specific training in this area of grammar that the Non-Immersion Subgroup receives in the classroom context. The progress of the Immersion Subgroup from level A2 to level B2 is not so notable, however level B2 did perform slightly better with Type 1 and Type 2 sentences, but Type 3 sentences show a minimal setback (0,74%). This slower progress might be due to the lack of classroom instruction, i.e. the lack of explicit teaching of specific grammar rules.

On level C1, the Non-Immersion Subgroup reached slightly higher scores than the Immersion Subgroup with Type 1 and Type 2 sentences, but the Immersion Subgroup performed minimally better with Type 3 sentences. This, again, might be due to the specific training in this area received in a classroom context but not in an immersion context. The Immersion Subgroup's results are slightly higher on level C1 than on level B2 with all three sentence types, thus showing generally gradual progress from level A2 to level C1. The Non-Immersion Subgroup's scores, however, indicate a minor setback on level C1 when compared with level B2 with respect to Type 2 and Type 3 sentences, and the figures only show progress with Type 1 sentences. Thus

the results of the Non-Immersion Subgroup generally peak at level B2. This is not an unusual phenomenon though, as we have seen, it might only mean that level C1 students are already being trained in classroom context in other areas of grammar than gender agreement.

We shall now move on to the discussion of results on determiners and adjectives of these two subgroups with respect to sentence types, presented in Table 42 and 43 below:

DET Score (%)						
	Type 1		Type 2		Type 3	
Level	Immersion	Non-Immersion	Immersion	Non-Immersion	Immersion	Non-Immersion
A2	76,47	82,35	82,35	74,87	94,12	81,28
B2	88,24	90,44	83,82	93,38	94,12	97,06
C1	89,41	92,44	84,71	85,71	96,47	94,96
Total	87,65	87,56	84,12	83,48	95,29	89,82

Table 42: Immersion and Non-Immersion Subgroups broken down to sentence types: scores on determiners

ADJ Score (%)						
	Type 1		Type 2		Type 3	
Level	Immersion	Non-Immersion	Immersion	Non-Immersion	Immersion	Non-Immersion
A2	82,35	80,75	82,35	73,80	94,12	78,61
B2	88,24	89,71	83,82	88,97	92,65	93,38
C1	89,41	92,44	83,53	87,39	95,29	94,96
Total	88,24	86,65	83,53	79,86	94,12	87,56

Table 43: Immersion and Non-Immersion Subgroups broken down to sentence types: scores on adjectives

The general tendency we have observed before for determiners and adjectives, namely, that determiners usually receive slightly higher scores, prevails in this case as well. This, once again, can be due to the adjacency of the determiner and the noun or that students are trained to memorize the nouns together with the determiners. However, the Immersion Subgroup breaks this tendency on level A2 with Type 1 sentences, (where 5,88% more adjectives were assigned correct gender values) along with the Non-Immersion Subgroup on level C1 with Type 2 sentences (1,68% more correct gender value for adjectives than for determiners).

Near native-like results can be found with Type 3 sentences: on level C1 among the Immersion group for determiners (96,47%) and on level B2 among the Non-Immersion Subgroup, also for determiners (97,06%). As for Type 3 adjectives, the Immersion Subgroup scored 95,29% on level C1, which can also be considered near native-like.

4.4.3. Data analysis of the Immersion vs. Non-Immersion Subgroups by noun groups

In the following section, we shall take a look at and compare the performances of the two subgroups in question with regard to the noun groups of this investigation. Table 44 below presents these data.

Once again, to be able to present the obtained data on the different noun groups, we use a color-coding system. This time, we established a scale with four degrees, from 60% to 100%, since our participants reached results no lower than 60%, which is higher than chance level (and also higher than some of the Romance L2 vs. Non-Romance L2 Subgroups' corresponding scores, where the lowest degree was at 40%, see Table 35).

Overall Score (%) (Part 1: Noun Groups 1-6)												
	Group 1 <-e>		Group 2 <-gem>		Group 3 <-dade>		Group 4 <-r>		Group 5 <-l>		Group 6 <-s/z>	
Level	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion
A2	75,00	77,27	100	60,61	100	100	100	62,12	100	84,85	66,67	71,21
B2	83,33	93,75	100	97,92	100	97,92	85,42	100	100	93,75	95,83	95,83
C1	86,67	85,71	93,33	95,24	100	100	90,00	95,24	100	100	100	92,86
Total	84,17	84,62	96,67	81,41	100	99,36	89,17	82,69	100	91,67	95,00	84,62

Overall Score (%) (Part 2: Noun Groups 7-12)												
	Group 7 <-é/á/ó>		Group 8 <-a> [+masc]		Group 9 <-ão>		Group 10 <-im>		Group 11 <-a> [+fem]		Group 12 <-o>	
Level	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion
A2	83,33	82,58	66,67	66,67	66,67	75,76	100	89,39	100	100	100	96,97
B2	83,33	81,25	87,50	93,75	79,17	83,33	66,67	91,67	100	93,75	95,83	100
C1	63,33	71,43	93,33	100	81,67	85,71	96,67	90,48	100	100	100	100
Total	73,33	79,17	88,33	83,97	79,17	80,77	85,00	90,38	100	98,08	98,33	98,72

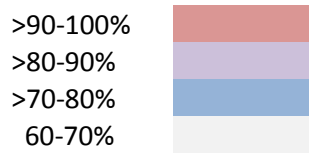


Table 44: Immersion and Non-Immersion Subgroups: overall scores broken down to noun groups

The lowest scores, ranging from 60-70%, were assigned to the light grey color and were attained mostly by speakers from level A2 among both the Immersion and the Non-Immersion Subgroups. The Non-Immersion Subgroup shows results within this range for Groups 2, 4, and 8 on level A2 (all of them being non-default word markers). The Immersion Subgroup however obtained the lowest scores with Groups 6, 8 and 9 on level A2, and, furthermore, on level B2 with Group 10 (word marker <-*im*>, the noun *amendoim* 'peanut' receiving only 25% correct answers, see Appendix 19) and on level C1 for Group 7 (*chaminé* 'chimney' received no correct answers, see Appendix 19), which are all non-default word markers. This, again, might indicate that these vocabulary items are introduced on lower levels of proficiency and do not continue to be trained later on at all.

The blue color assigned to the second lowest degree of the scale (>70-80%) appears with Group 1 on level A2 with both subgroups, with Group 6 and 9 on level A2 with the Non-Immersion Subgroup, with Group 7 on level C1 with the Non-Immersion Subgroup, and with Group 9 on level B2 with the Immersion Subgroup. Also, the total results of Group 7 fall within this range with both subgroups together with the total results of the Immersion Subgroup with Group 9 nouns. These results are still not close to the native-like levels.

The purple color, which is assigned to the degree ranging between >80-90% appears with both subgroups with Group 1 on level C1, with Group 4 total results, with Group 7 on level A2 and B2, with Group 8 total results, and with Group 9 on level C1, among others. If we compare the number of the blue and the purple colored cells, we can conclude that there were more results within he purple range, which means more higher performances (however still not native-like).

The red wine color, assigned to results ranging from >90-100% appears, among many others, with Groups 3, 11 and 12 on all levels of proficiency with both subgroups. As we have previously discussed, these groups contain the, apparently, highly salient disyllabic word marker <-*dade*> and the two default theme indices, respectively, thus appear to be the most accurately acquired noun groups. As for the Immersion Subgroup, they also obtained results within this range with Groups 2 (word marker <-*gem*>) and 5 (word marker <-*l*>)[66], where all levels of

[66] To the word marker <-*gem*> the [+feminine] value is only assigned in Portuguese, while the corresponding forms in other Romance languages have the [+masculine] gender value, e.g. Portuguese: *viagem* 'journey', Spanish: *viaje*, Italian: *viaggio*, French: *voyage*, Catalan: *viatge*. The fact that the Immersion Subgroup received such high results with this word marker might be due to its saliency and thus their sensitivity to the gender value in the input or to specific training.

proficiency scored within this range. The Non-Immersion Subgroup, other than the already mentioned results, scored within this range on level B2 and C1 with Groups 2, 4, 5, 6, 8 and 10, all of which are non-default word markers, however since these are not the lowest levels of proficiencies, these high results might be due the specific classroom training this subgroup receives with regard to non-default word marker gender assignment.

In some instances, the two subgroups either reached the native-like levels or performed even better: with Group 2 the Immersion Subgroup outperformed the control group on level A2 and B2 and reached 100%; also, Group 3 obtained higher results than the control group in almost all instances with both subgroups (expect for level B2 with the Non-Immersion Subgroup); with Group 4 level A2 of the Immersion Subgroup and level B2 of the Non-Immersion Subgroup obtained 100%, thus outperforming the native speakers; Group 5 received 100% results from all levels of proficiency of the Immersion Subgroup and from level C1 of the Non-Immersion Subgroup, this way performing on a native-like level; Group 6 received better than native-like scores (100%) with the Immersion Subgroup on level C1; the performance on Group 8 was native-like with the Non-Immersion Subgroup on level C1; furthermore, Group 9 received native-like results on level A2 of the Immersion Subgroup. All of the aforementioned native-like or better scores were with non-default word markers. We have already seen that very high scores were obtained for the default theme indices, reaching the native-like level (100%) almost on every level of proficiency (expect for the Non-Immersion Subgroup for level B2 and A2 for Group 11 and 12, respectively and for the Immersion Subgroup on level A2 with Group 11).

In the following section we are going to compare the scores of the two subgroups on determiners and adjectives broken down to noun group to be able to examine which item received higher scores. We shall present these data via the color-coding system we applied before for similar results in this study: the yellow cells show scores that are equal for both determiners and adjectives, the green cells represent higher scores and the red cells mean lower scores in comparison.

The word marker *<-l>* always appears with [+masculine] gender value, even among Romance languages in its corresponding forms (Portuguese: *pastel* 'pastry', Spanish: *pastel,* Catalan: *pastís*, Italian: *pastello*, French: pastel), thus might be more transparent and salient than a word marker that can appear with both gender values, such as *<-e>,* which is corroborated by the lower results obtained with the latter word marker.

