



A MORFOLOGIA DOS DESVIOS DE GÊNERO GRAMATICAL EM PLNM

THE MORPHOLOGY OF GRAMMATICAL GENDER DEVIATIONS IN PNNL

Tânia Santos Ferreira¹

RESUMO

Reconhecendo o gênero gramatical como uma área crítica na aquisição tardia do português como língua não materna (PLNM) (FERREIRA, 2011; 2019; MARTINS, 2015; PINTO, 2017), neste trabalho analisam-se desvios morfológicos detectados em produções escritas por falantes nativos de espanhol, alemão e inglês, a frequentar turmas em diferentes níveis de proficiência linguística. Mais especificamente, pretende-se apurar, a partir da análise de formas morfológicas desviantes que resultam da tentativa do aprendente em estabelecer a concordância nominal em gênero, padrões de aquisição desta estrutura que se correlacionem quer com fatores intralinguísticos, como a configuração do sistema linguístico (v.g. morfossemântico) nativo do aprendente e as características de atribuição de valores de gênero do português (VILLALVA, 1994; MOTA, 2020), quer com fatores extralinguísticos, como o nível de aprendizagem. Da análise empírica conclui-se que, além da transferência linguística, uma parte substancial dos desvios resulta da sobregeneralização dos padrões de atribuição de valores de gênero disponíveis em português (**os tempos livros*, **as carpetas*). Portanto, estes dados são reveladores de um certo grau de sensibilidade dos diferentes grupos de aprendentes quanto aos indícios formais de atribuição de gênero.

Palavras-chave: Sistema de classificação nominal; Gênero gramatical; Aquisição tardia de português como língua não materna; Transferência linguística.

ABSTRACT

Recognizing grammatical gender as a critical area in the context of late acquisition of the Portuguese as a non-native language (PNNL) (FERREIRA, 2011; 2019; MARTINS, 2015; PINTO, 2017), this work analyzes morphological deviations detected in productions written by native speakers of Spanish, German and English who attend classes of different levels of Portuguese learning. More specifically, it seeks to ascertain from the analysis of deviant morphological forms that result from the learner's attempt to establish nominal gender agreement, patterns of acquisition of this structure that could correlate either with intralinguistic factors, such as the configuration of the learner's native linguistic system (e.g. morphosemantic) and the characteristics of attribution of gender values of Portuguese (VILLALVA, 1993; MOTA, 2020), either with extralinguistic factors, such as the level of learning. From the empirical analysis it is concluded that, in addition to the linguistic transfer, a substantial part of the deviations results from the overgeneralization of the gender assignment patterns available in Portuguese (**os tempos livros*; **as carpetas*). Therefore, these data reveal a certain degree of sensitivity of different groups of learners to formal evidence of gender assignment.

Keywords: Nominal classification system; Grammatical gender; Late acquisition of Portuguese as non-native language.

¹ Professora Auxiliar Convidada, Faculdade de Letras da Universidade de Coimbra, CELGA-ILTEC, tania.ferreira@fl.uc.pt

Introdução

O presente artigo centra-se na aquisição tardia da categoria de género gramatical por aprendentes de português como língua não materna (PLNM), através da análise de formas morfológicas desviantes que resultam da tentativa do aprendente tardio em estabelecer, no sintagma nominal, a concordância em género. Os dados deste estudo foram selecionados de um conjunto de produções escritas por aprendentes adultos, falantes nativos de espanhol, alemão e inglês, a frequentar turmas de diferentes níveis de aprendizagem. Com base na análise empírica, pretende-se apurar padrões variáveis da aquisição tardia da categoria de género gramatical em PLNM que se possam correlacionar, por um lado, com fatores de natureza intralinguística, como a configuração do sistema linguístico nativo do aprendente, e em particular do sistema morfossemântico, e as características subjacentes à atribuição e marcação de valores de género nominal em português, a língua-alvo de aprendizagem, e, por outro, com fatores de natureza extralinguística, nomeadamente o respetivo nível de proficiência e de competência linguísticas do aprendente em português.

Estudos sobre a aquisição tardia de uma LNM são reveladores de que a atribuição de valores de género aos nomes corresponde a uma área crítica. Com efeito, a investigação produzida no âmbito da aquisição linguística tardia do PLNM, seja como língua segunda ou língua estrangeira (FERREIRA, 2011; 2019; MARIOTTO & LOURENÇO GOMES, 2013), seja como língua terceira ou língua adicional (LACSÁN, 2015; PINTO, 2017) demonstra que, embora se registre uma melhoria global no desempenho em função do nível de proficiência linguística dos aprendentes, os desvios de atribuição e de concordância nominal em género nunca chegam a ser totalmente erradicados das produções dos aprendentes.

A maioria destes trabalhos baseia-se, sobretudo, na análise dos desvios de atribuição de valores de género detetados em produções escritas. Nestes estudos, são colocadas diversas hipóteses para justificar a persistência dos comportamentos desviantes ao longo do desenvolvimento interlinguístico dos aprendentes, tendo-se em conta não só os possíveis constrangimentos associados ao perfil do aprendente, nomeadamente a idade e período de exposição à língua-alvo, o nível de proficiência e de competência linguísticas, a LM ou outras línguas previamente conhecidas, mas também as condicionantes associadas às características formais de atribuição e de concordância nominal em género do português.

Com o intuito de compreender que estratégias mobilizam diferentes grupos de aprendentes tardios no momento de atribuição de valores de género aos nomes e qual o respetivo grau de sensibilidade relativamente às propriedades formais de atribuição e de concordância nominal em género do português, ao longo das diferentes fases do desenvolvimento da aprendizagem, neste estudo analisa-se um conjunto de desvios que incidem sobre a forma morfológica dos constituintes internos do sintagma nominal, tais como especificadores (*o *minho pais*); modificadores (**aproveito os tempos livros*) e nomes (**os novos colegas*). Parte-se, assim, da assunção de que a análise destas formas morfológicas não convergentes permitirá aferir as particularidades do processo de aquisição tardia da categoria gramatical de género.

