

# A atribuição de valores de género nominal em produções orais de aprendentes tardios de português como língua não materna

The attribution of nominal gender in oral productions of late learners of Portuguese as a non-native language

TÂNIA SANTOS FERREIRA\*

PALAVRAS-CHAVE: Aquisição tardia de português como língua não materna, Valores de género gramatical, Produção oral, Conhecimentos linguísticos prévios.

KEYWORDS: Late acquisition of Portuguese as non-native language, Grammatical gender values, Oral production, Prior linguistic knowledge.

## 1. Introdução

Os desvios de atribuição de género e de concordância nominal são persistentes nas produções dos aprendentes de Português como Língua Não Materna (PLNM), mesmo nos níveis mais avançados de proficiência linguística (Godinho, 2010; Ferreira, 2011, 2019; Mariotto e Lourenço-Gomes, 2013; Martins, 2015; Lacsán, 2016; Pinto, 2017). Muita da investigação produzida neste âmbito assenta na análise de desvios detetados em produções escritas por aprendentes tardios, inseridos em contexto instrucional, sendo ainda pouco numerosos os trabalhos que assentam em dados de produções orais (Martins, 2020, p. 171). Surge, portanto, a necessidade de complementar os estudos já produzidos com dados de natureza oral, de modo a obter uma descrição mais detalhada acerca do processo de assimilação da categoria gramatical de género por aprendentes tardios de PLNM.

A base empírica deste estudo integra produções selecionadas do *Corpus* Oral de Português L2 – Coimbra (COraL-Co), produzidas por falantes nativos de espanhol e de inglês, a frequentar turmas de três níveis de proficiência linguística distintos: A2, B1 e B2, estipulados de acordo com o Quadro Europeu Comum de Referência para as Línguas (QECL) (Conselho da Europa, 2001).

\* Universidade de Coimbra, Centro de Estudos de Linguística Geral e Aplicada (CELGA-ILTEC), Portugal

Procura-se, deste modo, confrontar o desempenho linguístico destes aprendentes, tendo em conta, por um lado, a configuração do conhecimento linguístico prévio proveniente da Língua Materna (LM) e o respetivo nível de proficiência linguística (nível QECRL), e, por outro, as propriedades relativas ao sistema de atribuição de valores de género aos nomes e de concordância nominal em género do português. Os dados selecionados permitem efetuar uma Análise Contrastiva Interlíngua (Granger, 1996, 2009). Em função da observação dos desvios de atribuição de género gramatical produzidos por falantes de LM espanhola e de LM inglesa, será possível aferir padrões de aquisição correlacionados quer com a configuração do sistema linguístico nativo dos informantes, quer com padrões transversais a todos os aprendentes, independentemente da sua LM<sup>1</sup>.

Este trabalho encontra-se organizado da seguinte forma: após esta breve introdução, o ponto 2 é dedicado ao enquadramento teórico. Em 2.1, descreve-se o sistema de atribuição de género nominal do português e em 2.2 é feita a uma breve revisão da literatura disponível acerca da aquisição tardia de género gramatical, com especial enfoque para as hipóteses relativas ao papel dos conhecimentos linguísticos prévios para o desenvolvimento (inter)linguístico. No ponto 3, apresenta-se o estudo empírico. Após a apresentação dos procedimentos metodológicos (3.1), no ponto 3.2 apresentam-se e discutem-se os resultados apurados. Por fim, em 4, assinalam-se as considerações finais e as perspectivas de trabalhos futuros.

## **2. Enquadramento**

### **2.1. O sistema de atribuição de género aos nomes em português**

Na língua portuguesa, todos os nomes e alguns pronomes são especificados quanto ao valor de género gramatical, havendo dois valores em oposição: masculino e feminino. A associação de um valor de género afeta outras classes de palavras pela necessidade de estabelecerem relações de concordância sintática

<sup>1</sup> Como afirma Granger (1996, pp. 47-48) “[i]t is thus essential to conduct comprehensive IL comparisons in order to distinguish between L1-related features and cross-linguistic invariants, i.e. those features which are common to all learners, irrespective of mother tongue”.

com o nome (Mota, 2020). No interior do sintagma nominal (SN), a concordância estabelece-se entre o núcleo nominal e (i) as expressões de determinação e de quantificação que o antecedem (especificadores); e (ii) os adjetivos integrados em sintagmas adjetivais que funcionam como seus modificadores ou elementos apositivos (Brito, 2003, p. 328).

Segundo Mota (2020), a categoria de género gramatical encontra-se inerentemente especificada na matriz dos traços da entrada lexical nominal. Nas demais classes de palavras, que concordam com o nome, o género corresponde a uma categoria morfossintática «de concordância» (*Ibid.*, p. 2816), uma vez que esses elementos adquirem um dos dois valores possíveis por via da concordância sintática.

Quanto aos critérios de associação de valores de género, verifica-se que em grande parte dos itens com o traço semântico [+sexuado], o valor de género corresponde ao género natural da entidade designada. Existem formas nominais masculinas que se referem a seres do sexo masculino (o homem) e formas nominais femininas com referentes do sexo feminino (a mulher). Há ainda itens nominais, denominados comuns de dois, com referentes sexuados que admitem a associação a mais do que um valor de género, como o/a motorista, o/a viajante. Nestes casos, Mota (2016) considera que se trata de dois lexemas distintos, sendo que cada entrada lexical, masculina ou feminina, vai ser ativada na sintaxe em função do sexo do respetivo referente (*Ibid.*, p. 156). É também de assinalar os casos de itens cujo valor de género se associa a um conteúdo semântico particular. Por exemplo, o item polícia classifica-se como um nome (i) feminino; e (ii) comum de dois. Quando associado ao valor de género feminino, refere-se à instituição responsável pela segurança pública. Como comum de dois, a forma do feminino designa a pessoa, de sexo feminino, que desempenha as funções de agente da polícia, por oposição à forma marcada como masculina que designa o indivíduo, de sexo masculino, agente da polícia.

