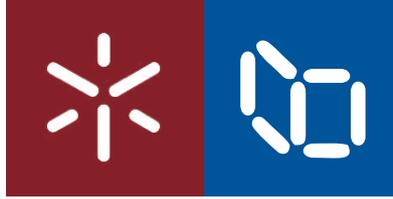


**Universidade do Minho**  
Instituto de Letras e Ciências Humanas

Liliana Manuela Osório Correia

**O efeito da exposição linguística sobre a  
competência lexical de crianças  
lusodescendentes residentes na Alemanha**





**Universidade do Minho**

Instituto de Letras e Ciências Humanas

Liliana Manuela Osório Correia

**O efeito da exposição linguística sobre a  
competência lexical de crianças  
lusodescendentes residentes na Alemanha**

Tese de Mestrado

Mestrado em Português Língua Não Materna (PLNM):  
Português Língua Estrangeira (PLE) e Português Língua  
Segunda (PL2)

Trabalho efetuado sob a orientação da

**Professora Doutora Cristina Maria Moreira  
Flores**

## DECLARAÇÃO

Nome: Liliana Manuela Osório Correia

Endereço eletrónico: lilianamocorreia@gmail.com

Número do Cartão de Cidadão: 12079255

Título da dissertação: **O efeito da exposição linguística sobre a competência lexical de crianças lusodescendentes residentes na Alemanha**

Orientador: Professora Doutora Cristina Maria Moreira Flores

Ano de conclusão: 2017

Designação do Mestrado: Mestrado em Português Língua Não Materna (PLNM): Português Língua Estrangeira (PLE) e Português Língua Segunda (PL2)

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

Universidade do Minho, \_\_\_/\_\_\_/\_\_\_\_\_

Assinatura: \_\_\_\_\_

## **Agradecimentos**

A realização deste projeto de investigação não teria sido possível sem o contributo de um conjunto de pessoas que, direta ou indiretamente, ajudaram a concretizá-lo.

Em primeiro lugar, gostaria de expressar a minha gratidão à comunidade portuguesa de Estugarda, que prontamente aceitou participar neste estudo e sem a qual o mesmo não se teria realizado.

Em segundo lugar, gostaria de agradecer à Professora Doutora Cristina Flores não só por ter, sem reservas, aceitado orientar esta dissertação, como também pelo seu apoio constante, sugestões de aperfeiçoamento e, sobretudo, pelo incentivo (e oportunidades) que me tem dado para continuar a realizar investigação dentro da área do bilinguismo.

Em terceiro lugar, gostaria também de agradecer à equipa de investigação do projeto EFFE que, de imediato, se disponibilizou a ceder os dados dos participantes monolíngues e a partilhar a sua experiência, bem como ao Diogo por ter acolhido este projeto na sua sala de aula, por ter distribuído/recolhido os questionários e por se ter prontificado a entrar em contacto com os pais das crianças bilingues sempre que surgia alguma dúvida.

E, naturalmente, não poderia terminar sem agradecer à minha família e aos meus amigos próximos, os quais, cada um a seu modo, também contribuíram para o desenvolvimento deste trabalho.