Refira-se, neste ponto, a estrutura deste artigo. A primeira parte é dedicada ao enquadramento teórico que sustenta este trabalho, analisando-se, em primeiro lugar, as questões relativas ao conceito de gênero gramatical e à descrição do sistema de atribuição e de concordância de gênero nominal em português. Posteriormente, consideram-se os contributos teóricos relativos à assimilação das propriedades morfossintáticas de um idioma não nativo, com especial enfoque para o papel que, neste contexto, é atribuído ao conhecimento linguístico prévio proveniente da gramática da LM ao longo do desenvolvimento interlinguístico dos aprendentes. Referidos os pressupostos teóricos, passar-se-á ao trabalho empírico, com a descrição, em primeiro lugar, do *corpus* bem como das variáveis de análise. Em seguida, procede-se ao tratamento e discussão dos dados apurados e, por fim, as considerações finais.

A categoria gramatical de gênero

É possível encontrar na bibliografia especializada diferentes contributos teóricos para a delimitação do conceito de gênero gramatical (IBRAHIM, 1973, CORBETT, 1991, RITTER, 1993, AIKHENVALD, 2000). Além de não ser uma categoria gramatical universal, já que não se encontra em todos os idiomas do mundo, corresponde a uma propriedade verdadeiramente particular, sobretudo se se tiver em conta (i) os critérios que cada língua seleciona e que servem de base à delimitação concetual desta categoria; (ii) a aplicação, em cada idioma, de critérios (semânticos e formais) de atribuição de valores de gênero nominal; e (iii) a manifestação formal dos valores de gênero.

Segundo Corbett (1991), o gênero corresponde a uma propriedade inerente aos nomes, não o sendo em outras classes de palavras como, por exemplo, determinantes, adjetivos, verbos, etc.² Para este investigador, a atribuição de um determinado valor de gênero ao item nominal pode estar dependente quer do significado deste quer da sua forma, sendo que tais propriedades (semânticas ou formais) só serão consideradas como instanciações de uma categoria de gênero numa língua se nela espoletarem o fenómeno de concordância sintática (CORBETT, 1991). Portanto, e de acordo com esta perspetiva, considera-se que numa língua há tantos valores de gênero quantas as possibilidades de concordância sintática desencadeadas pelo item nominal. A este respeito, Mota afirma:

[d]ado que nos nomes, o caso mais típico é não haver identificação entre o valor de gênero e o padrão temático, em última análise é através das palavras que não têm gênero inerente (os adjetivos e as palavras gramaticais pré-nominais), mas que entram na relação de concordância sintática com os nomes, que o valor de gênero destes últimos é evidenciado (MOTA, 2020, p. 2851).

O gênero surge tipicamente associado a sistemas de classificação nominal característicos de línguas de matriz indo-europeia e de línguas semíticas que possuem um número restrito

2 A este respeito, Mota (2020, p. 2850) refere que “[n]os nomes, o gênero é uma categoria lexical típica; está inerentemente marcada nos lexemas nominais com valor invariável, maioritariamente independente da forma do lexema (...) outros elementos não nominais não têm um valor especificado na sua entrada lexical (...)”.

de valores, essencialmente dois ou três (masculino, feminino e neutro), e segundo os quais os nomes se organizam. Nestes sistemas, todos os nomes possuem um valor de género intrínseco, apesar de este nem sempre estar explicitamente marcado na sua forma morfológica. Além disso, e apesar de se verificar, nos nomes que denominam entidades sexuadas, uma estreita relação entre o género natural do referente (macho/fêmea) e o género gramatical (masculino/feminino), a associação de valores de género aos nomes é, em grande parte dos casos, idiossincrática (CORBETT, 1991; AIKHENVALD, 2000; GRINEVALD, 2000).

A arbitrariedade e assistemática da associação dos valores de género aos nomes é também visível pelo facto de haver, em línguas da mesma família linguística, formas nominais com o mesmo referente, mas com valores de género distintos. Por exemplo, em português o nome *árvore* é feminino, não o sendo em espanhol (*árbol*), em italiano (*albero*) e em francês (*arbre*). Por sua vez, o item *leite* é masculino na língua portuguesa, sendo feminino em espanhol (*leche*), italiano (*latte*) e francês (*lait*) (VILLALVA, 1994, p. 236-237).

O sistema de classificação nominal do português

Em português, todos os nomes e alguns pronomes possuem um valor de género gramatical, havendo dois valores em oposição: masculino e feminino. Além do nome, a categoria de género afeta outras classes de palavras (lexicais e gramaticais) pela necessidade de com ele estabelecerem relações de concordância sintática³ (MOTA, 2016, 2020). Deste modo, são marcadas quanto ao valor de género, as palavras concordantes localizadas à esquerda e à direita do nome, o núcleo do sintagma, ou seja, especificadores (determinantes e quantificadores) e modificadores (adjetivos).

Assim sendo, no interior do sintagma nominal, a concordância estabelece-se entre o núcleo nominal e as expressões de determinação e de quantificação que o antecedem (especificadores); e (ii) o núcleo nominal e os adjetivos integrados em sintagmas adjetivais que funcionam como seus modificadores ou elementos apositivos (BRITO, 2003, p. 328).

Atendendo à estrutura formal dos itens nominais, tipicamente os valores de género masculino e feminino são associados às respetivas vogais átonas *-o* e *-a* que surgem à direita do radical nominal, simples ou derivado. Porém, não é possível considerar que estas vogais correspondem a morfemas de género, uma vez que encontramos nomes masculinos terminados em *-a* (*o dia*) e femininos terminados em *-o* (*a tribo*)⁴ (RIO-TORTO & RODRIGUES, 2016;

3 De acordo com Mota (2020, p.2816), nos nomes o género corresponde a uma categoria lexical, por se encontrar inerentemente especificado na matriz de traços da entrada lexical nominal, mas nos restantes elementos que com ele concordam, o género corresponde a uma categoria morfossintática, “de concordância”, uma vez que tais elementos possuem um valor de género especificado na sua entrada lexical, adquirindo um dos dois valores possíveis, masculino ou feminino, pela concordância sintática.

4 Ao analisar o funcionamento da categoria gramatical de género do português, e em particular a relação entre marcadores morfológicos do nome e valores de género, Mota (2020) constata que o índice temático em português corresponde a um sufixo *portmanteau*, já que, além de classificar a classe formal do nome, pode, em alguns casos, acumular a informação de valor de género (cf. MOTA 2020, p. 2850). Porém, a mesma autora salienta que “a possibilidade de haver acumulação de informação de género no índice temático não equivale a que os nomes flexionem em género” (MOTA, 2020, p.2901).