Existem também nomes que, apesar de terem referentes sexuados, possuem um único valor de género independentemente do sexo do seu referente, como a pessoa, o indivíduo, a testemunha. Já no caso dos itens com referentes não sexuados, verifica-se uma correlação parcial entre índices temáticos -o e -a e valores de género masculino e feminino, respetivamente, mas nem sempre essa correlação se observa, visto que se encontram na língua portuguesa nomes marcados como masculinos terminados em -a (o problema, o dilema, o dia) e nomes femininos terminados em -o (a líbido, a tribo, e formas truncadas como a foto, a moto) (Rio-Torto e Rodrigues, 2016, p. 158; Mota, 2020, p. 2901). Além disso, existem na língua portuguesa itens nominais com diferentes

terminações que se associam quer ao género masculino quer ao género feminino (por exemplo, o pontapé, a chaminé, o coração, a paixão, o papel, aatedral).

Em suma, em português a atribuição de um valor de género aos nomes resulta do cruzamento de critérios semânticos e formais que não atuam de modo exclusivo nem exaustivo. As características do sistema de atribuição de género aos nomes demonstram, por um lado, a fraca robustez dos dados disponíveis no *input*, seja numa perspectiva formal seja numa perspectiva semântica e, por outro, a relevância da concordância sintática para determinar o valor de género do item nominal. Por conseguinte, o facto de a concordância ser um mecanismo reconhecidamente problemático na aquisição tardia de uma LNM (Franceschina, 2005; Leiria, 2006; Godinho, 2010; Martins, 2015; Ferreira, 2019) contribuirá para justificar a relativa dificuldade dos aprendentes tardios em assimilar, na plenitude, os valores de género dos nomes da língua-alvo de aprendizagem.

## **2.2. Aquisição da categoria de género gramatical por falantes não nativos do português**

No âmbito da investigação sobre a aquisição/aprendizagem da categoria de género gramatical em PLN, verifica-se que itens cujo valor de género gramatical corresponde ao género natural da entidade designada tendem a apresentar índices reduzidos de desvios, por oposição aos itens cujo valor de género não depende nem de critérios de natureza semântica nem de critérios de natureza formal (Ferreira, 2011, 2019; Martins, 2015; Pinto, 2017). A partir dos dados analisados por Ferreira (2011, 2019) e por Martins (2015), extraídos da mesma base de dados, o *Corpus* de Produções Escritas por Aprendentes de Português L2 (PEAPL2) (Martins et al., 2019), verifica-se que os nomes atemáticos registam um maior número de desvios de atribuição de género nominal. Porém, também os nomes cujo valor de género seria corretamente dedutível a partir da aplicação da ‘pseudorregra’: “se o nome termina em -o é de género masculino e se o nome termina em -a é de género feminino”, registam índices de desvio consideravelmente elevados (Ferreira, 2011, p. 63; Ferreira, 2019, p. 230; Martins 2015, p. 41). Atendendo a estes dados, assume-se que a fraca robustez do *input* linguístico contribui para uma certa ‘desconfiança’ dos aprendentes relativamente ao poder preditivo dos indícios morfológicos (Ferreira, 2019, p. 254). Vejamos se os dados deste trabalho confirmam igualmente esta tendência.

No que concerne às estratégias adotadas ao longo do processo de assimilação da categoria de género em PLNM, verifica-se a adoção de uma “marcação *default* da forma masculina” (Mariotto e Lourenço-Gomes, 2013, p. 1282), ou seja, no momento de atribuir um valor de género aos nomes, os aprendentes tendem a preferir a seleção do masculino, a forma não-marcada, em detrimento do feminino (Ferreira, 2011, 2019; Mariotto e Lourenço-Gomes, 2013; Martins, 2015).

Quanto aos constituintes sintáticos nos quais se explicitam as marcas de atribuição desviante do valor de género, regista-se um maior número de ocorrências desviantes de atribuição de género nominal nos especificadores, por oposição aos modificadores (Godinho, 2010; Ferreira, 2011, 2019; Mariotto e Lourenço-Gomes, 2013; Martins, 2015; Pinto 2017). No confronto das produções desviantes de dois grupos de aprendentes, falantes nativos de espanhol e de inglês, Mariotto e Lourenço-Gomes (2013, p. 1283) verificaram que os aprendentes, falantes nativos de inglês, apresentam uma maior tendência para a alteração da forma morfológica do nome e conseqüente alteração do seu valor de género. Da análise global do desempenho dos aprendentes e dos desvios produzidos por cada segmento da amostra, as autoras concluem que “adultos que aprendem uma língua com estrutura morfológica semelhante à da sua língua materna tendem a ser mais consistentes na aplicação da concordância de género [...]” (*Ibid.*, p. 1283).