DET Score (%) (Part 1: Group 1-6)												
	Group 1 <-e>		Group 2 <-gem>		Group 3 <-dade>		Group 4 <-r>		Group 5 <-l>		Group 6 <-s/z>	
Level	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion
A2	66,67	77,27	100	60,61	100	100	100	62,12	100	84,85	66,67	74,24
B2	83,33	93,75	100	100	100	100	87,50	100	100	95,83	95,83	97,92
C1	86,67	85,71	93,33	95,24	100	100	90,00	95,24	100	100	100	92,86
Total	83,33	84,62	96,67	82,05	100	100	90,00	82,69	100	92,31	95,00	86,54

DET Score (%) (Part 2: Groups 7-12)												
	Group 7 <-é/á/ó>		Group 8 <-a> [+masc]		Group 9 <-ão>		Group 10 <-im>		Group 11 <-a> [+fem]		Group 12 <-o>	
Level	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion
A2	83,33	83,33	66,67	66,67	66,67	77,27	100	93,94	100	100	100	96,97
B2	83,33	81,25	91,67	95,83	79,17	83,33	66,67	91,67	100	95,83	91,67	100
C1	63,33	71,43	93,33	100	83,33	85,71	100	90,48	100	100	100	100
Total	73,33	79,49	90,00	84,62	80,00	81,41	86,67	92,31	100	98,72	96,67	98,72

Higher than ADJ Score
Equal to ADJ Score
Lower than ADJ Score

Table 45: Immersion and Non-Immersion Subgroups broken down to noun groups: scores on determiners

ADJ Score (%) (Part 1: Group 1-6)												
	Group 1 <-e>		Group 2 <-gem>		Group 3 <-dade>		Group 4 <-r>		Group 5 <-l>		Group 6 <-s/z>	
Level	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion
A2	83,33	77,27	100	60,61	100	100	100	62,12	100	84,85	66,67	68,18
B2	83,33	83,33	100	95,83	100	95,83	83,33	100	100	91,67	95,83	93,75
C1	86,67	90,48	93,33	95,24	100	100	90,00	95,24	100	100	100	92,86
Total	85,00	76,28	96,67	80,77	100	98,72	88,33	82,69	100	91,03	95,00	82,69

ADJ Score (%) (Part 2: Groups 7-12)												
	Group 7 <-é/á/ó>		Group 8 <-a> [+masc]		Group 9 <-ão>		Group 10 <-im>		Group 11 <-a> [+fem]		Group 12 <-o>	
Level	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion	Immer-sion	Non-Immer-sion
A2	83,33	81,82	66,67	66,67	66,67	74,24	100	84,85	100	100	100	96,97
B2	83,33	81,25	83,33	91,67	79,17	83,33	66,67	91,67	100	91,67	100	100
C1	63,33	71,43	93,33	100	80,00	85,71	93,33	90,48	100	100	100	100
Total	73,33	78,85	86,67	83,33	78,33	80,13	83,33	88,46	100	97,44	100	98,72

Higher than DET Score
Equal to DET Score
Lower than DET Score

Table 46: Immersion and Non-Immersion Subgroups broken down to noun groups: scores on adjectives

If we take a look at Table 45 and Table 46, we can see that most results are equal and that when this is not true, the determiners show high scores in general. The only instances when the adjectives received more correct answers are with Group 1 on level A2 of the Immersion Subgroup and level C1 of the Non-Immersion Subgroup, furthermore with Group 12 on level B2 of the Immersion Subgroup.

The equal or higher general scores on determiners fit into the tendency we have observed before with respect to the corresponding data of the L2 Portuguese speaker population. This, once again, might be due to several things: either it is because of a training effect where the determiner is memorized together with the noun, or because of the adjacency of the determiner and the noun or because of the determiner having a scope over the noun.

Chapter 5
Conclusions

In this chapter, we shall meticulously examine each of our six hypotheses based on our findings exposed in Chapter 4 and shall determine whether they were verified by our data or disproved. For ease of reference, each hypothesis is repeated before its corresponding analysis.

5.1. Analysis of Hypothesis 1

Hypothesis 1:

The increase in the rates of success of our L2 Portuguese participants is directly proportional to the increase in their levels of proficiency.

In other words, the more training these participants have received, the better their results are going to be (a progression effect). Therefore we expect level A2 to have the lowest scores, level B2 to have higher scores and level C1 to have the highest scores.

As we have attested in the previous chapter, the success rates did not necessarily increase in all instances with the increase of the levels of proficiency. While there was usually a general leap, in other words a notable improvement in performances from the elementary level (level A2) to the intermediate level (level B2), this tendency was not clearly visible for the transition from the intermediate level (level B2) to the advanced level (level C1). In fact, level C1 attained minimally lower scores than level B2 in numerous cases, see, for example Table 12, 18, 25, 32, 35, 38, 41 and 44.

The possible reason for level B2 tending to minimally outperform level C1 might lie in the fact that level B2 was most probably undergoing training in the area of gender agreement among other grammar structures at the time of the application of our cloze test, while level C1 has already moved on to different aspects of grammar.

Furthermore, since we did not employ a language proficiency placement test but merely relied on the respective classes our participants attended at university when assigning them to levels of proficiency, which could have led to the misplacement of some participants to incorrect levels of proficiency. Because the differences in performance between level B2 and C1 are minimal, it is possible that some level B2 participants actually possess level C1 competences (at

least in the area of gender agreement), regardless of them still attending level B2 classes. Moreover, since level B1 participants were eventually grouped together with level A2 participants (and not with level B2 participants) because of their low numbers (3 people belonged to level B1) and similar performances to level A2, the lower results of B1 participants compared to B2 participants did not affect the performance of level B2 in a negative way, but in fact might have had a positive effect on the results of level A2 participants as a whole.

Thus Hypothesis 1 was not conclusively confirmed or validated as a whole but it seems to hold true, at least, for the progress observed from level A2 to B2.

5.2. Analysis of Hypothesis 2

Hypothesis 2:

The increase of the distance between the noun and the determiner is inversely proportional to the rates of success in gender assignment of our L2 Portuguese participants.

In other words, the further the adjective is away from the noun it must agree with, the lower the success rates in gender assignment are going to be. Thus we expect Type 1 sentences (local agree construction, attributive adjectives, see in section 3.2.3.1.) to have the highest scores, Type 2 sentences (copulative constructions, see in section 3.2.3.2.) to have lower scores and Type 3 sentences (surface/linear long-distance agree construction with an intervener between the head noun and the adjective, see in section 3.2.3.3.) to have the lowest scores.

Hypothesis 2 above was constructed based on the performance of the control group, i.e. the productions of L1 Portuguese speakers, with the three types of sentences examined in this study. As we saw in Table 8, the overall results of the control group point to a tendency where the distance between the noun and the adjective does seem to matter with regard to correct gender assignment. Thus the L1 Portuguese speakers did, in fact, score the highest with Type 1 sentences (99,80%), slightly lower with Type 2 sentences (99,71%) and the lowest with Type 3 sentences (99,41%). We have to emphasize, however, that the differences between the scores by types of sentences were minimal (0,09% difference between Type 1 and Type 2, and 0,30% between Type 2 and Type 3), which merely indicate a tendency.

Nevertheless, the tendency that L1 speakers have the lowest results with Type 3 sentences (long-distance agreement) might reinforce Brito's (1991) analysis of relative clauses for European Portuguese. According to this analysis, the relative clause is adjunct to the right of the

head noun and then the agreement relation between this noun (which is external to the relative clause) and the adjective (which is internal to the relative and occurring after the gap) cannot be local, only a long-distance agreement one (see more in section 3.2.3.3.).

However, Brito's (1991) analysis only seems to hold true for L1 Portuguese, but not for L2 Portuguese. As we have seen in our findings (Chapter 4), Hypothesis 2 was not confirmed by all of our data from the Portuguese L2 speakers. Table 47 below presents the comparison of total scores by sentences types of all our examined populations and subgroups (see the detailed results in Table 8, 24, 32 and 41):

Comparison of total scores by sentence types (%)						
Scores	**Control Group**	**Overall Population**	**Romance L2**	**Non-Romance L2**	**Immersion**	**Non-Immersion**
Highest	**Type 1** (99,80)	**Type 3** (90,36)	**Type 3** (92,42)	**Type 1** (87,06)	**Type 3** (94,71)	**Type 3** (88,89)
Lower	**Type 2** (99,71)	**Type 1** (87,34)	**Type 1** (87,44)	**Type 3** (85,00)	**Type 1** (87,94)	**Type 1** (87,10)
Lowest	**Type 3** (99,41)	**Type 2** (83,33)	**Type 2** (83,82)	**Type 2** (82,06)	**Type 2** (83,82)	**Type 2** (83,14)

Table 47: Comparison of total scores by sentence types of all examined populations and subgroups

The results in Table 47 show us that, apart from the control group, no other population or subgroup exhibits the presupposed success rates by sentence types described in Hypothesis 2. What is more, except for the Non-Romance L2 Subgroup, the overall population and all other subgroups scored the highest with Type 3 sentences and scored lower with Type 1 sentences. The Non-Romance L2 Subgroup scored the highest with Type 1 sentences and scored lower with Type 3 sentences. Moreover, both the overall population and all examined subgroups scored the lowest with Type 2 sentences. Therefore these tendencies render Hypothesis 2 unconfirmed for L2 Portuguese speakers.

The highest results (or second highest results in the case of the Non-Romance L2 Subgroup) obtained for Type 3 sentences could be explained by Kayne's (1994) raising hypothesis (see section 3.2.3.3.) on the one hand, and by an unforeseen potential error in the construction of our experimental grammatical test on the other hand.

According to Kayne's (1994) head raising analysis, the relation between the head noun and the adjective is local before the head noun moves to the specifier position of the relative clause (with the adjective checking first its gender against the noun before it gets raised). Therefore Portuguese L2 speakers, as opposed to Portuguese L1 speakers, might not treat Type 3 sentences as long-distance agreement sentences, but as local-agreement ones, hence the highest results with this sentence type.

However, we did discover a probable error in the structure of the cloze test that we used as the basis of this research. As we mentioned in Chapter 3, we provided the Hungarian translation of each sentence in our experiment as an attempt to help those participants who had lower levels of proficiency to fully understand the target sentences, however we did not consider that this would facilitate the gender assignment of the target items, since Hungarian is a genderless language. We also described in Chapter 3 that Type 3 sentences involved a [+human] subject (noun) that had the opposite gender value to the head noun between the head noun and the adjective (see section 3.2.3.3.). We assumed that this intervener [+human] noun's gender value would be more salient due to the semantic cue it carried and on which our participants could rely with regard to biological sex and thus grammatical gender assignment. Furthermore, as was described in Chapter 3, all of our target-nouns (head nouns in the case of Type 3 sentences) were [-animate] thus [-human].

The potential problem is that this difference of animacy between the target-noun (head noun) and the intervener noun (subject of the relative clause) is reflected in the relative pronoun in the Hungarian translation, but it is not in the Portuguese original sentence. In Portuguese, the relative pronoun *que* 'that, which' can refer to [+human] and [-human, -animate] objects as well, while Hungarian exclusively uses the [+human] relative pronoun *aki* 'who' to refer to a [+human] object and the [-human] relative pronouns *ami/amely* 'which, that' to refer to a [-human] object. Note, moreover, that Hungarian relative pronouns also exhibit case and number marking, therefore the [-human] relative pronoun *amelyet* in (5.1) exhibits the accusative case marker *-t* and is singular in number:

(5.1) Finalmente vamos comer um jantar que a mãe não comprou preparado. (Portuguese)
that

Végre olyan vacsorát fogunk enni, amelyet anya nem készen vett. (Hungarian)
relative pronoun.ACC.SG.[+human]

'Finally, we are going to eat a dinner that mom did not buy prepared.'

Thus the occurrence of this [-human] relative pronoun in the Hungarian translations of the Portuguese target sentences could have given our participants the means to find the relation between the head noun and the target adjective, even if there was a [+human] intervener between the two target items.