MOTA, 2020). Além disso, há nomes que possuem diferentes terminações que tanto podem ser de gênero masculino ou feminino (*a pá, o pé, o coração, a mão, o elefante, a catedral*). Consequentemente, a principal função das vogais *-o* e *-a* não é a de marcar o valor de gênero, mas a de classificar os nomes, integrando-os em diferentes classes temáticas e designam-se índices temáticos (VILLALVA, 2003; MOTA, 2020).

No trabalho de investigação levado a cabo por Ferreira (2011), a autora procurou determinar a distribuição relativa dos nomes de gênero masculino e feminino, em função da classe temática a que pertencem. Deste modo, procedeu à análise dos índices de frequência dos lemas recenseados no *Corpus Léxico Fundamental do Português Contemporâneo* (CORLEX), contabilizando as frequências de nomes de gênero masculino e feminino, formas simples e derivadas, com referentes sexuais e não sexuais, distinguindo os que possuem índices temáticos *-o* e *-a* dos que pertencem a outras classes temáticas (FERREIRA, 2011, p.33-34).

Deste modo, constatou que a percentagem da frequência de nomes masculinos com índice temático *-o* é de, aproximadamente, 32%, havendo apenas cerca de 1,6% de frequências de formas masculinas com índice temático *-a*. Quanto aos nomes de gênero feminino, constatou que a percentagem da frequência de nomes com índice temático *-a* atinge quase os 27%, um valor bastante superior ao registado em formas femininas de tema em *-o*: cerca de 0,04%.

Por conseguinte, estes dados são reveladores de que em cerca de 60% das frequências de ocorrências nominais do português se verifica a correlação parcial entre índice temático (*-o* e *-a*) e valor de gênero. Nos restantes casos, não há qualquer indício de natureza morfológica que permita justificar a associação de valores de gênero aos nomes. Portanto, estes dados são reveladores da fraca robustez que existe na associação de índices temáticos e valores de gênero nominal em português, corroborando ainda o que afirmam Choupina *et al.*:

classe e gênero não são traços do mesmo tipo: os tipos de classe permitem inserir um Item de Vocabulário numa classe de nomes em que todos os elementos partilham o mesmo funcionamento morfológico; os traços de gênero promovem a organização dos nomes em grupos ou classes que determinam diferentes padrões de concordância (CHOUPINA *et al.*, 2016, p. 210)

Além dos índices temáticos, existem outros constituintes morfológicos que condicionam o valor de gênero nominal. Com efeito, existem alguns sufixos derivacionais que, além de determinarem a categoria sintática do produto final que originam, possuem informação de gênero quando criam produtos nominais (VILLALVA, 2003; RIO-TORTO & RODRIGUES, 2016, p.140-150; CHOUPINA, BAPTISTA & COSTA, 2014) e alteram, em alguns casos, o gênero da base nominal a que se associam. Por exemplo, todos os nomes derivados em *-ção* são femininos (*a preparação, a constipação*), sendo masculinas todas as formas nominais derivadas em *-ume* (*o vasilhame, o queixume*).

No que aos critérios de atribuição de valores de gênero nominal diz respeito, verifica-se em português a atuação de critérios semânticos e formais. Em grande parte dos itens com o traço semântico [+sexuado], o valor de gênero gramatical corresponde ao gênero natural do referente.

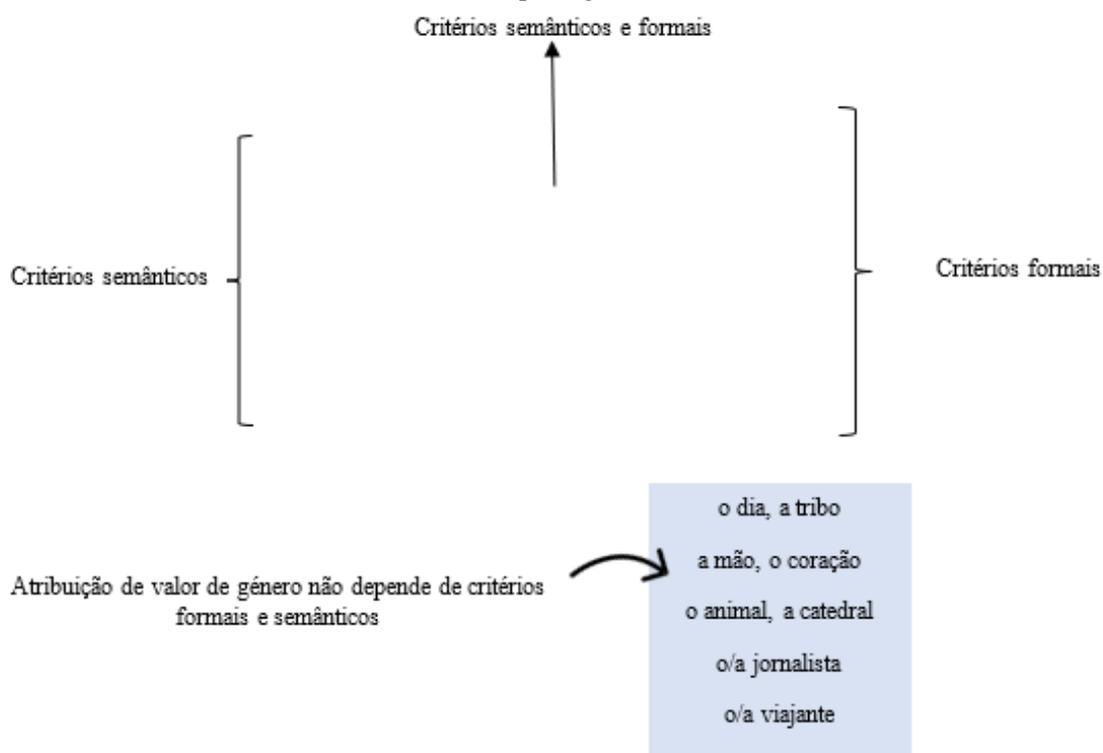
Assim, encontram-se formas nominais masculinas que designam seres do sexo masculino (*o homem, o rapaz, o menino*) e formas nominais femininas que designam seres do sexo feminino (*a mulher, a rapariga, a menina*). Nestes casos, a atribuição de um valor de género caracteriza-se pela aplicação do critério semântico que pode, em alguns casos, ser coincidente com o critério formal. Por exemplo, a forma nominal *menino* refere-se a uma entidade masculina e possui o índice temático *-o* que, tipicamente, está associado ao género masculino.