No que diz respeito à configuração da gramática do idioma nativo do aprendente, Martins (2015) observa que os falantes nativos de chinês registam um melhor desempenho global na atribuição dos valores de género nominal em português face ao registado em aprendentes falantes nativos de espanhol e de alemão. Segundo a autora, a partir do desempenho dos informantes de LM espanhola e alemã é possível inferir um certo efeito de transferência negativa “sob a forma de uma atitude de insegurança por parte do aprendente” (*Ibid.*, p. 44). Também Ferreira (2019), que analisou os desvios de atribuição e de concordância nominal de género em produções escritas por aprendentes tardios, falantes nativos de espanhol, italiano, alemão, inglês e chinês, verificou que a configuração do conhecimento linguístico prévio do aprendente interfere no ritmo em que se desenvolve o processo de assimilação da categoria gramatical de género, já que entre os falantes nativos das línguas românicas, como o espanhol e o italiano, se registam, nos níveis iniciais de aprendizagem, índices de desvios relativamente mais baixos aos verificados nos outros segmentos. Porém, observa-se, sobretudo entre os aprendentes de LM espanhola, um certo efeito de estabilização da aprendizagem, já que a incidência de desvios apurada no nível mais avançado, C1, não é consideravelmente mais baixa do que a

registada nos níveis elementares, A1 e A2. Deste modo, a investigadora considera que estes aprendentes “estabilizam cedo” e, por isso, “ficarão impermeáveis aos efeitos do *input* que [...] não geram [...] conflitos de hipóteses sobre a forma que deverá assumir a gramática da L[íngua]-A[alvo]” (*Ibid.*, p. 292).

Por fim, os dados aqui sucintamente apresentados indiciam que, no contexto da aquisição tardia do sistema de atribuição de género nominal em PLNM, a configuração do conhecimento linguístico prévio dos aprendentes parece atuar conjuntamente com outros fatores, característicos do próprio desenvolvimento interlinguístico.

### 3. Estudo empírico

#### 3.1. Metodologia

A base empírica do presente estudo foi extraída do *Corpus* Oral de Português L2 – Coimbra (COral-Co) (coord. Isabel Santos) (Santos et al., 2016)<sup>2</sup>. Os dados orais disponíveis neste acervo resultam da aplicação de um protocolo de recolha (Santos et al., 2016, p. 749). Para este trabalho em particular, selecionaram-se os dados transcritos das produções orais que resultaram da aplicação da tarefa 3 “Construção de um texto narrativo a partir de uma sequência de imagens”. Nesta tarefa, foi apresentada uma tira de banda desenhada aos informantes que deveriam verbalizar oralmente todos os pormenores da sequência narrativa representada<sup>3</sup> (Santos et al., 2016, p. 751). A estrutura narrativa envolve quatro personagens distintas: (i) uma senhora; (ii) um ladrão; (iii) um homem, que é agente da polícia; e (iv) um juiz. Estas personagens vão surgindo e interagindo em três espaços distintos: na rua, na esquadra da polícia e no tribunal, e a história desenvolve-se em torno do assalto da mala da senhora e do facto de o ladrão, o polícia e o juiz, que surge num tribunal, se conhecerem e serem amigos.

Atendendo aos objetivos do presente trabalho, selecionaram-se as produções orais de aprendentes de LM espanhola e inglesa. O espanhol apresenta, como

<sup>2</sup> O *COral-Co* encontra-se disponível para consulta, em regime de acesso aberto, em <http://teitok2.iltec.pt/coralco/>.

<sup>3</sup> A sequência de imagens apresentada foi recolhida da obra *Condições Humanas* da autoria de Quino e publicada em 1992.

o português, dois valores de género em oposição (masculino e feminino), tendo um processo de marcação morfossintática e de concordância nominal muito semelhante (Harris, 1991; Ambadiang, 1999). Em contrapartida, o inglês não exibe um sistema de classificação nominal baseado em valores de género nem a obrigatoriedade do estabelecimento de relações de concordância nominal (Ibrahim, 1973, p. 86), embora conserve vestígios da presença de um primitivo sistema de três valores (masculino, feminino e neutro) ao nível dos pronomes pessoais *he, she, it*.

Além da LM, selecionaram-se as produções em função do nível de proficiência e de competência linguísticas da turma frequentada pelos aprendentes no momento da recolha (nível QECRL). Tendo em conta os dados disponíveis no COral-Co, selecionaram-se as produções orais de aprendentes a frequentar turmas do nível elementar, A2, do nível pré-intermédio, B1, e do nível intermédio, B2. Dado o facto de o número absoluto de textos disponíveis no acervo consultado não ser equitativamente homogéneo pelos diferentes grupos de informantes e respetivos níveis QECRL<sup>4</sup>, foi necessário selecionar e agrupar as produções orais em dois subconjuntos distintos: um composto por produções de informantes a frequentar os níveis A2 e B1, e outro de informantes a frequentar turmas de nível B2. Portanto, analisou-se um conjunto global 12 produções orais, equitativamente distribuídas por LM e nível de proficiência linguísticas (cf. Tabela 1).

Os informantes são de ambos os sexos (50% do sexo feminino e 50% do sexo masculino) e têm uma média de idades de 31 anos (distribuídos entre os 20 e os 62 anos). A duração média global das produções orais é de aproximadamente 02 minutos, sendo o tempo médio da duração das produções do segmento da amostra de aprendentes de LM espanhola de 00:01:20, um valor médio inferior ao registado nas produções de LM inglesa (00:02:25). A diferença da duração poderá estar, porventura, correlacionada com a proximidade tipológica do espanhol e do português.

4 Com efeito, no COral-Co, na tarefa 3, encontram-se disponíveis 6 produções orais de aprendentes de LM inglesa (2 no nível A2; 1 no nível B1 e 3 no nível B2). Já entre os falantes de LM espanhola, encontram-se 12 produções orais (4 no nível A2; 4 no nível B1 e 4 no nível B2). Dada a diferença quantitativa entre os dois segmentos, foi necessário selecionar produções de forma a obter conjuntos equitativamente homogéneos nos diferentes níveis QECRL.