## Resumo

Nas últimas décadas têm surgido vários estudos sobre o desenvolvimento da competência lexical de crianças bilingues. Estes estudos demonstram que os falantes bilingues tendem a apresentar níveis de vocabulário inferiores aos seus pares monolingues quando apenas uma das suas línguas é avaliada (Cote e Bornstein, 2014). Devido à variedade de contextos em que a aquisição bilingue ocorre, estas investigações têm procurado simultaneamente identificar fatores intervenientes no processo de desenvolvimento da competência lexical bilingue. Neste âmbito, fatores como a quantidade e a qualidade da exposição linguística, entre outros, têm sido apresentados como variáveis preditivas do desenvolvimento lexical de crianças bilingues (Unsworth, 2016). Nesta linha de investigação, o presente estudo tem como um dos principais objetivos verificar se, no processo de aquisição linguística, falantes de herança do português europeu, entre os 6 e os 11 anos, residentes na Alemanha, têm um desempenho lexical idêntico ao de crianças monolingues do português europeu, no 1.º Ciclo do Ensino Básico. Procura-se, ainda, identificar fatores de *input* preditivos do desenvolvimento da competência lexical das crianças bilingues. Foram testadas 23 crianças bilingues e 21 crianças monolingues, com base numa tarefa de produção oral semi-espontânea, da qual se constituiu um *corpus*, composto por três *subcorpora* (categorias gramaticais nomes, verbos e adjetivos), que serviu de base às análises comparativas inter e intragrúpis. O perfil sociolinguístico das crianças bilingues foi delineado através de um questionário parental detalhado (baseado em Unsworth, 2013), do qual se extraíram dados sobre as diferentes fontes e quantidade de *input* linguístico. Análises estatísticas mostram que há diferenças significativas entre ambos os grupos relativamente ao *corpus* total e aos *subcorpora* nomes e verbos, mas não quanto aos adjetivos. Os resultados revelam ainda que há diferenças significativas dentro dos dois grupos na utilização de nomes, verbos e adjetivos, i.e. ambos utilizam mais nomes do que verbos e mais verbos do que adjetivos. Por sua vez, uma análise estatística centrada nas possíveis variáveis preditivas revela que a quantidade de *input* e de *output* em contexto doméstico influenciam significativamente a competência lexical destes falantes, isto é, quanto maior a quantidade de *input* e de *output*, mais elevado é, de facto, o seu nível de proficiência linguística. Verificou-se ainda que há uma correlação positiva significativa entre o número de progenitores falantes da língua minoritária e o desempenho lexical das crianças bilingues, ou seja, as crianças bilingues que recebem *input* em português europeu de ambos os pais apresentam melhores resultados lexicais.



## *Abstract*

The last decades have seen the emergence of numerous studies within the field of early bilingual lexical acquisition. These studies have shown that bilingual speakers tend to have smaller vocabulary sizes compared to their monolingual peers when only one of their languages is assessed (Cote and Bornstein, 2014). Due to the great variability of contexts in which bilingual language acquisition takes place, these studies have attempted to identify the factors that play a major role in the process of bilingual children's lexical development. In this line of investigation, factors such as the quantity and the quality of the input have been claimed to predict bilingual children's lexical development (Unsworth, 2016). The present study aims at investigating whether, in the process of language acquisition, heritage speakers of European Portuguese, between six and eleven years old, living in Germany, have a similar lexical development as their monolingual counterparts attending primary school. Moreover, it also aims at identifying input factors which may predict the lexical development of these bilingual children. Twenty-three bilingual children and 21 monolingual children were tested on a semi-spontaneous oral production task. The collected data were used to build a corpus, composed of three subcorpora (grammatical categories: nouns, verbs and adjectives), which served as a basis for the between- and within-group comparisons. The bilingual children's sociolinguistic profile was traced by means of a detailed parental questionnaire (following Unsworth, 2013), from which information about the different sources of input as well as its quantity was collected. A statistical analysis shows that there are significant differences between both groups regarding the total corpus and the subcorpora nouns and verbs, but no differences concerning the subcorpus adjectives. The results also show that there are significant differences in the production of nouns, verbs and adjectives, both within the bilingual group and within the monolingual one; i.e. both groups produce more nouns than verbs and more verbs than adjectives. On the other hand, a statistical analysis centred on the possible predictive variables reveals that bilingual children's lexical competence is significantly influenced by both the input and the output quantity at home; i.e. the larger the amount of input and output they receive and produce, the higher their level of linguistic proficiency. The analysis also shows a positive correlation between the number of parents who are speakers of the minority language and the lexical performance of the heritage speakers; that is, children who receive European Portuguese input from both parents have better lexical results.