Ainda que se verifique a atuação do critério semântico nos nomes de entidades sexuadas, registam-se, no entanto, algumas restrições. Na verdade, encontram-se, em português, nomes que, apesar de terem referentes sexuados, possuem um valor de género único e independente do sexo do referente, como *a testemunha*.

Já em alguns itens com o traço semântico [-sexuado], verifica-se uma correlação parcial entre marcadores morfológicos, como índices temáticos *-o* e *-a*, e respetivos valores de género masculino e feminino, havendo ainda alguns sufixos derivacionais que determinam o valor de género do produto nominal.

Com base nos indícios de atribuição de valores de género aos nomes em português, no Esquema 1 apresenta-se uma proposta de representação da distribuição dos nomes, atendendo à aplicação de critérios semânticos e formais.

Esquema 1 – Aplicação de critérios semânticos e formais de atribuição de valores de género nominal em português



Portanto, na perspectiva de quem aprende o português como LNM, as características acima descritas apontam, por um lado, para a pouca robustez dos dados disponíveis no *input*

no que ao nome diz respeito, quer numa perspectiva formal, quer numa perspectiva semântica. Por outro lado, parece evidente que a concordância sintática se constitui como um fator importante para determinar o valor de gênero do item nominal. Por conseguinte, sendo a concordância um mecanismo reconhecidamente problemático na aquisição tardia de uma LNM (FRANCESCHINA, 2005; LEIRIA, 2006; GODINHO, 2010; MARTINS, 2015) compreende-se que a atribuição de valores de gênero nominal por parte do aprendente tardio implica não só que ele reconheça o valor de gênero intrínseco do nome, núcleo do sintagma nominal, como também seja capaz de adequar a forma morfológica dos demais elementos concordantes.

Os conhecimentos linguísticos prévios na aquisição tardia do gênero nominal

Dados sobre a assimilação da categoria de gênero gramatical de um idioma não nativo são reveladores de que, e independentemente da língua-alvo em aquisição, quanto mais tardia for a exposição do aprendente, maiores dificuldades se registam na assimilação do sistema de atribuição e de concordância nominal em gênero. Diferentes fatores têm sido equacionados para justificar a relativa “resistência” da categoria de gênero gramatical ao processo de aquisição linguística tardia. Importa, neste contexto, destacar o papel que, na literatura, tem sido atribuído à configuração do conhecimento linguístico prévio do aprendente para o processo de assimilação tardia dos valores de gênero e de concordância nominal.

Selinker (1992 [1972]) define a transferência linguística como um dos distintos processos que atuam na construção das interlínguas dos aprendentes tardios. Genericamente, a interlíngua corresponde ao sistema linguístico intermédio que o aprendente tardio vai construindo, ao longo da aprendizagem. Do ponto de vista estrutural, as interlínguas não correspondem nem às estruturas da língua de origem, ou seja, do idioma nativo do aprendente, nem às estruturas da língua-alvo e são consideradas sistemas linguísticos dado que, durante a sua construção, os fenómenos desviantes são, de certo modo, previsíveis. Com efeito, e à semelhança do que é postulado para a aquisição linguística nativa, os desvios sistemáticos evidenciam um processo gradual e transicional, podendo ser o reflexo do que Corder (1992 [1967]) designa como o ‘programa interno’ (*built-in-syllabus*) do aprendente.

No que à assimilação de valores de gênero diz respeito, vários estudos procuram compreender em que medida a configuração do conhecimento linguístico dos informantes atua ao longo das diferentes fases de assimilação da categoria gramatical de gênero (WHITE *et al.*, 2004; FRANCESCHINA, 2005; SABOURIN *et al.*, 2006; MARTINS, 2015; PINTO, 2017; FERREIRA, 2019). A investigação produzida demonstra que a influência dos conhecimentos linguísticos prévios pode afetar, direta e indiretamente, o percurso da aprendizagem, contribuindo para a formulação de hipóteses sobre o *input* e levando o aprendente à adoção de diferentes estratégias ao longo do desenvolvimento interlinguístico. Portanto, nestes trabalhos o principal objetivo é compreender em que medida a presença na LM de sistemas de atribuição

e de concordância de gênero nominal com características muito próximas das da língua-alvo de aprendizagem afeta o desempenho linguístico dos aprendentes e se, em última instância, este é um fator determinante para a aquisição plena desta estrutura.

Ao analisar as implicações da transferência linguística para a assimilação do sistema de atribuição de gênero nominal do holandês por aprendentes tardios, falantes nativos de alemão, de inglês e de línguas românicas, SABOURIN *et al.* (2006) verificaram que, pese embora se registre uma melhoria global no desempenho linguístico em níveis de proficiência mais avançados nos diferentes grupos de aprendentes analisados, os falantes cuja LM não possui um sistema de atribuição de gênero gramatical apresentam maiores dificuldades, quer em tarefas de associação do valor de gênero ao item lexical, quer em tarefas de estabelecimento de relações de concordância em gênero. Neste estudo, distinguem-se, com base no tratamento dos dados, dois tipos distintos de transferência linguística: a transferência de superfície (*surface transfer*) e a transferência profunda (*deep transfer*).

A transferência de superfície corresponde a “direct transfer of morphologically similar gender realization between L1 and L2” (SABOURIN *et al.*, 2006, p. 10) e, portanto, está intimamente correlacionada com o facto de na LM do aprendente haver um sistema de atribuição de gênero nominal com características muito próximas do da língua-alvo de aprendizagem, nomeadamente no que concerne às propriedades formais subjacentes à atribuição dos valores de gênero nominal. Trata-se, por conseguinte, da possibilidade de os aprendentes utilizarem, nas suas produções, a representação morfológica dos itens e dos respetivos valores de gênero que são, assim, diretamente importados da sua LM. Segundo os investigadores, a presença de cognatos entre as línguas potencia este tipo de transferência linguística e, dado o grau de semelhança entre as estruturas, auxilia o aprendente tardio na tarefa de assimilação dos valores de gênero nominal.