Informantes			Produções orais	
LM	Nível QECRL	#	#	Tempo médio da gravação hh:mm:ss
Espanhol	A2 / B1	3	3	00:01:18
	B2	3	3	00:01:22
<b>Σ</b>		<b>6</b>	<b>6</b>	<b>00:01:20</b>
Inglês	A2 / B1	3	3	00:02:10
	B2	3	3	00:02:40
<b>Σ</b>		<b>6</b>	<b>6</b>	<b>00:02:25</b>
<b>Σ</b>		<b>12</b>	<b>12</b>	<b>00:01:52</b>

Tabela 1: Distribuição do número de informantes, de produções orais e respetivo tempo médio da duração das produções por LM e Nível QECRL.

A fim de averiguar o desempenho linguístico destes aprendentes no que à associação de valores de género e à concordância nominal de género diz respeito, procedeu-se ao levantamento, nas transcrições dos ficheiros selecionados, das ocorrências dos itens nominais por produção oral. Estabelecido o universo de nomes foi, então, possível calcular, a partir dele, a proporção de itens integrados em sintagmas nominais (SN) constituídos por um núcleo, o nome, que coocorre com especificadores (determinantes e/ou quantificadores) e/ou adjetivos (com função atributiva e/ou predicativa), por oposição à ocorrência de nomes isolados. O valor apurado de SN servirá, depois, de referência para o cálculo das percentagens relativas de desvios por segmentos da amostra empírica, visto que só estes contextos permitem aferir o valor de género atribuído pelo aprendente.

Para o tratamento e análise dos casos desviantes assinalados, identificaram-se as ocorrências desviantes que incidem sobre os especificadores (determinantes e quantificadores), sobre os adjetivos (com função atributiva e/ou com função predicativa) e sobre os nomes<sup>5</sup>.

<sup>5</sup> Assinale-se que nas situações que resultam da alteração da forma morfológica do nome foram considerados desvios sempre que a formatação do item resulta também no uso de um valor de género gramatical oposto ao da forma correspondente em português. Por exemplo, no segmento “\*na processa da justicia” assinalaram-se dois tipos de desvios: um relativo à forma morfológica do nome (uso de “processa” em vez de “processo”), e outro respeitante à forma do determinante que coocorre com o nome no SN (uso da forma feminina “a” em vez do masculino “o”).



A Tabela 2 contém a distribuição dos informantes por LM e nível QECRL, bem como o número absoluto (#) de sintagmas nominais (SN) apurado por segmento da amostra, o número absoluto (#) de desvios de atribuição e de concordância nominal em género e, por fim, a média de desvios por produção oral.

Informantes				Produções orais						
LM	Nível QECRL	#		#		# SN		# Desvios		Média de desvios por produção oral
Espanhol	A2/B1	3	6	3	6	55	108	7	13	2
	B2	3		3		53		6		2
Inglês	A2/B1	3	6	3	6	58	150	22	34	5
	B2	3		3		92		12		4
$\Sigma$		12		12		258		47		3

Tabela 2: Caracterização da base empírica do estudo.

A Tabela 3 contempla alguns exemplos de desvios retirados do *corpus* deste trabalho e que foram codificados em função do constituinte afetado.

	Constituinte afetado	Exemplos
Especificador	Determinante	<i>Um juíza (...) está a anda</i> (Inglês.B2)
		<i>E a polícia e o homem são amigos (...)</i> (Espanhol.B2)
	Quantificador	<i>A ladrão na pistola (...)</i> (Inglês.A2)
		<i>E todos os pessoas presentes (...)</i> (Espanhol.A2)
		<i>(...) um pistola</i> (Inglês.B2)
Adjetivo	Em posição atributiva	<i>A velho amiga dela (...)</i> (Inglês.A2)
	Em posição predicativa	<i>E todos os pessoas presentes (...) quedaron sorprendidos</i> (Espanhol.A2)
	Nome	<i>O bolso da senhora</i> (Espanhol.B2)
		<i>Na processa da justicia (...)</i> (Inglês.B2)

Tabela 3: Distribuição de desvios de atribuição e de concordância nominal em género por constituintes afetados.

### 3.2. Resultados e discussão

Do conjunto global de 258 ocorrências de itens nominais (cf. Tabela 2), registou-se um total de 47 desvios de atribuição e de concordância nominal em género que afetam 44 itens. Portanto, em termos proporcionais, verifica-se que os desvios relativos à categoria gramatical de género recaem sobre cerca de 18% das ocorrências de nomes registadas nas produções orais selecionadas (cf. Gráfico 1).

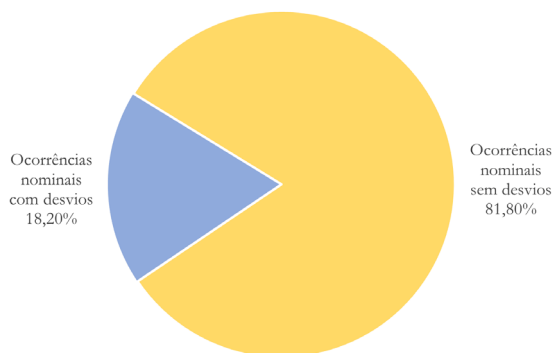


Gráfico 1: Distribuição (em valores percentuais) das ocorrências nominais com desvios de atribuição e de concordância nominal em género por oposição às ocorrências nominais sem desvios.