Moving on to the lower results of Type 1 sentences, we could assume that these results were due to the difference of 'noun-adjective' order between Hungarian and Portuguese. As we have seen in Chapter 3, Hungarian attributive adjectives immediately precede the noun, while in Portuguese they immediately follow the noun. Since there is no distance between the determiner and the noun or the noun and the adjective in Type 1 sentences, it might be this typological difference between the two languages that could explain the slight difficulty in the acquisition of Type 1 sentences compared to Type 3 sentences.

Lastly, the lowest obtained results were those of Type 2 (copulative) sentences, which we assume were not due to the distance between the noun and the adjective (although there is a copula between the two items), but to the difficulty in the acquisition of the L2 copulative construction itself. As we have described in Chapter 3, in 3SG and 3PL present indicative sentences Hungarian copulas are dropped, which does not happen in any case in Portuguese. Even though most of our Type 2 target sentences were past indicative where the copulas do appear in Hungarian, our participants still had the most errors with this sentence type. Therefore we assume that the difficulty in the acquisition of L2 Portuguese copulative sentences is at least partly due to a typological difference between the two languages, but other, currently unknown psycholinguistic factors in language processing could also have a played a role. However, to confirm the latter assumption, more research is needed in the future.

In sum, the increase of distance between the noun and the adjective does not seem to make the acquisition of L2 Portuguese gender agreement more problematic, and what is more, distance does not seem to play a role in most errors made by our L2 Portuguese participants, thus

Hypothesis 2 was not confirmed for Portuguese L2 speakers of L1 Hungarian. What does seem to cause difficulties are the typological differences between the two languages and perhaps the sheer genderless nature of the L1, i.e. the complete lack of gender features onto which it is impossible to map new gender related to morphophonological material (as is claimed by the FFFH, Hawkins & Chan 1997), and also some currently unknown other psycholinguistic factors in the processing of L2 Portuguese gender agreement that require more future investigation.

5.3. Analysis of Hypothesis 3

Hypothesis 3:

The noun groups comprising the default theme indices (Group 11-12, see section 3.2.1) will have higher rates of success in gender assignment than the noun groups consisting of non-default word markers (Group 1-10, see section 3.2.1.).

We expect the default theme indices to be acquired and mastered before the non-default word markers. Since the default theme indices *<-o> ([+masculine])* and *<-a> ([+feminine])* possess the most transparent cues for gender value assignment, these are going to be the best controlled noun groups compared to the non-default word markers that are characterized by more opaque and ambiguous gender value assignments.

To provide an overview and comparison of the performances of all populations and subgroups with regard to the 12 noun groups of this research, we present the total results of each noun group by each population and subgroup in Table 48 below (see detailed results in Table 9, 25, 35 and 44). To facilitate the comprehension of this synthesized data, we applied a color coding system where the green colored cells indicate those results that were equal to or higher than the scores of the control group (i.e. are native-like), and the blue colored cells indicate results where the scores were maximum 5% lower than those of the control group (i.e. near native-like).

Comparison of total scores by noun groups (%)												
Total Scores	Group 1 <-e>	Group 2 <-gem>	Group 3 <-dade>	Group 4 <-r>	Group 5 <-l>	Group 6 <-s/z>	Group 7 <-é/á/ó>	Group 8 <-a> [+masc]	Group 9 <-ão>	Group 10 <-im>	Group 11 <-a> [+fem]	Group 12 <-o>
Control Group	**100**	**99,44**	**98,89**	**99,44**	**100**	**99,72**	**99,17**	**100**	**99,44**	**100**	**100**	**100**
Overall	84,49	85,65	99,54	84,49	93,98	87,50	77,55	85,19	80,32	88,89	98,61	98,61
Romance L2	84,29	85,90	100	88,14	94,87	90,71	75,64	89,74	80,13	88,46	98,08	99,36
Non-Romance L2	85,00	85,00	98,33	75,00	91,67	79,17	82,50	73,33	80,83	90,00	100	96,67
Immersion	84,17	96,67	100	89,17	100	95,00	73,33	88,33	79,17	85,00	100	98,33
Non-Immersion	84,62	81,41	99,36	82,69	91,67	84,62	79,17	83,97	80,77	90,38	98,08	98,72

Equal to or higher than Control Group Score
Maximum 5% lower than Control Group Score

Table 48: Comparison of total scores by noun groups of all examined populations and subgroups

As we can see in Table 48, Group 1, 4, 7, 8, 9 and 10 did not receive any coloring therefore these groups of nouns (all of them comprising non-default word markers) did not even perform on a near native-like level, only lower. However, the lowest performance, 75%, is still much higher than chance level.

As we presupposed in Hypothesis 3, our L2 Portuguese participants received very high scores with the default theme indices (Group 11 and 12), which were mostly near native-like, but also native-like among the Non-Romance L2 and Immersion Subgroups. This might be due to the absolute default nature of these theme indices, hence their transparency and easy learnability.

Nevertheless, all of our L2 Portuguese participants seemed to perform even more effortlessly (mostly higher than the native-like level) with another noun group that included the non-default [+feminine] word marker <*-dade*>. As we have explained before, the near-perfect results with this group of nouns might be due to the saliency of this word marker, being disyllabic and always being assigned to the [+feminine] gender value in other Romance languages as well, being similar in its respective forms. For instance, 'liberty' is *liberdade* in Portuguese, *liberdad* in Spanish, *liberté* in French, *libertà* in Italian and *llibertat* in Catalan.

Furthermore, the Immersion Subgroup obtained native-like results with Group 5 (non-default word marker <*-l*>), and near native-like scores with Group 2 (non-default word marker <*-gem*>) and Group 6 (non-default word marker <*-s/z*>). As for the high results with <*-l*>, as we have mentioned before, the gender assignment of nouns that end in this word marker in Romance languages is uniform, these words having the [+masculine] gender value, e.g. 'paper': *papel* in Spanish and Portuguese, *paper* in Catalan, and *papier* in French.

The near native-like scores of the Immersion Subgroup on nouns with the word marker <*-gem*> might be due to its saliency, being a three-letter morpheme and only being assigned to the [+feminine] gender value, hence its transparency. Also, L2 Portuguese learners receive specific training with respect to the plural form of such nouns, since this word marker becomes <*–gens*> in its plural form, which includes a qualitative change in one of the graphemes (<*m*> → <*n*>). Furthermore, throughout the Romance languages, this word marker only assumes the [+feminine] gender value in Portuguese, but is [+masculine] in all others (e.g.: 'journey': Portuguese *a viagem*, Spanish *el viaje*, Italian *il viaggio*, Catalan *el viatge*, French *le voyage*). Thus, since 26 of our 36 L2 Portuguese speakers speak another Romance language, they are specifically trained to pay attention to the different gender values of nouns that have the same/similar forms between Romance languages.

Finally, the high rates of success with the word marker <*-s/z*> among the Immersion Subgroup might be due to the fact that [+feminine] nouns with this word marker are monosyllabic (e.g.: *paz* 'peace', *luz* 'light', *cruz* 'cross'), and [+masculine] nouns with this word marker are polysyllabic (e.g. *lápis* 'pencil', *nariz* 'nose', *país* 'country') (this is true for at least the nouns used in this study). The number of syllables, therefore, could also be a relevant cue for the gender assignment of this word marker.

In sum, Hypothesis 3 did not prove to be valid, since our L2 participants could also acquire groups of nouns with non-default word markers on a near native-like and even on native-like level, not only the two groups of noun with the default theme indices. These high results, once again, could be due to the transparency of the aforementioned word markers, transfer from other Romance languages and to specific training effects, described by Selinker (1972) (see section 1.1.3.3.). The specific successes of the Immersion Subgroup could also have been influenced by the effect of living in an immersion context and having invested in their cultural capitals, thus further improving their language skills (Peirce 1995).

On the other hand, the lack of native-likeness (or even near native-likeness) in the remaining noun groups (Groups 1, 4, 7, 8, 9 and 10) confirms the Bottleneck Hypothesis (Slabakova 2008), according to which functional morphology (to which the marking of Portuguese grammatical gender also belongs) cannot fully be attained by L2 learners.

Furthermore, the findings for Hypothesis 3 also indicate that UG is still partially available after the end of the critical period (see Chapter 1), since some noun-groups were, in fact, acquired in a native-like manner. However, since the parameters for gender features are fixed at [-gender] in L1 Hungarian (thus morphophonological segments for gender do not exist in the L1), our participants have no way of mapping the new morphophonological material for [+gender] features of L2 Portuguese onto any such Hungarian segments (Hawkins & Chan 1997).

Thus we assume that because of the lack of gender features in L1 Hungarian and due to a critical period effect, during the acquisition of gender agreement in L2 Portuguese our L1 Hungarians do not have access to parameterized functional features, therefore we cannot talk about any resetting of such parameter values, thus confirming Hawkins & Chan's (1997) FFFH. This way, instead of an LAD, what seem to be guiding the acquisition of L2 Portuguese gender agreement are the psycholinguistic processes of a latent psychological structure instead described by Selinker (1972) (see section 1.1.3.3.), and perhaps transfer from other acquired Romance languages (*Typological Primacy Model*, Rothman 2010, 2011, 2013).

5.4. Analysis of Hypothesis 4

Hypothesis 4:

Assuming transfer from L1 to L2, we predict that our L2 Portuguese participants will have higher success rates in gender assignment with adjectives compared to determiners.

We expect that, based on the characteristics of the L1, our participants will assign more correct gender values to adjectives than to determiners. Since in Hungarian the case marking is in the end of the noun, we assume that the post-nominal position (the position of the adjective in Portuguese (N+ADJ)) is more salient in the acquisition of L2 Portuguese gender agreement.

According to our findings in Chapter 4, Hypothesis 4 proved to be unconfirmed and, what is more, our L2 Portuguese participants assigned slightly more correct gender values to determiners than to adjectives, see Tables 13, 14, 16, 17, 19, 20, 22, 23, 26, 27, 30, 31, 33, 34, 36, 37, 39, 40, 42, 43, 45 and 46. The only instances where our participants scored slightly better with adjectives than with determiners was with Type 2 sentences by the Non-Romance L2 Subgroup (see Table 33 and 34) and with Type 1 sentences by the Immersion Subgroup (see Table 42 and 43). In all other cases, scores on determiners were higher.

We expected to encounter just the opposite, since Hungarian determiners are invariable as we have seen in Chapter 3, thus do not carry information neither on gender nor on number or case, etc. Hungarian nouns, however, demonstrate case marking (in the end of the noun, i.e. to the right of the radical as is the case of Portuguese gender marking) and can receive many kinds of other suffixes. Therefore, we expected our participants to pay more attention to morphemes in the ends of our target nouns and to the right position (i.e. the position of Portuguese adjectives) in general.