Por sua vez, a transferência profunda diz respeito à transferência de estruturas sintáticas mais abstratas e manifesta-se quando não existe uma correspondência direta relativamente às marcas morfológicas associadas à realização dessas mesmas categorias gramaticais na LM e na língua-alvo (SABOURIN *et al.*, 2006, p. 3). Ou seja, a transferência profunda está implicada nas situações em que o falante aprendente possui na sua LM a categoria de gênero gramatical, mas com um sistema de atribuição de gênero nominal com propriedades formais de associação dos valores distintas do da língua em aprendizagem.

Os autores verificaram que este tipo de transferência também é relevante para a assimilação da categoria de gênero, já que o segmento de falantes nativos de línguas românicas, pese embora a impossibilidade de transferirem as marcas morfológicas de atribuição dos valores de gênero da sua LM para a língua-alvo, apresentam um desempenho globalmente superior ao que é registado pelo grupo de falantes cuja LM não possui um sistema de atribuição de gênero nominal, como é o caso dos falantes nativos de inglês. Estes últimos, por não terem representada

na sua língua a categoria de gênero gramatical, apresentam piores resultados relativamente aos restantes grupos, quer em tarefas de associação de valor de gênero aos itens lexicais, quer em tarefas de concordância nominal em gênero (SABOURIN *et al.*, 2006, p. 23).

Por fim, em função dos resultados obtidos, os autores concluem que:

the relatively good performance of the German group compared to the Romance group suggests that while transfer of abstract features plays a role in the L2 acquisition of gender, direct surface transfer of a congruent system is even more helpful, at least for acquiring gender assignment. Surface transfer, thus, probably accounted for some part of the advantage (SABOURIN *et al.*, 2006, p. 24).

Porém, e como os próprios autores sublinham, não é possível assumir uma posição clara relativa ao papel desempenhado pela transferência (profunda ou de superfície) para a aquisição tardia da categoria de gênero gramatical, já que a assimilação desta estrutura pode depender da atuação de outros fatores que não estão intrinsecamente relacionados com a configuração do conhecimento linguístico prévio (cf. FRANCESCHINA, 2005; SABOURIN *et al.*, 2006).

Alguns trabalhos relativos à aquisição tardia de PLNM (LEIRIA, 2006; MARTINS, 2015; FERREIRA, 2019) revelam que o facto de a LM do aprendente exibir um sistema de gênero não constitui *por si só* um fator facilitador no processo de assimilação dos valores de gênero nominal em português. A este respeito, Martins (2015) observa que os falantes nativos de chinês, idioma sem categoria gramatical de gênero, registam um melhor desempenho global na atribuição dos valores de gênero nominal face ao registado nos grupos de aprendentes cuja LM detêm um sistema de classificação nominal baseado em valores de gênero. Para a investigadora, a partir dos seus dados infere-se, assim, um efeito inibidor da transferência “não tanto direta de valores de gênero da LM para a LA em vocábulos semanticamente equivalentes (...), mas mais propriamente sob a forma de uma atitude de insegurança por parte do aprendente nestas circunstâncias” (MARTINS, 2015, p. 44).

Dados empíricos

A base empírica deste estudo compreende um conjunto de textos selecionados do *corpus* de Ferreira (2019) que, por sua vez, integra produções escritas por aprendentes tardios de PLNM com desvios de atribuição e de concordância nominal em gênero. Todas as produções deste *corpus* foram escritas por informantes adultos que, no momento da recolha, se encontravam a frequentar cursos de português europeu como língua estrangeira em instituições de ensino superior (cf. FERREIRA, 2019, p. 157-161).

Para este trabalho em particular, e tendo em conta os objetivos estipulados, selecionou-se um conjunto de textos escritos por falantes nativos de espanhol, alemão e inglês, distribuídos em

função do respetivo nível de proficiência linguística da turma frequentada pelos informantes. O critério que determinou a seleção deste subconjunto de informantes prende-se, essencialmente, com o número de valores de género que os respetivos idiomas nativos apresentam, bem como com a pertinência da própria categoria na língua em questão. Deste modo, selecionaram-se produções de falantes nativos de:

- I. espanhol**, idioma românico com um sistema de classificação nominal baseado em género bipartido, i.e., com dois valores de género em oposição (masculino e feminino) e com um sistema de marcação morfossintática e de concordância nominal muito semelhante ao da língua portuguesa;
- II. alemão**, uma língua com três valores de género (masculino, feminino e neutro) e com um sistema de concordância complexo, já que além do género e do número, os nomes e restantes elementos relacionados exibem a concordância em caso (nominativo, dativo, acusativo e genitivo); e
- III. inglês**, idioma sem um sistema de atribuição de género, conservando apenas vestígios de um primitivo sistema de género entre as formas pronominais, havendo ainda alguns itens nominais com formas lexicais distintas que refletem uma oposição de género natural das entidades designada, como *boy* ‘rapaz’ e *girl* ‘rapariga’.

Além da LM, outra variável considerada para o tratamento dos dados foi o nível de competência e de proficiência linguísticas da turma frequentada pelos aprendentes no momento da recolha dos textos. As produções encontram-se, assim, organizadas em três grupos distintos, devidamente identificados com os descritores propostos no âmbito do *Quadro Europeu Comum de Referência para as Línguas* (QECRL) (CONSELHO DA EUROPA, 2001). Deste modo, o *corpus* analisado contempla produções de aprendentes a frequentar turmas do nível de iniciação (A1 A2), intermédio (B1-B2) e avançado (C1). Esta distinção é, a nosso ver, essencial para a análise dos dados empíricos, uma vez que nos permitirá identificar as diferentes fases do desenvolvimento interlinguístico dos diferentes segmentos da amostra empírica em estudo.

Tendo em conta os objetivos da presente investigação, procedeu-se à análise dos textos com vista à identificação dos constituintes sintáticos (determinantes, quantificadores, nomes e adjetivos) que apresentavam formas morfológicas não convergentes com as correspondentes do português e que resultam da tentativa do aprendente em estabelecer, no sintagma nominal, a concordância em género (cf. Quadro 1).