Os valores apurados indiciam que, em geral, os aprendentes selecionam adequadamente o género dos nomes que utilizam. Por produção oral, registou-se, em média, a ocorrência de três desvios, sendo que é nas produções dos informantes de LM inglesa que se regista o valor médio de ocorrências desviantes mais elevado (cf. Tabela 2).

Com vista a aferir a proporção global de desvios por segmentos da amostra, calcularam-se percentagens relativas de desvios tomando em linha de conta o número absoluto de ocorrências de sintagmas nominais e de desvios, por conjuntos de informantes (cf. Tabela 2). Os resultados encontram-se cartografados no Gráfico 2.

O segmento de informantes de LM espanhola regista, em termos proporcionais, valores de incidência de desvios mais baixos relativamente aos registados no segmento dos aprendentes que são falantes nativos de inglês. Com efeito, entre os níveis A2 e B1, verifica-se uma proporção de desvios mais reduzida nas produções dos informantes de LM espanhola (12.73%) face à

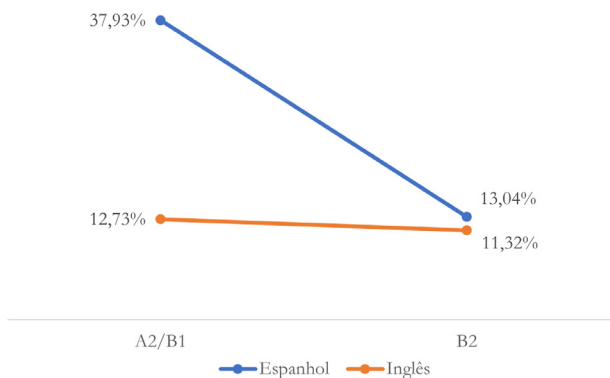


Gráfico 2: Distribuição das percentagens relativas de desvios por LM e Nível QECRL da turma frequentada pelos aprendentes.

registada, nos mesmos níveis, nas produções dos informantes de LM inglesa (37.93%). A diferença de desempenho entre os dois segmentos é muito menos expressiva no nível B2, já que os valores apurados nos dois grupos são muito próximos. Com efeito, as produções orais de aprendentes de LM inglesa do nível B2 registam uma proporção de desvios de 13.04%, ligeiramente superior à registada nas produções dos aprendentes de LM espanhola (11.32%).

Portanto, os dados indiciam que, embora os falantes nativos de inglês apresentem inicialmente maiores dificuldades na atribuição de valores de género aos nomes, ao longo da progressão da aprendizagem do português o seu desempenho linguístico regista uma considerável melhoria, já que no nível B2, a proporção de desvios é manifestamente inferior. Por sua vez, as produções orais dos informantes de LM espanhola evidenciam um certo efeito de estabilização da aprendizagem, dado que mesmo em B2 a proporção de desvios não é substancialmente mais reduzida do que a registada nos níveis precedentes, A2 e B1. Assim, e à semelhança do que já observara Ferreira (2019), e também Martins (2015), poder-se-á admitir que a proximidade tipológica dos idiomas terá efeitos relativamente ao ritmo em que se processa o desenvolvimento da aprendizagem do sistema de atribuição e de concordância nominal em género do português.

Atendendo a que as produções orais resultam da descrição de um mesmo estímulo pictórico, as seqüências narrativas criadas pelos aprendentes evidenciam o uso de um conjunto de nomes relativamente restrito. No segmento dos informantes de LM espanhola, os 13 desvios assinalados afetam apenas 4 formas

nominais distintas (bolsa, pessoa, polícia e quartel) (cf. Tabela 4), e entre os informantes de LM inglesa, os 34 desvios assinalados recaem sobre 14 nomes (amiga, arma, carteira, final, juiz, juíza, ladrão, mala, mão, mulher, pistola, polícia, processo e saco) (cf. Tabela 5).

Nomes afetados	Nível QECRL			
	A2/B1		B2	
	O* / D**	% de D**	O* / D**	% de D**
bolsa	–	–	2/1	50
pessoa	3/3	100	–	–
quartel	2/1	50	–	–
polícia	10/3	30	10/5	50

O\* = Ocorrências; D\*\* = Desvios

Tabela 4: Distribuição do número absoluto de ocorrências e de desvios dos nomes afetados e das percentagens relativas de desvios apuradas nos segmentos dos informantes de LM espanhola.

Nomes afetados	Nível QECRL			
	A2/B1		B2	
	O* / D**	% de D**	O* / D**	% de D**
amiga	1/1	100	–	–
arma	1/1	100	3/0	0
carteira	1/1	100	–	–
final	–	–	2/1	50
juiz	2/1	50	5/1	20
juíza	–	–	3/1	33.3
ladrão	12/4	33.3	19/1	5.2
mala	2/1	50	4/0	0
mãos	5/5	100	–	–
mulher	5/1	20	5/0	0
pistola	2/1	50	1/1	100
polícia	9/6	66.7	15/4	26.7
processo	–	–	2/2	100
saco	–	–	1/1	100

O\* = Ocorrências; D\*\* = Desvios

Tabela 5: Distribuição do número absoluto de ocorrências e de desvios dos nomes afetados e das percentagens relativas de desvios apuradas no segmentos dos informantes de LM inglesa.

Como se pode observar nas Tabelas 4 e 5, o peso relativo de desvios face ao número de ocorrências dos nomes afetados no *corpus* por segmento da amostra é, salvo alguns casos pontuais, elevado, isto é, ao utilizarem determinado nome, os aprendentes erram mais do que acertam.