As we have seen in section 4.1.1., the control group assigned slightly more correct gender values to determiners than to adjectives, which could be due to a variety of reasons, as we explained: on the one hand, in the target sentences of this study the distance between the determiner and the noun is always constant, which means that the determiners always immediately precede the nouns (see Chapter 3); secondly, native speakers use a ‘top-down constructive processing’ (Liceras, Díaz & Mongeon 2000), which means that in the first stages of acquisition they treat the determiner and the noun as one syntactic unit (Carroll 1989, MacWhinney 1992), and the triggering of the gender feature only comes later, when they already treat the determiner and the noun as separate syntactic units; and thirdly, according to Franceschina (2005), agreement between the determiner and the noun is acquired before the agreement between the noun and the adjective. This way, following Montrul, Foote & Perpiñán (2008:510), “the gender of the determiner is often used as evidence for lexical assignment of

gender because it appears that monolingual children use the gender of the determiner when predicting the gender of nouns (Carroll 1989, Lew-Williams & Fernald 2008)".

There are two possible mistakes that can occur in the production of gender agreement. One is when both the determiner and the adjective are assigned different gender values than the noun (**o casa bonito* 'the beautiful house). This error is lexical and has to do with the gender assignment of the noun itself. The other mistake happens when either the determiner or the adjective is assigned a different gender value than the noun (**a casa bonito/o casa bonita* 'the beautiful house') which point to a syntactic error (Montrul, Foote & Perpiñán 2008). Our data indicate that there were more correctly assigned gender values to determiners than to adjectives, which means that our participants made more mistakes with adjectives. Consequently, the majority of errors they made was not lexical, but syntactic. Thus our L2 Portuguese speakers seem to have been more successful at acquiring gender as the lexical property of nouns, but did not seem to succeed so well at the acquisition of gender agreement which is a syntactic process.

There might be various reasons behind the fact that L1 Hungarian participants performed similarly to the control group with respect to this particular issue.

First of all, as we have mentioned before, L2 Portuguese students usually learn nouns by the mnemonic memorization of lists where the noun appears together with a determiner or memorize passages of texts where the noun appears preceded by the determiner, thus the students memorize/learn the two items together. Even though L2 learners do not acquire the L2 via a top-down process and most probably do not treat the determiner and the noun as one syntactic unit, (since they tend to rely on visual cues where, orthographically, the two are clearly indicated as separate items with a space between them), the memorization of such DET+N lists could facilitate the acquisition of correct gender values for both items simply because of morphophonological factors (e.g. if the noun ends in <*-a*> then the determiner will also end in <*-a*> and vice versa, e.g. *uma casa* 'a house'). This might also be the reason behind Franceschina's claim formulated for L2 Spanish, according to which "gender agreement between the article and the noun is mastered before the agreement between the adjective and the noun" (2005:113), which is also corroborated by our findings for L2 Portuguese.

Secondly, since the determiners always immediately preceded the nouns in our target sentences, it could have facilitated the correct gender assignment between the determiner and the noun, while the adjectives were placed increasingly further away from the noun, which could

have made the correct gender assignment of the adjective more difficult for some participants. However, as we have seen in section 5.2., distance did not seem to play a notable role in the issue of gender assignment, but sentence structure might have had an effect instead. Once again, the differences between correctly assigned gender values to determiners and adjectives were minimal.

Thirdly, as we have seen in Chapter 3, Portuguese and Hungarian differ in the 'noun-adjective' order. While Portuguese adjectives typically appear on the right of the noun, Hungarian adjectives precede the noun. Therefore, in the end, the left position, where Hungarian nouns appear, might have been more salient for our L2 Portuguese participants. In sum, this typological difference between the two languages could also have contributed to the obtained results.

Finally, transfer from other Romance languages could also have influenced our L2 Portuguese participants, since the 'noun-adjective' and 'determiner-noun' orders in these languages are the same as in Portuguese. Therefore if a participants had mastered the gender agreement between determiner and noun already (which is mastered before agreement between noun and adjective, as Franceschina 2005 stated), then this knowledge might have been transferred onto L2 (or L_n in the case of the study) Portuguese as well.

5.5. Analysis of Hypothesis 5

Hypothesis 5:

Assuming transfer from L2 to L3, we predict that those participants who, apart from Portuguese, learned at least one other Romance language will have higher rates of success compared to those participants who did not learn any other Romance language.

We expect that even if the L1 (Hungarian) does not contain any gender cues and has its value of gender feature fixed for [-gender] (as the FFFH claims, see Chapter 1), previous training in other Romance languages, which are morphosyntactically similar and genetically linked to Portuguese, facilitates the acquisition of Portuguese gender agreement, be it through transfer from this/these other Romance language(s) or via the application of other psycholinguistic processes (Selinker 1972, see Chapter 1) already acquired for the other Romance language(s).

Hypothesis 5 was formulated based on three theories: the Cumulative Enhancement Model (CEM) (Flynn, Foley & Vinnitskya 2004), the L2 status factor model (Bardel & Falk

2007, Falk & Bardel 2010; 2011) and the Typological Primacy Model (TPM) (Rothman 2010, 2011, 2013).

The CEM, in general terms, claims that already acquired language systems can either facilitate L3 or L_n acquisition or their effect remains neutral, but never negative. Flynn, Foley & Vinnitskya (2004) proved that experience in any previously learned language (L1 and L2) can be utilized in the acquisition of any subsequent language (L3 or L_n). Similarly, the L2 status factor model maintains that the L2 has a significant status for morphosyntactic transfer, while the TPM states that typological similarity across languages is ultimately the decisive factor conditioning initial stages of L3 transfer, since the underlying syntax of either the L1 or the L2 is transferred completely. Furthermore the TPM predicts the possibility of facilitative and non-facilitative (negative) transfer as well.

As we have seen, our Romance L2 Subgroup consisted of 26 participants and our Non-Romance L2 Subgroup comprised 10 participants and the latter unfortunately completely lacked participants from level C1. This error was unforeseen when we applied the experimental test on our participants and, unfortunately, could not be corrected due to the very low number of L2 Portuguese university students in Hungary. Nevertheless, we obtained interesting results from both subgroups, which only partly confirm Hypothesis 5.

The general tendency that could be observed when comparing the performances of the Romance L2 and Non-Romance L2 Subgroups was that on the elementary level (level A2) the Romance L2 Subgroup participants notably outperformed the Non-Romance L2 Subgroup participants (see the overall results in Table 29 and Table 32 for the results on Type 2 and Type 3 sentences). However, this initial advantage that speaking another typologically similar L2 could have meant for the Romance L2 Subgroup disappeared on intermediate level (level B2), where the Non-Romance L2 Subgroup outperformed the Romance L2 Subgroup. What is more, level B2 Non-Romance L2 Subgroup participants had similar or even higher scores than advanced level (level C1) Romance L2 speakers.

These findings might corroborate Rothman's (2010, 2011, 2013) TPM in this particular case, since transfer from typologically similar Romance L2s appears to facilitate the acquisition of L3/L_n Portuguese gender agreement in the initial stages (level A2). Furthermore, the initial state of L3/L_n acquisition might even be the typologically similar already acquired Romance L2. The TPM also explains why the Non-Romance L2 Subgroups outperforms the Romance L2

Subgroup from level B2 forward: Romance L2 speakers might experience non-facilitative or negative transfer or transfer errors from the L2(s) to the L3/L_n (Portuguese) (specifically, as we have seen, some of our target nouns could have the opposite gender value in another Romance language than in Portuguese) (see also Montrul 2011). This, therefore, goes against the CEM, because transfer from the L2(s) to the L3/L_n can, in fact, be negative. Similarly, the L2 status factor model's claim, according to which the L2 is the decisive factor for morphosyntactic transfer, was not corroborated by our data, since our Non-Romance L2 Subgroup did outperform the Romance L2 Subgroup starting from the intermediate level. Those participants who did not speak another Romance language did, in fact, speak other languages, but these either did not exhibit gender classes (like English) or did exhibit gender classes but had completely different gender morphologies and classifications (like German) (see Appendix 5). Therefore the morphosyntactic transfer from the non-Romance L2 was probably either not possible or not applicable to a Romance language, like Portuguese.

The overall results of level B2 of the Non-Romance L2 Subgroup could even be classified as near native-like and prove that these participants have at least partial access to UG, since they were able to acquire gender features that are not at all present in their L1 and L2(s) (or are completely different at least). Their results might also mean that specific formal training in a classroom environment is more successful when there is no negative transfer or interference from a different but typologically similar L2. Therefore solely relying on formal instruction and on the psychological processes described in Chapter 1 (Selinker 1972) might lead to a more native-like attainment of the L3/L_n (Portuguese) than also having to filter out negative transfer from another typologically similar L2.

However, the lack of native-likeness in both subgroup's performances point to the fact that parameterized functional features, like gender features, are subject to a critical period (as is claimed by the FFFH, Hawkins & Chan 1997) and haven't been fully acquired (Bottleneck Hypothesis, Slabakova 2008) in spite of (or perhaps because of) the Romance L2 Subgroup already dominating a typologically similar other Romance L2.

This way, Hypothesis 5 only proved to be confirmed for the initial stages of L3/L_n Portuguese acquisition, but not for intermediate or advanced levels, thus, as a whole, Hypothesis 5 was rendered invalid.

5.6. Analysis of Hypothesis 6

Hypothesis 6:

Those L2 Portuguese participants who not only studied Portuguese in a classroom context but were also immersed in Portuguese language and culture for at least 3 months will have higher success rates than those L2 Portuguese participants who exclusively studied Portuguese in a classroom context.

We expect that those participants who lived in an immersion context received more PLD (i.e. naturalistic input) from L1 speakers and also made an 'investment' according to Peirce (1995) (see Chapter 1), therefore they were more successful at language acquisition than those participants who had no access to input from L1 Portuguese speakers and who did not have the need to increase their 'cultural capitals' (Peirce 1995, see Chapter 1).

Based on data from section 4.4, we can conclude that Hypothesis 6 is unconfirmed. As we could see in Table 38 for the overall results and in Table 41 for Type 2 and Type 3 sentences of the Immersion and Non-Immersion Subgroups, it was only on the elementary level/initial stage (level A2) that the Immersion Subgroup outperformed the Non-Immersion Subgroup, but from the intermediate level forward (level B2 and C1) the Non-Immersion Subgroup performed (mostly only minimally) better or very similarly to the Immersion Subgroup.

These results might be due to various factors. The investment made by level A2 of the Immersion Subgroup to strengthen their cultural capital living abroad in an immersion context might have paid off together with their exposure to more and better quality PLD, hence their initial advantage (see Peirce 1995).

However, since the length of immersion was a minimum of 3 months and a maximum of one year (see Appendix 5), this relatively short period might not have been enough for our Immersion Subgroup participants to receive enough PLD to be able to correct their gender errors by positive evidence. Also, according to Patkowski (1980), Birdsong (1992) and Flege & Liu (2001) even if the naturalistic exposure to PLD, i.e. living in an immersion context and interacting with native speakers of the given language, is very lengthy, L2 attainment is still not necessarily native-like (also, see Franceschina 2005:202). Therefore without the negative evidence provided by formal/classroom training, the Immersion Subgroup might not have been able to correct their gender mistakes. Perhaps this is exactly why the Non-Immersion Subgroup was more successful from level B2 forwards, because they were specifically trained in

Portuguese gender agreement and their mistakes were corrected by their teachers, thus they could progress faster to a more native-like state. Even though the Non-Immersion Subgroup had no need to make an investment to further their cultural capital by acquiring the L2 more correctly, they seem to have done so through explicit training and probably by applying the psycholinguistic processes described in Chapter 1 (Selinker 1972).