Quadro 1 – Formas morfológicas desviantes distribuídas em função do constituinte afetado

Constituinte afetado		Exemplos
Especificadores	Determinantes	«Durante esso tempo» (Alemão.B1-B2) « las subidas» (Espanhol.A1-A2) «Todos os minhos amigos também» (Alemão.C1)
	Quantificador	«lembrar cado dia» (Alemão.B1-B2) «nos anos novento » (Inglês.B1-B2)
Adjetivos		«aproveito os tempos livros » (Alemão.B1-B2) «personalidas célebras » (Alemão.A1-A2) «ao dia seguinto » (Alemão.A1-A2)
Nomes		«num tentativo » (Inglês.C1) «Esta noita » (Alemão.A1-A2) «tomas umas copas » (Espanhol.A1-A2)

Atendendo à LM e respetivo nível de proficiência da turma frequentada pelos informantes, analisou-se um acervo total de 368 textos e apurou-se um total de 168 formas morfológicas desviantes. Nos textos selecionados apurou-se ainda um total de 12.955 ocorrências de sintagmas nominais, constituídos por um núcleo, o nome, acompanhado de especificadores (determinantes e quantificadores) e/ou modificadores (adjetivos). O número total de ocorrências de sintagmas nominais servirá de referência para o cálculo das percentagens relativas de desvios, uma vez que os casos assinalados incidem, precisamente, nas formas dos respetivos constituintes sintáticos que integram o sintagma nominal.

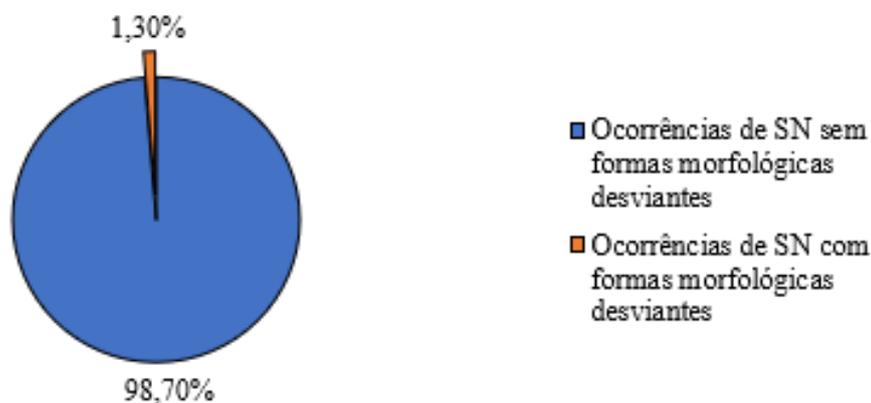
A Tabela 1 contém a distribuição dos informantes por LM e nível QECRL da turma frequentada, bem como o número (#) de sintagmas nominais (SN) apurado por segmento da amostra e respetivo número de formas morfológicas desviantes (FMD).

Tabela 1 – Caracterização da base empírica do estudo

Informantes			Textos						
LM	Nível QECRL	#	#	# de SN	#FMD				
Espanhol	A1-A2	42	76	46	84	1280	2811	64	74
	B1-B2	24		26		959		8	
	C1	10		12		572		2	
Alemão	A1-A2	70	134	75	155	1870	5321	21	48
	B1-B2	48		60		2483		21	
	C1	16		20		968		6	
Inglês	A1-A2	38	100	46	129	940	4823	19	46
	B1-B2	42		54		2756		20	
	C1	20		29		1127		7	
Σ		310		368		12 955		168	

Em função do número total de sintagmas nominais produzido, verifica-se que, em valores percentuais, as formas morfológicas desviantes afetam 1,30% de todas as ocorrências de sintagmas nominais no *corpus* selecionado (cf. Gráfico 1).

Gráfico 1 – Distribuição de ocorrências de sintagmas nominais (SN) sem formas morfológicas desviantes por oposição às ocorrências de SN com formas morfológicas desviantes

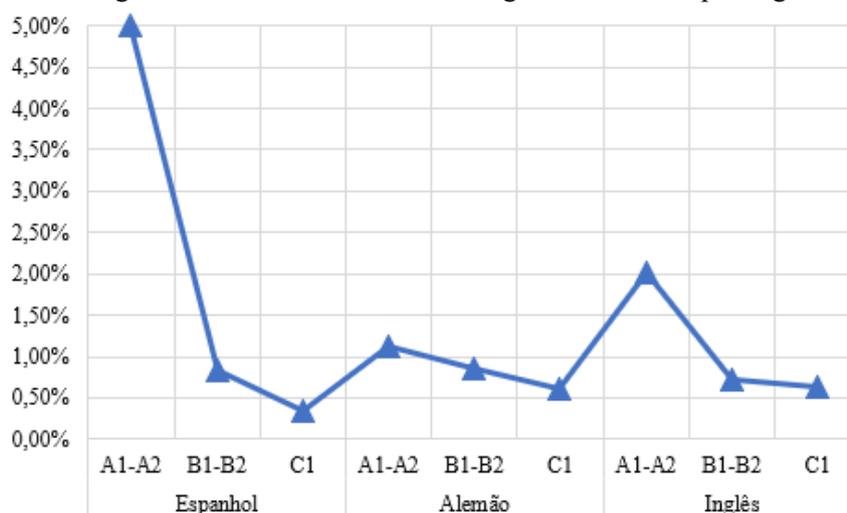


Portanto, estes resultados globais revelam que uma parte substancial dos informantes, no momento da atribuição de valores de gênero nominal, seleciona adequadamente a forma morfológica dos constituintes que compõem o sintagma nominal.

Resultados e discussão

Tendo como valor de referência o número absoluto de ocorrências de sintagmas nominais produzido por segmento da amostra empírica deste estudo (cf. Tabela 1), procedeu-se, num primeiro momento, ao cálculo das percentagens relativas de desvios morfológicos. Os resultados encontram-se cartografados no Gráfico 2.

Gráfico 2 – Percentagens relativas de formas morfológicas desviantes por segmentos da amostra



Mediante a leitura do Gráfico 2, constata-se que em todos os segmentos de informantes por LM se regista o mesmo padrão. Ou seja, são os níveis iniciais (A1-A2) que apresentam uma

maior proporção de desvios morfológicos face aos níveis mais avançados, sendo que no nível C1 a proporção de desvios é a mais baixa em todos os grupos de informantes. Portanto, infere-se que a assimilação desta estrutura parece ser favorecida com a progressão do conhecimento da língua.