A proporção de nomes afetados nos dois segmentos da amostra ilustra, além disso, uma diferença qualitativa de desempenho, traduzindo uma maior dificuldade, por parte dos aprendentes, falantes nativos de inglês, em associar adequadamente o valor de género aos diferentes nomes que produzem. Poder-se-á, deste modo, admitir que estes aprendentes ainda não assimilaram, na plenitude, o valor de género de cada nome. O melhor desempenho dos informantes de LM espanhola poderá estar correlacionado com a proximidade tipológica dos idiomas, LM e língua-alvo de aprendizagem. Assim, o facto de haver, no espanhol e no português, formas nominais que apresentam uma estrutura formal semelhante e o mesmo valor de género, contribuirá, porventura, para fenómenos de transferência direta e uma conseqüente maior taxa de acertos, mesmo em fases iniciais da aprendizagem.

Em seguida, pretendeu-se averiguar se os desvios assinalados resultam preferencialmente do uso generalizado de um valor de género, em detrimento de outro, já que os dados apurados em outros estudos (Mariotto e Lourenço-Gomes, 2013, p. 1282; Martins, 2015, p. 39; Ferreira, 2019, p. 204) apontam para a adoção, por parte do aprendente, de uma estratégia de atribuição por defeito do valor de género masculino, a forma não-marcada. Assim, calculou-se a proporção de desvios por segmentos da amostra em função do valor de género nominal da forma afetada (cf. Gráfico 3).

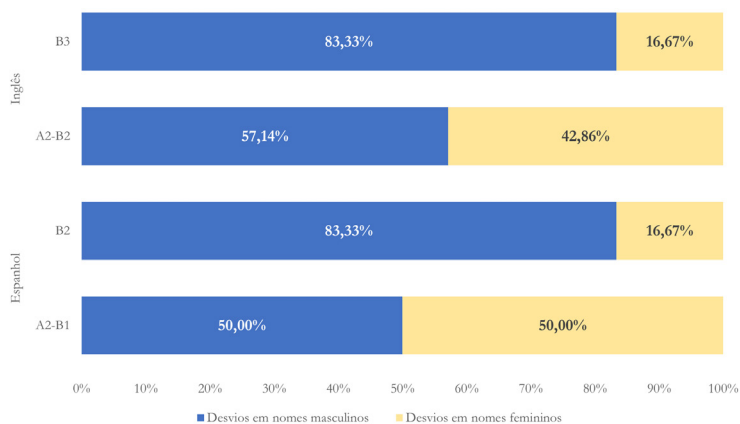


Gráfico 3: Distribuição, por segmentos da amostra empírica, da proporção de desvios em função do valor de género nominal da forma afetada.

Como se pode constatar, os dados apurados não revelam o recurso a uma estratégia de marcação do valor de género masculino por defeito o que poderá estar correlacionado com o facto de uma parte substancial dos desvios no *corpus* incidir na associação do valor de género feminino ao nome comum de dois “pólicia” que, em função do contexto da sequência narrativa criada, deveria ser marcado como masculino (cf. (1)-(4)):

- (1) O ladrão levanta os mãos depois **a pólicia** também levanta os mãos (...) (Inglês.B1)
- (2) **Uma pólicia** chega (...) (Inglês.A2)
- (3) Olhou **a pólicia** (...) Espanhol.A2)
- (4) E **a pólicia** e o homem são amigos (...) (Espanhol.B2)

A tendência observada poderá, por um lado, estar correlacionada com o facto de estes aprendentes assumirem implicitamente que o item nominal “pólicia” é sempre marcado como feminino na língua portuguesa e de não reconhecerem, por isso, a possibilidade de associar a este item o valor de género masculino. Também a ocorrência da expressão “a pólicia” para designar a entidade de segurança pública poderá contribuir para que os aprendentes assumam que este item é somente marcado como feminino. Os aprendentes, independentemente do seu conhecimento linguístico prévio terão, assim, internalizado, na sua gramática mental, que este nome tem um valor de género único. Por outro lado, a incidência de desvios também pode resultar de um efeito de sobregeneralização da ‘pseudorregra’ de atribuição de valores de género aos nomes em português, “se o nome termina em -o é de género masculino, se termina em -a é feminino”.

Procurou-se, por fim, determinar a proporção de desvios por segmentos da amostra em função do constituinte sintático afetado. Os resultados encontram-se representados no Gráfico 4.

No que diz respeito aos constituintes afetados, verifica-se o mesmo padrão nos diferentes segmentos da amostra empírica, já que, é nos especificadores, nomeadamente no determinante que recai o maior número de desvios. Em contrapartida, a proporção de desvios que afetam os adjetivos, sobretudo em posição atributiva, é, em geral, pouco expressiva.

Poderemos analisar estes resultados atendendo, em primeiro lugar, ao facto de, nas produções orais analisadas, se ter verificado uma frequência de uso de sintagmas constituídos apenas por determinante e nome muito mais elevada do que a de sintagmas constituídos por especificadores e modificadores, quer

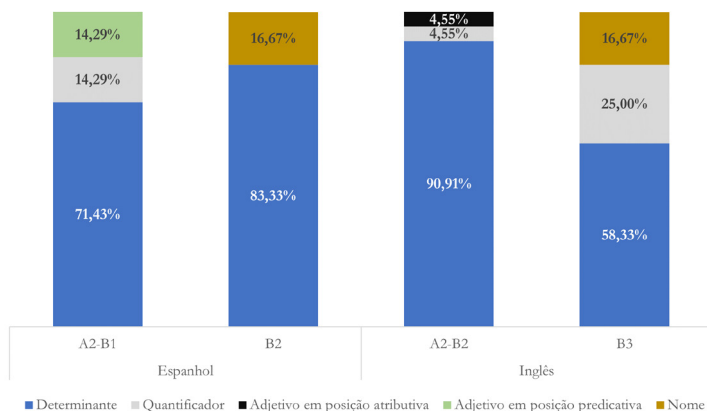


Gráfico 4: Distribuição da proporção de desvios por constituintes afetados.

com função atributiva quer com função predicativa. Assim, sendo mais frequentes, a probabilidade de ocorrência de desvios será proporcionalmente maior.