5.7. Final thoughts

In the end, all of our six initial hypotheses were (at least partially) disproved based on the tendencies that were indicated by our collected data from Portuguese L2 speakers of L1 Hungarian. Nevertheless, we need to emphasize that due to our small sample size; to the unforeseen errors in our experimental test; and because we did not apply qualitative statistical analyses to our data, these results are in no way absolutely decisive or completely reliable and because of the minimal differences between most scores, they merely indicate tendencies.

As for future research in the specific field of acquisition of gender agreement in L2 Portuguese by Hungarian L1 speakers, it would be worthwhile to apply more types of tests and methods (like grammaticality judgement tests and/or the elicitation of spontaneous speech) to obtain a deeper and wider understanding of how the acquisition of L2 Portuguese gender agreement is processed in the brain of Hungarian L1 speakers. It might also lead to interesting findings to compare the results on the acquisition of L2 Portuguese gender agreement of L1 Hungarians with the results of other L1 [-gender] and L1 [+gender] speakers.

Nevertheless, our study confirmed that there seems to exist a critical period effect for the acquisition of L2 gender features and that even if our participants were exposed to extended PLD, they still did not attain native-like Portuguese gender agreement in all instances (thus proving Orwell's problem), all of which indicate the validity of the FFFH (Hawkins & Chan 1997) to the specific issue of the acquisition of L2 Portuguese gender agreement. We also found that, since the vast majority of the results were notably higher than chance level, our L1 Hungarian speakers were, in fact, capable of acquiring L2 Portuguese gender agreement to a lesser or higher extent (sometimes even in a near/native-like manner), which indicates that they have partial access to UG. Furthermore, transfer from other Romance L2s and the effect of immersion also seemed to have contributed to the higher rates of success of our participants, but predominantly in the initial stages of SLA. Finally, the psycholinguistic processes described in Chapter 1 by Selinker (1972)

(also, see Franceschina 2005:198-201) and specific formal training also seemed to be relevant for the successful acquisition of L2 Portuguese gender agreement.

References

Afonso, C., Gonçalves, A. & Freitas, M. J. (2014). A princesa ficou *adormir ou a dormir? Dados sobre a consciência da unidade palavra em Português europeu. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto, vol. 9,* Porto: FLUP.

Adger, D. (2003). *Core Syntax. A minimalist approach*. Oxford: OUP.

Adger, D. & Ramchand, G. (2003). Predication and equation. *Linguistic Inquiry 34 (3),* pp. 325-359.

Adger, D. & Ramchand, G. (2006). Psych nouns and predication. In David, C., Deal, A. & Zabbal, Y. (eds.), *Proceedings of NELS 36,* Amherst: GLSA, pp. 89-102.

Alexandre, N. (2000). *A estratégiar resumptiva em relativas restritivas do português europeu.* Master's Dissertation, Lisbon: Universidade de Lisboa.

Al-Hamad, M. (2002). The role of input in SLA of syntax. Paper presented at *Looking at Language Acquisition II, RCEAL,* University of Cambridge, 7 May 2002.

Balazs, J. E. (2012). *The syntax of small clauses*. Thesis, Cornell University

Baldé, N. R. (2011). *A aquisição do artigo em português L2 por falantes de L1 russo.* Master's Dissertation. Lisbon: Universidade de Lisboa

Balogh, J., Haader, L., Keszler, B., Kugler, N., Laczkó, K. & Lengyel, K. (2000). *Magyar Grammatika*. Budapest: Nemzeti Tankönyvkiadó.

Bardel, C. & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research 23(4),* pp. 459–484.

Bartning, I. (2000). Gender agreement in L2 French: Pre-advanced vs advanced learners. *Studia Linguistica, 54* (2). *Special Issue: A selection of papers from EUROSLA 9*, Lund, Sweden, June 10–12, 1999, pp. 225–237.

Benkő, L. & Imre, S. (eds.) (1972). *The Hungarian language.* Budapest: Akadémiai Kiadó.

Birdsong, D. (1992). Ultimate attainment in SLA. *Language, 68* (4), pp. 706–747.

Birdsong, D. (1994). Decision making in second language acquisition. *Studies in Second Language Acquisition, 16*, pp. 169–182.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? *Gass,* S. *M.& Schachter, J. (eds.), Linguistic Perspectives on SLA,* pp. 41–68. Cambridge: CUP.

Brito, A. M. (1991). *Sintaxe das Orações Relativas em Português: Estrutura, Mecanismos Interpretativos e Condições sobre a Distribuição dos Morfemas Relativos,* PhD Dissertation, Porto: INIC.

Bruhn de Garavito, J. & White, L. (2000). Second language acquisition of Spanish DPs: The status of grammatical features. *Howell, S. C., Fish, S., & Keith-Lucas, T. (eds.), BUCLD 24: Proceedings from the 24th annual Boston University Conference on Language Development* (1999), pp. 164–175. Somerville, MA: Cascadilla.

Brysbaert, M., Mitchell, D.C., (2000). The failure to use gender information in parsing: A comment on van Berkum, Brown, and Hagoort., 1999. *Journal of Psycholinguistic Research 29*, pp. 453-466.

Burton-Roberts, N. (1997). *Analysing sentences: An introduction to English grammar*. London: Longman.

Carroll, S. E. (1989). Second language acquisition and the computational paradigm. *Language Learning, 39* (4), pp. 535–594.

Carvalho, J. G. Herculano de (1967, 1984). *Teoria da linguagem/Natureza do fenómeno linguístico e a análise das línguas II*. Coimbra: Coimbra Editora.

Castro, A. (2010). Aquisição de morfologia de plural em português europeu: sobre a produtividade das regras. *Textos Seleccionados, XXV Encontro Nacional da Associação Portuguesa de Linguística,* Porto, APL, 2010, pp. 277-289.

Chini, M. (1995). Un aspect du syntagme nominal en italien L2: Le genre. *AILE, 5*, pp. 15-42.

Chomsky, N. (1986). *Knowledge of Language: Its nature, origins, and use*. New York, NY: Praeger.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Clahsen, H. & Muysken, P. (1986). The availability of universal grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research, 2*, pp. 93–119.

Clahsen,H., Eisenbeiss, S.,& Vainikka,A. (1994). The seeds of structure. *Schwartz, B. D. & Hoekstra, T. (eds.), Language Acquisition Studies in Generative Grammar*, pp. 85–118. Amsterdam: John Benjamins.

Corrêa, L., Name M. C. & Ferrari-Neto, J. (2004). O processamento de informação de interface na aquisição de género e de número no português brasileiro. *Letras de Hoje, Porto Alegre, vol. 39 (3)*, pp. 123-137.

Corbett, G. G. (1991). *Gender*. Cambridge: CUP.

Corbett G G (2005). 'The number of genders,' 'Sex-based and non-sex-based gender systems,' and 'Gender assignment systems' [three chapters and maps]. *Haspelmath, M, Dryer, M, Gil, D & Comrie, B (eds.) World Atlas of Language Structures*. Oxford: Oxford University Press.

Corbett, G. G. (2006). Gender, grammatical. *Encyclopedia of Language & Linguistics*, pp. 749-756.

Corbett, G. G. (2013). Number of Genders. *Dryer, M. S. & Haspelmath, M. (eds.) The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Corder, S.P. (1967). The significance of learners' errors. *International Review of Applied Linguistics 5/4*, pp. 161-170.

Corder, S.P. (1981). *Error Analysis and Interlanguage.* Oxford: Oxford University Press.

Cruz Payão, L. (2010). Aquisição de fonologia: a influência do acento e o preenchimento de unidades prosódicas em dados de fala de duas crianças entre 1;0.4 e 2;1.10 de idade, em contato com o português brasileiro falado em Alagoas e Pernambuco. Maceió: UFAL, PhD Dissertation.

Costa, J. (ed.) (2000). *Portuguese syntax. New comparative studies*. Oxford: Oxford University Press.

Csaba, M., Fodor, A. & Szijj, I. (2006). *Grão a grão… Exercícios de gramática portuguesa. Portugál nyelvtani feladatgyűjtemény.* Budapest: Typotex.

Csaba, M., Perjés, M. & Szijj, I. (2007). Como se diz em português? Portugál nyelvkönyv kezdőknek és haladóknak. Budapest: Typotex.

Croft, W. (1991). *Syntactic Categories and Grammatical Relations*. Chicago, London: University of Chicago Press.

Dato, D. P. (1975). On psycholinguistic universals in children's learning of Spanish. *Dato, D. P. (ed.), Developmental Linguistics: Theory and applications*, pp. 235–254. Washington, DC: Georgetown University Press.

De Angelis, G. (2007). *Third or Additional Language Acquisition*. Clevedon: Multilingual Matters.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *SSLA, 22*, pp. 499-533.

Dewaele, J. M. & Veronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, *4* (3), pp.275–297.

Doughty, C. J. & Long, M. (eds.) (2005). *The handbook of second language acquisition. (Blackwell handbooks in linguistics).* MA: Blackwell Publishing.

É. Kiss, K., Kiefer, F. & Siptár, P. (2003). *Új magyar nyelvtan*. Budapest: Osiris Kiadó.

Ellis, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.

Fábregas, A. & Pérez, I. (2008). Gender agreement on adverbs in Spanish. *Journal of Portuguese Linguistics, vol. 7 (2)*, pp. 25-45.

Falk, Y. & Bardel, C. (2010). The study of the role of the background languages in third language acquisition. The state of the art. *International Review of Applied Linguistics in Language Teaching. IRAL 48(2-3)*, pp. 185–220.

Falk, Y. & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research 27(1),* pp. 59–82.

Felix, S. W. (1985). More evidence on competing cognitive systems. *Second Language Research, 1*, pp. 47–72.

Felix, S.W. (1987). *Cognition and Language Growth*. Dordrecht: Foris.

Fernández García, M. (1999). Patterns in gender agreement in the speech of L2 learners. *Gutierrez Rexach, J. & Martinez Gil, F. (eds.), Advances in Hispanic Linguistics. Papers from the Second Hispanic Linguistics Symposium,* pp. 3-15. Somerville, MA: Cascadilla.

Ferreira, T. (2011). *Padrões na aquisição/aprendizagem da marcação do género nominal em português como L2.* Universidade de Coimbra, Faculdade de Letras. Coimbra: Master's Dissertation.

Finnemann, M. D. (1992). Learning agreement in the NP: The strategies of three first-year Spanish students. *International Review of Applied Linguistics* (*IRAL*), *30* (2), pp.121–136.

Flege, J. E. & Liu, S. (2001). The effect of experience on adults' acquisition of a second language [December 2001]. *Studies in Second Language Acquisition, 23* (4), pp. 527–552.

Flynn, S. (1987). *A Parameter-setting Model of SLA*. Dordrecht: Reidel.

Flynn, S., Foley, C. & Vinnitskaya, I. (2004). The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition of relative clauses. *International Journal of Multilingualism, vol. 1,* pp. 3-16.