Em função do desempenho linguístico observado por grupo de LM, o segmento da amostra de falantes nativos de espanhol regista a maior proporção de formas morfológicas desviantes que, nos níveis A1-A2, afetam cerca de 5% do total de ocorrências de sintagmas nominais produzido neste segmento da amostra. Nos demais grupos, a proporção de desvios morfológicos é inferior. Deste modo, a distribuição quantitativa dos desvios parece indicar que uma maior proximidade estrutural dos sistemas linguísticos, LM do aprendente e língua-alvo, parece conduzir, sobretudo numa fase inicial da aprendizagem, a uma maior proporção de ocorrências formas morfológicas desviantes sendo, por isso, necessário averiguar, nos dados apurados, que fatores estão na base dos desvios.

Para tal, procedeu-se a uma análise qualitativa dos dados apurados e partindo dessa análise, constatou-se que uma parte substancial dos desvios resulta da transferência linguística, sobretudo nos casos assinalados nos textos de aprendentes cuja LM é o espanhol:

1. *la familia* (Espanhol.A1-A2)
2. *la situação* (Espanhol.A1-A2)
3. *estos quartos* (Espanhol.A1-A2)
4. *ir al fútbol* (Espanhol.A1-A2)
5. *a la rua* (Espanhol.B1-B2)
6. *um equipo feminino* (Espanhol.A1-A2)

Com efeito, regista-se neste subconjunto de aprendentes que o facto de existirem cognatos entre a LM e a língua-alvo de aprendizagem promove a utilização, nas suas produções, da representação morfológica dos itens e respetivos valores de género que são, assim, diretamente importados do idioma nativo. Por conseguinte, poder-se-á admitir que as ocorrências de formas morfológicas desviantes se devem a um fenómeno de transferência linguística de superfície, em que o aluno, ao desconhecer a forma morfológica da língua-alvo, recorre à forma correspondente do seu idioma nativo, sobretudo quando se encontra numa fase inicial de aprendizagem.

Do mesmo modo, verifica-se no segmento de informantes de LM alemã e de LM inglesa a ocorrência de formas morfológicas desviantes que resultam de um fenómeno de interferência linguística. Porém, nestes casos não se trata da interferência do sistema linguístico nativo, mas de outra língua românica previamente conhecida pelos informantes, nomeadamente o espanhol, conforme atestam os exemplos a seguir apresentados:

7. *ando a la universidade* (Inglês.A1-A2)

8. *en una rua* (Inglês.A1-A2)

9. *poder escolher o suo estilo* (Alemão.B1-B2)

Além da transferência linguística, os desvios morfológicos assinalados, sobretudo nos textos de informantes de LM alemã e inglesa, parecem resultar da tendência do aprendente para a sobregeneralização da “pseudorregra” de atribuição de valores de género nominal do português, em que o informante assume que se o item termina em *-o* é de género masculino e se termina em *-a* é do género feminino. Tal sobregeneralização incide quer sobre as classes de palavras gramaticais, determinantes e quantificadores, quer sobre as classes de palavras lexicais, nomes e adjetivos:

10. *O minho quarto* (Alemão.A1-A2)

11. *Lembrar cado dia* (Alemão.B1-B2)

12. *Nos anos novento* (Inglês.B1-B2)

13. *Os tempos livros* (Alemão.B1-B2)

14. *Ao dia seguinte* (Alemão.B1-B2)

15. *A minha amiga melhora* (Inglês.A1-A2)

16. *As alimentações grando* (Inglês.B1-B2)

17. *as carpetas* (Alemão.C1)

18. *a historia representa a viagem do casal depois do seu morto* (Inglês.A1-A2)

19. *os semestros* (Alemão.B1-B2)

Estes casos parecem, assim, indicar que os aprendentes assimilaram que o género corresponde a uma propriedade intrínseca do nome e que ele tem de ser marcado nas palavras concordantes. Porém, no momento de estabelecerem a concordância não reconhecem as respetivas formas das palavras concordantes e adaptam-nas em função dos padrões de atribuição de género nominal que interiorizaram a partir do *input* linguístico. O facto de haver, em português, e como se verificou, uma correlação parcial entre índices temáticos *-o* e *-a* e valores de género masculino e feminino, respetivamente, contribui para a assunção, por parte dos aprendentes, que estas vogais correspondem a marcadores morfológicos de género gramatical. Em (18), por exemplo, o aprendente formatou o item lexical feminino “morte” para “morto”, assumindo, conseqüentemente, que se trata de uma forma masculina: *do seu morto*.

Considerações finais

Tendo em conta a análise empreendida, verifica-se que à medida que os aprendentes avançam na aprendizagem do português, regista-se uma descida considerável da frequência de uso de formas morfológicas desviantes, o que parece indicar que a assimilação da estrutura morfológica dos itens corresponde a um processo gradual e progressivo que melhora significativamente ao longo do processo de aprendizagem.

No *corpus* analisado, regista-se ainda que a maior incidência de formas morfológicas desviantes se regista no segmento de informantes de LM espanhola, decorrendo do uso de formas diretamente importadas do espanhol. Assim sendo, infere-se que quanto mais próximo da estrutura formal do português é o sistema linguístico nativo do aprendente, há uma maior proporção de desvios morfológicos, sobretudo em fases iniciais da aprendizagem.

As formas morfológicas desviantes assinaladas neste estudo parecem também resultar de um efeito de sobregeneralização dos padrões de atribuição de valores de gênero disponíveis no português, sendo igualmente reveladores de um certo grau de sensibilidade dos diferentes grupos de aprendentes relativamente aos indícios formais de marcação de valores de gênero nominal em português. Por fim, estes dados permitem-nos inferir algumas das estratégias mobilizadas pelos aprendentes ao longo da aprendizagem do sistema de atribuição e concordância nominal em gênero do português.

REFERÊNCIAS

AIKHENVALD, A. Y. *Classifiers. A Typology of Noun Categorization Devices*. Oxford: University Press, 2000.

BRITO, A. M. Categorias Sintáticas. In.: MATEUS, M. H.M. *et al. Gramática da Língua Portuguesa* (5ª ed., revista e aumentada), Lisboa: Caminho, 2003, p. 323-432.