No *corpus* analisado, verificam-se ainda manifestações desviantes que resultam da ambivalência de atribuição de valores de género. Nos exemplos (5) e (6) verifica-se que os informantes, apesar de selecionarem adequadamente a forma do adjetivo, não utilizam corretamente a forma do especificador.

(5) *A **polícia** está **acusado*** (Inglês.A2)

(6) *Supreendeu **à ladrão** que ficou também **assustado*** (Inglês.B2)

Estes exemplos revelam, conseqüentemente, que é sobretudo nos elementos que constituem a estrutura funcional do sintagma que se manifestam as maiores dificuldades em estabelecer a concordância nominal em género.

#### 4. Considerações finais

O presente trabalho permitiu-nos aferir padrões de atribuição de valores de género nominal por aprendentes tardios de PLNLM, falantes nativos de espanhol e de inglês, a frequentar turmas de três níveis de aprendizagem, de A2 a B2. O estudo empírico baseou-se na análise de desvios de atribuição e de concordância nominal detetados em produções orais e no qual se procurou obter dados que permitam a descrição do desempenho dos aprendentes em diferentes estádios de aprendizagem e em função dos conhecimentos linguísticos prévios.

Na perspectiva da Análise Contrastiva Interlíngua (Granger, 1996, 2009), os dados empíricos permitem-nos aferir que o facto de os aprendentes terem representada na sua LM a categoria gramatical de género e de possuírem um sistema de atribuição e de concordância de género semelhante ao do português, como é o caso do espanhol, se traduz por um reduzido número de desvios relativos à atribuição de valores de género aos nomes, logo em fases iniciais da aprendizagem linguística. Em contrapartida, à medida que progredem na aprendizagem, verifica-se um certo efeito de estabilização, já que a proporção de desvios assinalada entre os segmentos dos informantes de LM espanhola e inglesa no nível B2 é muito próxima. Ou seja, o efeito potencialmente benéfico dos conhecimentos linguísticos prévios deixa de se fazer sentir em fases mais avançadas do desenvolvimento interlinguístico dos aprendentes.

Os dados revelam ainda diferenças qualitativas de desempenho entre os dois grupos de informantes, já que a variedade de itens nominais com desvios é diferente. A maior proporção de nomes afetados registada nas produções dos informantes de LM inglesa face à proporção de itens afetados nas produções dos informantes de LM espanhola constitui um indicador de que uma maior proximidade estrutural auxilia o processo de assimilação dos valores de género associados a cada item.

Relativamente às estratégias adotadas no momento de atribuir um valor de género aos nomes, não se evidencia, nas produções analisadas, a marcação preferencial de um valor em detrimento de outro. No que concerne aos constituintes nos quais se evidenciam as marcas de desvio, verifica-se em ambos os segmentos o mesmo padrão, sendo os especificadores os que apresentam maiores incidências de desvios.

Em suma, pese embora o reduzido número de textos orais analisados, neste trabalho foi possível assinalar um conjunto de particularidades subjacentes ao processo de aquisição tardia da categoria de género em função dos conhecimentos linguísticos prévios dos aprendentes e dos respetivos níveis de proficiência linguística. Reconhecemos, por fim, a necessidade empreender mais estudos que complementem estes resultados, com vista a uma melhor compreensão acerca do processo de assimilação da categoria de género gramatical no contexto da aquisição tardia de PLNM.



## Referências bibliográficas

- AMBADIANG, T. (1999). La flexión nominal. Género y número. In I. Bosque, y V. Demonte (eds.), *Gramática descriptiva de la lengua española* (vol. 3, pp. 4843-4913), Madrid: Espasa Calpe.
- BRITO, A. M. (2003). Categorías Sintácticas. In M. H. M. Mateus, A. M. Brito, I. Duarte, I. H. Faria, S. Frota, G. Matos, F. Oliveira, M. Vigário, e A. Villalva (eds.), *Gramática da Língua Portuguesa* (vol. I, pp. 323-432), Lisboa: Caminho.
- CONSELHO DA EUROPA (2001). *Quadro Europeu Comum de Referência para as Línguas. Aprendizagem, ensino, avaliação*. Edições Asa.
- FERREIRA, T. S. (2011). *Padrões na aquisição/aprendizagem da marcação de género nominal em português como L2*. [Dissertação de mestrado em Português Língua Estrangeira, Língua Segunda. Faculdade de Letras da Universidade de Coimbra].
- FERREIRA, T. S. (2019). *Aquisição/aprendizagem do sistema de atribuição de género nominal em PLNLM*. [Tese de doutoramento em Linguística do Português: Investigação e Ensino, Faculdade de Letras da Universidade de Coimbra].
- FRANCESCHINA, F. (2005). *Fossilized Second Language Grammars – the Acquisition of Grammatical Gender*. John Benjamins Publishing Company.
- GODINHO, A. P. (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. In M. J. Marçalo, M. Lima-Hernandes, E. Esteves, M. C. Fonseca, O. Gonçalves, A. L. Vilela e A. A. Silva (eds.), *Língua portuguesa: ultrapassar fronteiras juntar culturas*. Évora: Universidade de Évora.
- GRANGER, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, B. Altenberg & M. Johansson (eds.), *Languages in Contrast. Text-based cross-linguistic studies* (pp. 37-51). Lund: Lund University Press.
- GRANGER, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: a critical evaluation. In K. Aijmer (ed.), *Corpora and Language Teaching* (pp. 13-32). Amsterdam / Philadelphia: John Benjamins.
- HARRIS, J. W. (1991). The Exponence of Gender in Spanish, *Linguistic Inquiry*, 22(1), 2762.
- IBRAHIM, M. H. (1973). *Grammatical Gender*. Mouton.