Foley, W. A. & Van Valin, R. D. J. (1984). *Functional Syntax and Universal Grammar*. Cambridge: CUP.

Foucart, A. (2008). *Grammatical gender processing in French as a first and second language.* Université Aix- Marseille I / University of Edinburgh. PhD Dissertation.

França, A. (1997). Comparação entre aquisição de português e inglês como L2. *Revista de Crioulos de Base Ibérica , 9*, pp.43-51.

Franceschina, F. (2001a). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research, 17* (3), pp.213–247.

Franceschina, F. (2001b). Where lies the difference between native and non-native grammars? Evidence from the L2A of Spanish. *Foster-Cohen, S. & Nizegorodcew, A. (eds.), EUROSLA Yearbook 1*, pp. 143–158. Amsterdam: John Benjamins.

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam: Benjamins.

Freitas, J. M., Alves, D. & Costa, T. (2007). *O conhecimento da língua: Desenvolver a consciência fonológica.* Ministério da Educação.

Freitas, J. M., Gonçalves, A. & Gonçalves, F. (2010). Aspectos fonológicos e morfossintácticos da aquisição dos ditongos nasais em português europeu. *Aquisição do Português europeu: Recursos e resultados linguísticos*. Lisbon: Centro de Linguística da Universidade de Lisboa (CLUL).

Friedmann, N., Belletti, A. & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua 119*, pp. 67–88.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: an introdutory course.* New York: Taylor & Francis e-Library.

Gass, S. M. & Selinker, L. (1992). Language transfer in language learning. *Applied Linguistics, vol. 16 (4)*, 1995, pp. 517-521.

Godinho, A. (2010). A aquisição da concordância de número e a sua relação com a aquisição da concordância de género: um estudo realizado com aprendentes chineses de português L2. *Língua portuguesa: ultrapassar fronteiras, juntar culturas.* Marçalo, M. J., Lima-Hernandes, M. C., & Esteves, E. (eds.), pp. 28-55.

Gonçalves, F., Guerreiro, P., Freitas, J. M. & Sousa, O. (2011). *O conhecimento da língua: Percurso de desenvolvimento*. Ministério da Educação.

González, G. (1970). *The acquisition of Spanish grammar by native Spanish-speaking children.* Ph.D. dissertation, University of Texas.

Gouveia, M. (2004). Considerações sobre a categoria gramatical de género: sua evolução do latim ao português arcaico. *BIBLOS*, vol. II, pp. 443-475.

Han, Z. (2004). *Fossilization in Adult Second Language Acquisition*, Clevedon, UK.: Multilingual Matters.

Hawkins, R. D. & Chan, C. (1997). The partial availability of UG in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research, 13* (3), pp. 187–226.

Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. *Prévost, P. & Paradis, J. (eds.), The acquisition of French in different contexts,* pp.175–205. Amsterdam: Benjamins.

Hengeveld, K. (1992). *Non-verbal Predication*. Berlin, New York: Mouton de Gruyter.

Hernández Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Madrid: Siglo XXI de España Editores.

Hockett, C. F. (1958). *A Course in Modern Linguistics*. New York, NY: Macmillan.

Hyltenstam, K. & Abrahamsson, N. (2003). Maturational constraints in SLA. *Doughty, C. J. & Long, M. H. (eds.) Handbook of Second Language Acquisition,* pp. 53-588. Malden, MA: Blackwell.

Ionin, T., Ko, H. & Wexler, K. (2004). Article Semantics in L2 Acquisition: The Role of Specificity. *Language Acquisition* 12(1), pp. 3-69.

Johnson, J. & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, pp. 60–99.

Kayne, R. S. (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kenesei, I., Vágó, R. M. & Fenyvesi, A. (1998). *Hungarian*. London, New York: Routledge.

Krashen, S. D. (1982). *Principles and practice in SLA*. Oxford: Pergamon Press.

Krashen, S. D. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.

Krashen, S. D. (1981, 1988), *Second Language Acquisition and Second Language Learning,* Oxford: Pergamon Press.

Kuchenbrandt, I. (2005). Gender Acquisition in Bilingual Spanish. *Cohen, J., McAlister, K.T., Rolstad, K. & MacSwan J. (eds.) Proceedings of the 4th International Symposium on Bilingualism,* pp. 1252-1263. Somerville, MA: Cascadilla Press.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
Lardiere, D. (2000). Mapping features to forms in SLA. In J. Archibald (Ed.), *Second language acquisition and linguistic theory*, pp. 102–129. Oxford: Blackwell.

Lardiere, D. (2008). Feature-Assembly in Second Language Acquisition. *Liceras, J., Zobl, H. & Goodluck H. (eds.), The role of formal features in second language acquisition*, pp. 106– 140. New York: Lawrence Erlbaum Associates.

Lardiere, D. (2013). Native language influence at the L2 steady state. *International Congress of Linguistics*, Geneva, July 2013.

Leiria, I. (2000). Português língua segunda e língua estrageira: investigação e ensino. 1º *Congresso do Português Língua Não-Materna*, October 1999, Lisbon.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York, NY: JohnWiley.

Lew-Williams, C., & Fernald, A. (2008). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18* (3), pp. 193–281.

Liceras, J. M., Diaz, L., & Mongeon, C. (2000). N-drop and determiners in native and non-native Spanish: More on the role of morphology in the acquisition of syntactic knowledge. *Leow, R. P. & Sanz, C. (eds.), Spanish Applied Linguistics at the Turn of the Millennium. Papers from the 1999 Conference on the L1 and L2 Acquisition of Spanish and Portuguese, Georgetown University*, pp. 67–96. Somerville,MA: Cascadilla.

Lightbown, P. M. (1985). Great Expectations: SLA research and classroom teaching. *Applied Linguistics, 6* (2), pp. 173–189.

López Ornat, S. (1997). What lies between a pre-grammatical and a grammatical representation? Evidence from nominal and verbal form-function mappings in Spanish from 1;7 to 2;1. *Perez-Leroux, A. T & Glass, W. R. (eds.), Contemporary Perspectives on the Acquisition of Spanish*, pp. 3–20. Somerville, MA: Cascadilla.

MacWhinney, B. (1992). Transfer and competition in L2 learning. *Harris, R. J (ed.), Cognitive Processing in Bilinguals*, pp. 371–390. Amsterdam: North-Holland.

Mariotto, E. (2014). Processamento da concordância de género por aprendentes de português como língua estrangeira: Evidências de um estudo de leitura automonitorada. FLUL: Master's Dissertation.

Mateus, M. H. & d'Andrade, E. (2000). *The phonology of Portuguese*. Oxford: Oxford University Press.

Mateus, M. H. *et al.* (2003), *Gramática da Língua Portuguesa*. Lisbon: Editorial Caminho.

Mattoso Câmara, J. (1966), *Considerações sobre o gênero em português*. In: *Revista Brasileira da Lingüística Teórica e Aplicada, volume I, n. 2,* São Paulo: Centro de Lingüística Aplicada, pp. 19.

Mattoso Câmara, J. (1994). *Estrutura da língua portuguesa*. Petrópolis: Vozes.

Meisel, J. M. (2011) *First and second language acquisition. Parallels and differences*. Cambridge: Cambridge University Press.

Miguel, S. M. (2004). O sintagma nominal em português europeu. Posições de sujeito. PhD Dissertation. Lisbon: FLUL.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: the effects of age and context of acquisition. *Language Learning*, pp. 503-553.

Montrul, S. (2011). Multiple interfaces and incomplete acquisition. *Lingua, 121*, pp. 591-604.

Montrul, S. (2013). How "Native" Are Heritage Speakers? *Heritage Language Journal*, 10 (2), pp. 15-39.

Nichols, J. (1992), *Linguistic Diversity in Space and Time*, Chicago, IL: University of Chicago Press.

Nunes, J. J. (1919, 1975). *Compêndio da gramática histórica portuguesa*. Lisbon: Livraria Clássica Editora.

Oliveira, I. (2006). *A aquisição do sistema vocálico por falantes de português europeu como língua não materna.* Master's Dissertation. Lisbon: FLUL.

Partee, B. H. (1998). Copula inversion puzzles in English and Russian. *UMOP 23: Issues in Semantics and its Interface, Kusumoto, K. and Villalta, E. (eds).* (2000) University of Massachusetts, Amherst: GLSA, pp. 198-208.

Pascual Y. Cabo, D. & Rothman, J. (2012). The (Il)Logical Problem of Heritage Speaker Bilingualism and Incomplete Acquisition. *Applied Linguistics 2012: 33/4,* pp. 450–455.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30*, pp. 449–472.

Penfield,W. & Roberts, L. (1959). *Speech and Brain Mechanisms*. New York, NY: Athenaeum.

Peirce, B. N. (1995). Social identity, investment and language learning. *TESOL Quarterly, 29* (1), pp. 9–31.

Pinto, J. (2015). *A aquisição do género e da concordância de género, em português língua terceira ou língua adicional.* Lisbon: CLUL.

Polinsky, M. (2008). Gender under Incomplete Acquisition: Heritage Speakers' Knowledge of Noun Categorization. *Heritage Language Journal*, *6*(1), pp. 40-71.

Prévost, P. & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16* (2), pp. 103–133.

Pustet ,R. (2003). *Copulas. Universals in the categorization of the lexicon*. Oxford: Oxford University Press.

Radford, A. (2004). *English syntax: An introduction.* Cambridge, UK: Cambridge University Press.

Reuland, E. (1993). Reflections on knowledge and language. *Reuland, E. & Abraham, W. (eds.), Knowledge and Language. vol. I*: *From Orwell's problem to Plato's problem,* pp. 11–27. Dordrecht: Kluwer Academic.

Rothman, J. (2010). On the Typological Economy of Syntactic Transfer: Word Order and Relative Clause Attachment Preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics, IRAL*, 48, 2/:3, pp. 243-271.

Rothman, J. *(*2011). L3 Syntactic Transfer Selectivity and Typological Determinacy: The Typological Primacy Model. *Second Language Research,* 27,1, pp. 107-127.

Rothman, J. (2013). Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. *Baauw, S., Dirjkoningen, F. & Pinto, M. (eds.), Romance languages and linguistic theory, 2011*, pp. 217–247. Amsterdam: John Benjamins.

Rothman, J. & Cabrelli, J. (2010). What variables condition syntactic transfer: A look at the L3 initial state. *Second Language Research* 26, pp. 189–218.

Rourke, P. & Van Petten, C. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain Research, 1392:* pp. 62-79.

Schachter, J. (1988). Second language acquisition and its relationship to UG. *Applied Linguistics, 9*, pp. 219–235.

Schnell de Acedo, B. (1994). Early morphological development: The acquisition of articles in Spanish. *Sokolov, J. L. & Snow, C. E. (eds.), Handbook of Research in Language Development using CHILDES,* pp. 210–253. Hillsdale, NJ: Lawrence Erlbaum.

Schouten, A. (2009). The Critical Period Hypothesis: Support, Challenge, and Reconceptualization. *TESOL & Applied Linguistics* , *9*, pp. 1-16.