CHOUPINA, C. M.; BAPTISTA, M. A.; COSTA, J. A. A gramática intuitiva, o conhecimento linguístico e o ensino-aprendizagem do gênero em PE. *Anais do IV SIELP - Simpósio Internacional de Ensino da Língua Portuguesa*, V. 3, n. 1. Uberlândia, 2014.

CHOUPINA, C. M.; BAPTISTA, M. A.; COSTA, J. A.; OLIVEIRA, I.; QUERIDO, J. Conhecimentos e regras explícitos e implícitos sobre o gênero linguístico nos alunos dos 1º e 2º ciclos do Ensino Básico. *Revista da Associação Portuguesa de Linguística*, 1 (1), p. 201-231, 2016.

CORBETT, G. G. *Gender*. Cambridge: Cambridge University Press, 1991.

CORDER, S. P. La importancia de los errors del que aprende una lengua segunda. In.: LICERAS,

J. M. (org.). *La Adquisición de las lenguas extranjeras: hacia un modelo de análisis de la interlengua*. Madrid: Visor, p.31-40, 1992. Tradução espanhola do original inglês de 1967.

FERREIRA, T. S. *Padrões na aquisição/aprendizagem da marcação de gênero nominal em português como L2*. Coimbra: FLUC, 2011. Dissertação (Mestrado em Português Língua Estrangeira/ Língua Segunda), Faculdade de Letras, UC, Coimbra, 2011.

FERREIRA, T. S. *Aquisição/aprendizagem do sistema de atribuição de gênero nominal em PLNM*. Coimbra: FLUC, 2019. Tese (Doutoramento em Linguística do Português), Faculdade de Letras, UC, Coimbra, 2019.

FRANCESCHINA, F. *Fossilized Second Language Grammars – the Acquisition of Grammatical Gender*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2005.

GODINHO, A. P. A aquisição da concordância de número e a sua relação com a aquisição da concordância de gênero: um estudo realizado com aprendentes chineses de português L2. In.: MARÇALO, M. J.; LIMA-HERNANDES, M.; ESTEVES, E.; FONSECA, M.C.; GONÇALVES, O.; VILELA, A.L.; SILVA, A. A. (eds.). *Língua portuguesa: ultrapassar fronteiras juntar culturas*. Universidade de Évora, p. 28-55, 2010.

GRINEVALD, C. Classifiers. In.: BOOIJ, G.; LEHMANN, C.; MUGDAN, J.; SKOPETEAS, S. (eds.). *Morphology: an international handbook on inflection and word formation*. Berlin: Walter de Gruyter, 2000, p. 1016-1031.

IBRAHIM, M. H. *Grammatical Gender*. Paris: Mouton, 1973.

LACSÁN, V. *The acquisition of gender agreement in L2 Portuguese by adult Hungarian speakers*. Lisboa: FLUL, 2015. Dissertação (Mestrado em Linguística do Português), Faculdade de Letras, UL, 2015.

LEIRIA, I. *Léxico, Aquisição e Ensino do Português Europeu língua não materna*. Lisboa: Fundação Calouste Gulbenkian, Fundação para a Ciência e Tecnologia, 2006.

MARIOTTO, E. M. C. & LOURENÇO-GOMES, M. C. Análise de erros na escrita relacionados à aprendizagem da concordância de gênero por falantes nativos do inglês, aprendentes de português europeu como língua estrangeira. In.: *Anais do IV Simpósio Mundial de Estudos de Língua Portuguesa (SIMELP)*. Língua Portuguesa: ultrapassando fronteiras, unindo culturas. Faculdade de Letras/UFG, Goiânia, Goiás, Brasil, p. 1278-1285, 2013.

MARTINS, C. S. P. Número e gênero nominais no desenvolvimento das interlínguas de aprendentes de português europeu como língua estrangeira. *Revista Científica da Universidade Eduardo Mondlane*, Série Letras e Ciências Sociais, 1 (1), p. 26-51, 2015.

MOTA, M. A. Morfologia nas interfaces. In.: MARTINS, A. M.; CARRILHO, E. (eds.). *Manual de Linguística Portuguesa*, Berlin/Boston: De Gruyter, 2016a, p. 156-177.

MOTA, M. A. Morfologia do nome e do adjetivo. In.: RAPOSO, E. B. P.; DO NASCIMENTO, M. F. B.; MOTA, M. A.; SEGURA, L.; MENDES, A. (eds.). *Gramática do Português*, V. III. Lisboa: Fundação Calouste Gulbenkian, 2020, p. 2833-2930.

PINTO, J. A aquisição do gênero e da concordância de gênero em português língua terceira ou língua adicional. In.: Osório, P.; Grosso, M. J. (eds.). *Teorias e Usos Linguísticos – Aplicações ao Português Língua Não Materna*. Lisboa: Lidel, p. 91-110, 2017.

RIO-TORTO, G.; RODRIGUES, A. S. Formação de nomes. In.: RIO-TORTO, G. *et al. Gramática Derivacional do Português* (2ª edição), Coimbra: Imprensa da Universidade de Coimbra, 2016, p. 135-240.

RITTER, E. Where's Gender?. *Linguistic Inquiry*, 24 (4), p. 795-803, 1993.

SABOURIN, L.; STOWE, L. A.; DE HAAN, G. D. Transfer effects in learning a second language grammatical gender system. *Second Language Research*, SAGE Publications, 22(1), p. 1-29, 2006.

SELINKER, L. La Interlengua. In.: LICERAS, J. M. (org.). *La Adquisición de las lenguas extranjeras: hacia un modelo de análisis de la interlengua*. Madrid: Visor, p.79-101, 1992. Tradução espanhola do original inglês de 1972.

VILLALVA, A. *Estruturas morfológicas. Unidades e Hierarquias nas Palavras do Português*. Lisboa: Fundação Calouste Gulbenkian, 1994.

VILLALVA, A. Estrutura mórfica básica. In Mateus, In.: MATEUS, M. H.M. *et al. Gramática da Língua Portuguesa* (5ª ed., revista e aumentada), Lisboa: Caminho, 2003, p. 919-931.

WHITE, L.; VALENZUELA, E.; KOZLOWSKA-MACGREGOR, M.; LEUNG, Y. I. Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 2, p. 105-133, 2004.