- LACSÁN, V. (2016). *The acquisition of gender agreement in L2 Portuguese by adult Hungarian speakers*. [Dissertação de Mestrado, Faculdade de Letras da Universidade de Lisboa].
- LEIRIA, I. (2006). *Léxico, Aquisição e Ensino do Português Europeu língua não materna*. Fundação Calouste Gulbenkian / Fundação para a Ciência e Tecnologia.
- MARIOTTO, E. M. C. e LOURENÇO-GOMES, M. C. (2013). 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* (pp. 1278-1285). Goiânia / Goiás: Faculdade de Letras (UFG).
- MARTINS, C., FERREIRA, T., SITO, M., ABRANTES, C., JANSSEN, M., FERNANDES, A., SILVA, A., LOPES, I., PEREIRA, I. e SANTOS, J. (2019). *Corpus de Produções Escritas de Aprendentes de PL2 (PEAPL2): Subcorpus Português Língua Estrangeira*. CELGA-ILTEC.
- MARTINS, C. (2015). 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), 26-51.
- MARTINS, C. (2020). Estudos sobre a aquisição/aprendizagem do gênero nominal por aprendentes de português língua não materna: valências pedagógicas. *Quaderns de Filologia: Estudis Lingüistics XXV*, 169-184. URL: <https://doi.org/10.7203/QF.25.19078> (Acesso em 09.01.2023).
- MOTA, M. A. (2016). A categoria gramatical gênero, nos nomes e adjetivos do português: algumas reflexões. *Diadorim, Especial*, 150-164.
- MOTA, M. A. (2020). Morfologia do nome e do adjetivo. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. Mota, L. Segura e A. Mendes (eds.), *Gramática do Português* (vol. III, pp. 2833-2930). Lisboa: Fundação Calouste Gulbenkian.
- PINTO, J. (2017). A aquisição do gênero e da concordância de gênero em portuguesa língua terceira ou língua adicional. In Osório, P. (org.), *Teorias e Usos Linguísticos* (pp. 91-110). Lisboa: Lidel.
- RIO-TORTO, G. e RODRIGUES, A. S. (2016). Formação de nomes. In G. Rio-Torto e A. S. Rodrigues, I. Pereira, R. Pereira e S. Ribeiro, *Gramática Derivacional do Português*, (2.<sup>a</sup> edição), (pp. 135-240.). Coimbra: Imprensa da Universidade de Coimbra.

SANTOS, I. A., PEREIRA, I., MARTINS, C. S. P., LOPES, A.C.M., CARAPINHA, C., SILVA, A. (2016). CorpOral: PL2 – Um novo recurso para o estudo do português língua não materna. In A. Moreno, F. Silva e J. Veloso (eds.), *Textos Seleccionados do XXX Encontro Nacional da Associação Portuguesa de Linguística* (pp. 103-112). Associação Portuguesa de Linguística.

TÍTULO: A atribuição de valores de género nominal em produções orais de aprendentes tardios de português como língua não materna

RESUMO: Partindo da análise de desvios de atribuição e de concordância nominal em género, em produções orais, o presente estudo visa aferir padrões relativos à aquisição da categoria de género gramatical em português como língua não materna (PLNM) por aprendentes tardios, falantes nativos de espanhol e de inglês, a frequentar turmas dos níveis A2, B1 e B2. Mais especificamente, procurou-se averiguar se a configuração do conhecimento linguístico prévio condiciona o desempenho dos aprendentes no que respeita à associação de valores de género nominal em português. O trabalho empírico permitiu constatar que uma maior proximidade tipológica dos idiomas, língua materna e língua-alvo de aprendizagem, não se traduz num número reduzido de desvios, já que, em termos proporcionais, os índices de desvios nos dois segmentos da amostra é muito próximo, sobretudo no nível mais avançado. Por fim, as produções orais dos informantes de LM espanhola revelam, ainda, um certo efeito de estabilização na aprendizagem, visto que no nível B2 a proporção de desvios não é consideravelmente inferior à registada nos níveis precedentes, A2 e B1.

TITLE: The attribution of nominal gender in oral productions of late learners of Portuguese as a nonnative language

ABSTRACT: Based on the analysis of gender assignment and nominal gender agreement deviations in oral productions, the present study aims to assess patterns of acquisition of the grammatical gender category in Portuguese as a non-native language (PNNL) by late learners, native speakers of Spanish and English, attending classes at levels A2, B1 and B2. More specifically, we aim to investigate whether the configuration of prior linguistic knowledge affects learners' performance with respect to the association of nominal gender values in Portuguese. The empirical work showed that a greater typological proximity of the languages, mother tongue and learning target language, does not translate into a reduced number of deviations, since, in proportional terms, the deviation rates in the two segments are very close, especially at the most advanced level. Finally, the oral productions of the Spanish informants also reveal a certain stabilization effect in learning, since at level B2 the proportion of deviations is not considerably lower than that recorded at the previous levels, A2 and B1.