Schumann, J. H. (1978). *The Pidginization Hypothesis: A model for second language acquisition*. Rowley, MA: Newbury House.

Schwartz, B. D. (1986). The epistemological status of second language acquisition. *Second Language Research, 2*, pp. 120-159.

Schwartz, B. D. & Sprouse, R. A. (1996). L2 cognitive states and the 'full transfer/full access' model. *Second Language Research, 12* (1), pp. 40–72.

Scovel, T. (1988). *A Time to Speak: A psycholinguistic inquiry into the critical period for human speech*. New York, NY: Newbury House.

Seliger, H. W. (1978). The implications of a multiple critical period hypothesis for second language learning. *Ritchie, W. C (ed.), Second Language Acquisition Research. Issues and implications,* pp. 11–19. New York, NY: Academic Press.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics* (*IRAL*), *10* (2), pp. 209–231.

Singleton, D. M. (1995). Introduction: A critical look at the Critical Period Hypothesis in SLA. *Singleton, D. M & Lengyel, Z. (eds.), The Age Factor in Second Language Acquisition*, pp. 1–29. Clevedon: Multilingual Matters.

Slabakova, R. (2008). *Meaning in the Second Language.* Mouton de Gruyter, Berlin.

Slabakova, R. (2009). How is inflectional morphology learned?. *EUROSLA Yearbook* 9, pp. 56-75.

Smith, N. & Tsimpli, I. M. (1995). *The Mind of a Savant: Language learning and modularity*. Oxford: Blackwell.

Soler, M. R. (1984). Adquisición y utilización del artículo. *Siguan, M. (ed.), Estudios sobre psicología del lenguaje infnatil*. Madrid: Pirámide.

Spencer, A. (1999). Gender as an inflectional category. *Essex Research Reports in Linguistics, 25*, 35–72.

Stassen, L. (1997). *Intransitive Predication*. Oxford: Clarendon.

Tarone, E. (1977). Conscious communication strategies in interlanguage. *Brown, H. D., Yorio, C. A. & Crymes, R. C. (eds.), On TESOL '77*: *Teaching and Learning English as a Second Language* , pp. 194-203. Washington, DC: Teachers of English to Speakers of Other Languages.

Tarone, E. (2006). Interlanguage. *Brown, Keith. Encyclopedia of Language and Linguistics*. Boston: Elsevier. pp. 747–751.

Tavares, A. (2008). *Português XXI Nível A1*. Lisbon: Lidel.

Tolbert, K. (1978). *The Acquisition of Grammatical Morphemes: A cross-linguistic study with reference to Mayan and Spanish.* PhD Dissertation, Harvard University.

Tsimpli, I. M. & Smith, N. (1991). Second language learning: Evidence from a polyglot savant. *UCL Working Papers in Linguistics, 3*, pp. 171–184.

Tsimpli, I. & Sorace, A. (2006). Differentiating Interfaces: L2 Performance in Syntax-Semantics and Syntax-Discourse Phenomena. *Proceedings of the Annual Boston University Conference on Language Development, 30 (2*), pp. 653-664.

Van Patten, B. & Benati, A. (2010). *Key terms in second language acquisition*. NY: Continuum.

Van Patten, B. & Lee, J. F. (1990). *Second Language Acquisition and Foreign Language Learning: Perspectives on Research and Practice*. Clevedon, UK: Multilingual Matters.

Vilela, M. (1973), Considerações sobre o género gramatical. *Revista do Porto da Faculdade de Letras da Universidade do Porto,* vol. I, Porto: Série de Filologia, pp. 139-150.

Villalva, A. (1994). *Estruturas Morfológicas.* Universidade de Lisboa, Faculdade de Letras. Lisboa: Tese de Doutoramento.

Walsh, T. M. & Diller, K. C. (1981). Neurolinguistic considerations on the optimum age for second language learning. *Diller, K. C (ed.), Individual Differences and Universals in Language Learning Aptitude,* pp. 3–21. Rowley, MA: Newbury House.

White, L. (1986). Markedness and parameter-setting: some implications for a theory of adult second language acquisition. *Eckman, F., Moravcsik, E. & Wirth, J. (eds.), Markedness*, pp. 309-327. Plenum Press.

White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: CUP.

White, L. (2008). Different? Yes. Fundamentally? No. Definiteness Effects in the L2 English of Mandarin Speakers. *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007), ed. Slabakova, R. et al.,* pp. 251-261. Somerville, MA: Cascadilla Proceedings Project.

Appendix 1: Declaration of Consent

DECLARAÇÃO DE CONSENTIMENTO

No âmbito da realização de um estudo para a obtenção do grau de Mestrado na Faculdade de Letras da Universidade de Lisboa, eu, Veronika Lacsán, peço a sua colaboração para usar os dados relativos à aquisição de Português L2. A participação no estudo implica preencher espaços em branco que ocorrem em certas frases. No tratamento dos dados recolhidos, é salvaguardado o seu anonimato e os dados serão usados exclusivamente no âmbito da investigação.

Lisboa, 27. 01. 2015

_____

(Veronika Lacsán)

✂ ----------

Eu, _____, aceito participar no estudo da responsabilidade de Veronika Lacsán.

Assinatura:_____

Data:_____

Appendix 2: Linguistic questionnaire

Perfil Linguístico do Informante

1.**Nome e Apelido**:_____

2. **Email**: _____

3. **Data do nascimento (dd/mm/aaaa**): ____/____/_____

4. **Língua(s) materna(s)**:_____ _____ _____

5. **Que língua(s) fala em casa?** _____ _____ _____

6. **Língua(s) de ensino**:

Antes da Universidade:_____

Na Universidade: _____

7. **Língua(s) estrangeira(s) (por ordem decrescente de proficiência)**

1._____ 2._____ 3._____

4. _____ 5._____ 6._____

8. **Nível de proficiência em português:** A2 B1 B2 C1 C2

9. **Há quanto tempo estuda português?**_____

10. **Onde aprendeu português?**_____

11. **Em que situações usa o português?**

Família e amigos Contextos académicos Trabalho Nunca

12. **Estadia num país lusófono:** Sim/Não

Onde?_____

Quando?_____

Quanto tempo?_____

Appendix 3: Target sentences and distractors

Target sentences:

Type 1 sentences:

Group 1:

(1) Segundo a previsão do tempo, hoje vai ser uma noite fria na Serra da Estrela.

Az időjárás előrejelzés szerint ma éjjel hideg lesz a Serra da Estrelán.

According to the weather forecast, tonight is going to be a cold night in the Serra da Estrela.

(2) Amanhã vamos ver um filme italiano no Campo Pequeno com dois amigos meus.

Holnap két barátommal elmegyünk megnézni egy olasz filmet a Campo Pequeno-ban.

Tomorrow, we are going to see an Italian movie in Campo Pequeno with two friends of mine.

Group 2:

(3) É melhor dormirmos mais algumas horas, pois hoje fizemos uma viagem longa.

Jobb lesz, ha alszunk még pár órát, mert ma hosszú utat tettünk meg.

We'd better sleep a few more hours, since we have done a long journey today.

Group 3:

(4) O Pedro gosta das cidades antigas da tua terra, porque têm uma história rica.

Pedro kedveli az országod ókori városait, mivel gazdag történelemmel büszkélkedhetnek.

Pedro likes the ancient cities of your country because they have a rich history.

Group 4:

(5) A minha prima só tem roupas vermelhas porque é a cor favorita dela.

Az unokahúgomnak csak piros ruhái vannak, mert ez a kedvenc színe.

My cousin only has red dresses because it's her favorite color.

(6) O açúcar branco não é muito saudável.

A kristálycukor nem túl egészséges.

White sugar is not very healthy.

Group 5:

(7) Na culinária portuguesa há uns pastéis deliciosos que se chamam queijadas de Sintra.

A portugál konyhaművészethez tartozik egy finom sütemény, amelyet sintrai sajtos kosárkának hívnak.

There are some delicious pastries in Portuguese cuisine called Sintra cheesecakes.

Group 6:

(8) O meu professor usa sempre um lápis vermelho para corrigir os nossos testes.

A tanárom mindig egy piros ceruzát használ a dolgozataink kijavításához.

My teacher always uses a red pencil to correct our tests.

(9) A primeira coisa que os atores veem no palco é a luz branca das lâmpadas.

A legelső dolog, amelyet a színészek meglátnak a színpadon, az a reflektorok fehér fénye.

The first thing that actors see on stage is the white light of the reflectors.

Group 7:

(10) É muito importante realizar inspeções ao aquecimento das casas, pois uma chaminé suja pode causar problemas graves.

Nagyon fontos, hogy ellenőriztessük a házak fűtőrendszerét, mert egy koszos kémény komoly problémákat tud okozni.

It's very important to carry out inspections of the heating (systems) of houses, since a dirty chimney can cause serious problems.

(11) Se não queres estar acordado até à uma de manhã, bebe um café fraco.

Ha nem akarsz ébren lenni egészen hajnali egyig, akkor csak egy gyenge kávét igyál.

If you don't want to be awake until 1 in the morning, just drink a weak coffee.

Group 8:

(12) Ouvi nas notícias que a National Geographic publicou um mapa modernissimo sobre a superfície de Marte.

Hallottam a hírekben, hogy a National Geographic közzétett egy hihetetlenül modern térképet a Mars felszínéről.

I heard it in the news that the National Geographic published a super modern map of the surface of Mars.

Group 9:

(13) O rato tem um coração pequeno.

Az egérnek kis méretű szíve van.

The mouse has a small heart.

(14) Alguém que tem predisposição para usar a mão esquerda chama-se canhoto.

Azt, aki hajlamosabb a bal kezét használni, balkezesnek hívjuk.

Someone who has the predisposition to use their left hand is called left-handed.

Group 10:

(15) Quando fomos à Madeira visitámos o jardim botânico do Monte, onde vimos milhares de plantas exóticas.

Amikor Madeirára látogattunk, megnéztük a montei füvészkertet, ahol egzotikus növények ezreivel találkoztunk.

When we went to Madeira, we visited the botanical garden of Monte where we saw thousands of exotic plants.

Group 11:

(16) Passei três horas ao computador à procura de uma receita fabulosa que pudesse fazer para o almoço com os pais do José.

Három órán át kutattam a számítógépen egy mesés recept után, amelyet majd elkészíthetek ebédre, amikor ott lesznek José szülei.

I spent three hours on the computer searching for a fabulous recipe that I could prepare for the lunch with José's parents.

Group 12:

(17) A Mariana ficou tão bêbada na tua festa de anos que perdeu o sapato direito.

Mariana annyira berúgott a szülinapi bulidon, hogy elveszítette a cipőjét a jobb lábáról.

Mariana got so drunk on your birthday party that she lost her right shoe.

Type 2 sentences:

Group 1:

(18) Com a tempestade de ontem uma árvore ficou partida.

A tegnapi vihar miatt kettéhasadt az egyik fa.

Because of yesterday's storm a tree got broken.

(19) O Francisco teve uma reacção alérgica porque o leite estava estragado.

Francisco allergiás rohamot kapott, mert a tej romlott volt.