# Índice

Agradecimentos .....	iii
Resumo .....	v
<i>Abstract</i> .....	vii
Lista de siglas e acrónimos .....	xi
Lista de tabelas .....	xiii
Lista de gráficos .....	xiii
Introdução .....	1
Capítulo 1 – Conceitos Preliminares.....	7
1.1 Língua Materna .....	7
1.2 Língua Segunda e Língua Estrangeira .....	9
1.3 Língua de Herança .....	10
Capítulo 2 – A aquisição da língua materna .....	15
2.1 Pressupostos teóricos da aquisição da linguagem .....	15
2.2 Estádios iniciais de desenvolvimento linguístico .....	19
2.3 Composição do léxico inicial .....	21
2.4 O papel do <i>input</i> em ambientes monolíngues .....	27
Capítulo 3 – A aquisição de duas (ou mais) línguas durante a infância .....	33
3.1 Diferenças e semelhanças entre a aquisição linguística monolíngue e a aquisição bilingue precoce .....	36
3.2 O papel do <i>input</i> em ambientes bilingues .....	44
3.2.1 A quantidade e a qualidade da exposição linguística .....	45
3.2.2 Sumário .....	53
Capítulo 4 – O presente estudo.....	57
4.1 Recolha de dados.....	57
4.1.1. Autorização.....	57

4.1.2. Tarefa de produção oral semi-espontânea .....	57
4.1.3. Constituição do <i>corpus</i> .....	59
4.1.4. Questionário sociolinguístico .....	60
4.2. Participantes.....	62
4.2.1. Falantes de herança .....	62
4.2.2. Falantes monolíngues .....	65
4.3. Questões de investigação.....	65
Capítulo 5 – Resultados .....	69
5.1 Tarefa de produção oral semi-espontânea .....	69
5.2 Relação entre os fatores de <i>input</i> e o desempenho lexical dos falantes de herança.....	74
Capítulo 6 – Discussão.....	79
Conclusão .....	89
Referências .....	95
Anexos.....	117
Anexo A – Autorização .....	119
Anexo B – Imagem-cenário <i>Sala</i> .....	121
Anexo C – Questionário sociolinguístico .....	123
Anexo D – Dados sociolinguísticos dos participantes bilíngues .....	129
Anexo E – Resultados dos testes estatísticos de normalidade e de homogeneidade.	131
Anexo F – Resultados das comparações intragrupois por método <i>pairwise</i> de <i>Bonferroni</i> .....	135

## Lista de siglas e acrónimos

ANOVA	Análise de Variância
CLUL	Centro de Linguística da Universidade de Lisboa
CPLP	Comunidade dos Países de Língua Portuguesa
DAL	Dispositivo de Aquisição da Linguagem
DP	Desvio-Padrão
EFFE	Escreves como falas – falas como escreves?
EPE	Ensino de Português no Estrangeiro
EUA	Estados Unidos da América
FH	Falante de Herança
FM	Falante Monolíngue
GU	Gramática Universal
IIQ	Intervalo Interquartilício
L1	Língua Primeira
L2	Língua Segunda
LE	Língua Estrangeira
LH	Língua de Herança
LM	Língua Materna
M	Média
PE	Português Europeu
PLH	Português Língua de Herança
SSAL	Sistema de Suporte à Aquisição da Linguagem
TEA	Tempo de Exposição Acumulada
UBiLEC	Utrecht Bilingual Language Exposure Calculator