Francisco had an allergic reaction because the milk was spoiled.

Group 2:

(20) Não passou no seu exame de fotografia porque a imagem ficou tremida.

Nem ment át a fényképészet vizsgáján, mert elmosódott a kép.

She/he did not pass her/his photography exam because the image got blurred.

Group 3:

(21) Depois de muitos anos de protestos, a liberdade foi garantida.

Sokévnyi tiltakozás után végre garantálták a szabadságot.

After many years of protests, liberty was guaranteed.

Group 4:

(22) A Cristina começou a gritar porque a dor era demasiada.

Cristina elkezdett kiabálni, mert a fájdalom túl erős volt.

Cristina started to scream because the pain was too much.

(23) O lugar foi escolhido após uma discussão intensa.

A helyszínt egy heves vita után választották ki.

The place was chosen after an intense discussion.

Group 5:

(24) O papel ficou molhado porque começou a chover quando estava na varanda a desenhar.

A papír elázott, mert elkezdett esni, amikor a verandán rajzolgatott.

The paper became wet because it started to rain when he/she was drawing on the veranda.

Group 6:

(25) Apanhei uma gripe, por isso o meu nariz está entupido.

Influenzás vagyok, úgyhogy eldugult az orrom.

I caught the flu, therefore my nose is blocked.

(26) Após tanto sofrimento, assinaram o tratado e a paz ficou estabelecida.

Megannyi szenvedés után végre aláírták a szerződést és a béke megszilárdult.

After so much suffering, the treaty was signed and peace was established.

Group 7:

(27) Varri todos os quartos até a pá ficar cheia.

Addig söprögettem a szobákban, amíg a lapát tele nem lett.

I swept all the rooms until the dustpan got full.

(28) O David esteve tanto tempo ao telefone que o chá ficou frio.

David annyi ideig telefonált, hogy kihűlt a tea.

David was on the phone for so long that his tea got cold.

Group 8:

(29) O Marco só se apercebeu que o problema era sério quando entrou na cozinha.

Marco csak akkor értette meg, hogy a probléma komoly volt, amikor belépett a konyhába.

Marco only realized that the problem was serious when he entered the kitchen.

Group 9:

(30) O limão está maduro demais.

A citrom túl érett.

The lemon is too ripe.

(31) A televisão ficou avariada, assim não vimos o fim do jogo.

Elromlott a televízió, így nem láttuk a meccs végét.

The television got broken, this way we didn't see the end of the game.

Group 10:

(32) Meninos, o pudim está feito, já podem vir lanchar.

Gyerekek, kész van a puding, jöhettek uzsonnázni!

Kids, the pudding is ready, you can come and eat!

Group 11:

(33) Por causa da chuva a janela está embaciada, assim não se vê nada.

Az eső miatt párás az ablak, ezért nem lehet látni semmit.

Because of the rain the window is foggy, therefore we cannot see anything.

Group 12:

(34) O vinho está aberto, só faltam os copos.

A bor már ki van bontva, csak a poharak hiányoznak.

The wine is open, the glasses are the only things missing.

Type 3 sentences

Group 1:

(35) Esta é a chave que o advogado achou pequena demais para a tua fechadura.

Ez az a kulcs, amelyről az ügyvéd azt hitte, hogy túl kicsi lesz a záradhoz.

This is the key that the lawyer thought was too small for your lock.

(36) Há ali um tomate que a menina comprou estragado por acidente.

Van ott egy paradicsom, amelyet a lány véletlenül romlottan vett.

There is a tomato that the girl bought rotten by accident.

Group 2:

(37) Quanto aos bascos, falou-se de uma origem que o aluno não achou clara.

Egy olyan eredetről volt szó a baszkokat illetően, amelyet a diák nem vélt egyértelműnek.

With regard to the Basques, they talked about an origin that the student did not consider clear.

Group 3:

(38) Esta é a sociedade que os políticos julgam desenvolvida, mas na realidade não é bem assim.

Ez az a társadalom, amelyet a politikusok fejlettnek vélnek, de a valóságban nem egészen így van.

This is the society that politicians consider developed, but in reality this is not exactly true.

Group 4:

(39) Neste parque há uma flor que o Vítor não acha linda.

Ebben a parkban van egy virág, amelyet Vítor nem tart szépnek.

There is a flower in this park that Victor does not consider beautiful.

(40) Finalmente vamos comer um jantar que a mãe não comprou preparado.

Végre olyan vacsorát fogunk enni, amelyet anya nem készen vett.

Finally, we are going to eat a dinner that mom did not buy prepared.

Group 5:

(41) Enfim, há aqui um jornal que a Catarina não achou chato.

Végre van itt egy újság, amelyet Catarina nem vélt unalmasnak.

At last, there's a newspaper here that Catarina did not consider boring.

Group 6:

(42) Amanhã vamos viajar para um país que a tia acha belíssimo.

Holnap elutazunk egy olyan országba, amelyet a nénikém gyönyörűnek tart.

Tomorrow we are going to travel to a country that my aunt considers gorgeous.

(43) No fim da rua vimos a cruz que o meu tio tinha encontrado partida.

Az utca végén láttunk azt a keresztet, amelyet a bácsikám törötten talált.

In the end of the street we saw the cross that my uncle had found broken.

Group 7:

(44) No Irão encontrámos uma fé que o professor julgou antiquíssima.

Találtunk Iránban egy olyan hitet, amit a professzor ősinek vélt.

We found a faith in Iran that our professor considered very ancient.

(45) Havia naquela sala um pó que a professora achou espesso.

A tanárnő szerint vastag por volt abban a teremben.

There was some dust in the room that the teacher considered thick.

Group 8:

(46) Hoje é um dia que a Luísa declarou péssimo, porque perdeu o seu bilhete de identidade.

A mai egy olyan nap, amelyet Luísa borzasztónak nyilvánított, mert elvesztette a személyi igazolványát.

Today is a day that Luísa declared horrible because she lost her identity card.

Group 9:

(47) Nesta aldeia celebra se uma tradição que o avô não acha adequada.

Ebben a faluban él egy olyan hagyomány, amelyet a nagypapa nem tart illendőnek.

In this village they celebrate a tradition that grandfather does not consider adequate.

(48) O Rui disse que encontraram o avião que as notícias anunciaram perdido.

Rui azt mondta, hogy megtalálták azt a repülőt, amelyet a hírekben még elveszettnek hittek.

Rui said that they found the airplane that had been announced lost in the news.

Group 10:

(49) Foi a minha irmã que comprou o amendoim que a Sílvia tornou triturado.

A nővérem vette azt a mogyorót, amelyet aztán Sílvia ledarált.

It was my sister who bought the peanut that Sílvia ground up.

Group 11:

(50) No ano passado organizámos uma festa que o meu patrão achou fantástica.

Tavaly szerveztünk egy olyan bulit, amelyet a főnököm fantasztikusnak vélt.

Last year we organized a party that my boss thought was fantastic.

Group 12:

(51) Não tinha muita esperança, mas finalmente consegui fazer um almoço que a minha namorada achou delicioso.

Nem igazán reménykedtem, de végül sikerült olyan ebédet készítenem, amely ízlett a barátnőmnek.

I didn't have much hope, but I finally managed to prepare a lunch that my girlfriend thought was delicious.

Distractors:

(52) A universidade não emite cartões de estudante durante este mês, portanto tens de esperar mais.

Az egyetem nem ad ki diákigazolványokat ebben a hónapban, úgyhogy még várnod kell.

The university does not issue student cards this month, therefore you have to wait more.

(53) Uma das raças de cães húngaros é muito famosa por ter o pelo com madeixas grossas.

Az egyik magyar kutyafajta arról híres, hogy a szőre vastag tincsekből áll.

One of the Hungarian dog breeds is very famous for having dreadlocks for fur.

(54) Há salmões que passam quase toda a vida em águas salgadas.

Vannak olyan lazacok, amelyek majdnem az egész életüket sós vizekben töltik.

There are salmons that spend their whole lives in saltwater.

(55) Já ouvi canções daquela época, mas não gostei muito.

Már hallottam dalokat abból az időszakból, de nem igazán tetszettek.

I have heard songs from that era but I didn't like them very much.

(56) Ainda existem falcões na Europa, mas estão em perigo de extinção.

Még élnek sólymok Európában, de a kihalás veszélye fenyegeti őket.

There still are falcons in Europe, however they are on the verge of extinction.

(57) Em Portugal cultivam-se melões de muitos tipos.

Portugáliába sokféle dinnyét termesztenek.

In Portugal, many kinds of melons are grown.

(58) Na exposição do vestuário do século XVII também vimos botões.

A XVII. századi öltözékekről szóló kiállításon láttunk gombokat is.

We also saw buttons in the exposition about 17th century clothing.

(59) Comprámos um quilo de camarões para a tua festa.

Vettünk egy kiló garnélarákot a bulidra.

We bought a kilo of prawns for your party.

(60) Tenho dúzias de notificações na minha página web.

Több tucat értesítés érkezett a weboldalamra.

I have dozens of notifications on my web page.

(61) Fomos a uma loja de colchões e comprámos algumas almofadas anatómicas.

Elmentünk egy matracokat árusító üzletbe és vettünk pár anatómiai párnát.

We went to a mattress store and bought some anatomic pillows.

(62) Depois de ter bebido uma garrafa inteira de vinho, o Paulo não está em condições de guiar.

Miután megivott egy egész üveg bort, Paulo nincs olyan állapotban, hogy vezetni tudjon.

After having drunken a whole bottle of wine, Paulo is in no condition to drive.

(63) A Amália passou toda a manhã em reuniões.

Amália egész délelőtt megbeszéléseken volt.

Amália spent the whole morning at meetings.

(64) É importante limpar as feridas muito bem para não contrair infeções.

Nagyon fontos, hogy jól kitisztítsuk a sebeket, nehogy elfertőződjenek.

It’s important to clean wounds very well so that they don’t get infected.

(65) Aparecem dragões em todos os contos de fadas.

Az összes tündérmesében találkozhatunk sárkányokkal.

Dragons appear in all fairy tales.

(66) O João dá explicações de física a estudantes universitários todas as sextas feiras.

João minden pénteken különórákat ad fizikából egyetemi diákoknak.

João gives extra physics classes for university students every Friday.

(67) Os Estados Unidos de América têm o maior número de furacões por ano.

Az Amerikai Egyesült Államokban pusztít évente a legtöbb hurrikán.

The USA has the highest number of hurricanes per year.

(68) Entrámos na pastelaria para pedir direções.

Bementünk egy cukrászdába, hogy útbaigazítást kérjünk.

We entered a pastry shop to ask for directions.

(69) As pessoas que sofrem de stress frequentemente desenvolvem perturbações do sono.

Azoknak, akik stressztől szenvednek, gyakran alvási zavaraik lesznek.

Those people who suffer from stress often develop sleep disorders.

(70) A vítima foi levada diretamente para a sala de operações depois do acidente.

Az áldozatot egyenesen a műtőbe szállították a baleset után.

The victim was taken directly to the operating theatre after the accident

Appendix 4: Snapshot of the online cloze test (learnclick.com)