## Lista de tabelas

- Tabela 1** Distribuição e variação lexical por grupo de falantes relativamente aos *subcorpora*.
- Tabela 2** Correlações entre os fatores de *input* e o desempenho lexical dos FH.
- Tabela D1** Dados sociolinguísticos dos participantes bilingues.
- Tabela E1** Resultados dos testes de normalidade Kolmogorov-Smirnov e Shapiro-Wilk referentes à distribuição dos dados do *corpus* total e dos *subcorpora* por grupo de participantes (com *outlier*).
- Tabela E2** Resultados dos testes de normalidade Kolmogorov-Smirnov e Shapiro-Wilk referentes à distribuição dos dados do *corpus* total e dos *subcorpora* por grupo de participantes (sem *outlier*).
- Tabela E3** Resultados dos testes de homogeneidade das variâncias dos dois grupos em comparação (com *outlier*).
- Tabela E4** Resultados dos testes de homogeneidade das variâncias dos dois grupos em comparação (sem *outlier*).
- Tabela E5** Resultados dos testes de normalidade Kolmogorov-Smirnov e Shapiro-Wilk referentes aos fatores de *input* (sem *outlier*).
- Tabela F1** Resultados das comparações por método *pairwise* de *Bonferroni* dentro do grupo de FH.
- Tabela F2** Resultados das comparações por método *pairwise* de *Bonferroni* dentro do grupo de FM.

## Lista de gráficos

- Gráfico 1** Quantidade de itens lexicais produzidos por grupo de falantes em cada categoria gramatical.
- Gráfico 2** Caixa de Bigodes: distribuição e variação lexical por grupo de falantes relativamente ao *corpus* total (nomes, verbos e adjetivos).
- Gráfico 3** Caixa de Bigodes: distribuição e variação lexical por grupo de falantes relativamente aos *subcorpora*.



## Introdução

Uma das áreas que tem sido investigada nas últimas décadas no campo da linguística, nomeadamente na área da aquisição de línguas, está relacionada com a *língua de herança* (LH), termo problematizado pela primeira vez nos anos setenta, no Canadá (Cummins, 2005), e que tem sido utilizado para designar uma língua, cultural e etnolinguisticamente minoritária, que se desenvolve em contextos bilíngues, nos quais outra língua, maioritária a nível sociopolítico, é falada (Montrul, 2016). Embora esteja também associado às línguas minoritárias faladas por comunidades não emigrantes (por exemplo, as minorias indígenas nos Estados Unidos da América (EUA)), o termo *língua de herança* é sobretudo usado para referir as línguas faladas pelas segundas e terceiras gerações de emigrantes (Flores e Barbosa, 2014; Montrul, 2016).

Portugal é um país com uma longa história de emigração, sendo muitos os filhos e netos de emigrantes portugueses de primeira geração que adquirem o português europeu (PE) como língua de herança. Embora seja difícil estabelecer o número exato de portugueses e lusodescendentes na diáspora, os dados do Observatório da Emigração (Pires *et al.*, 2016) revelam que, atualmente, mais de 20% da população portuguesa (i.e. mais de dois milhões de portugueses), nascida em Portugal, reside fora do país. De acordo com a mesma fonte, Portugal assume-se hoje, entre os países da União Europeia com mais de um milhão de habitantes, como o país com a maior proporção de emigrantes relativamente à população residente.

Quanto aos principais destinos da emigração portuguesa, até meados do século XX, os portugueses emigravam predominantemente para os continentes Americano e Africano. Contudo, o fim da Segunda Guerra Mundial e a necessidade de reconstrução dos países da Europa do Norte e Central apoiada pelo Programa de Recuperação Europeu, também conhecido por Plano Marshall, fizeram com que, a partir dos anos sessenta, a maior parte do fluxo migratório português se direcionasse para dentro do espaço europeu (Padilla e Ortiz, 2012). Motivados pelas precárias condições de vida no território nacional e atraídos pela recuperação económica que ocorria na Europa Central, os portugueses elegeram, entre 1950 e 1988, como principais destinos de emigração, países europeus como a França e a Alemanha (Baganha, 1994). Atualmente, o Reino Unido é o território europeu que mais emigrantes portugueses atrai, seguindo-se, por ordem decrescente, a França, a Suíça e a Alemanha. O número de emigrantes portugueses de primeira geração a residirem nos principais destinos emigratórios da segunda metade do século XX

continua a ser, ainda hoje, bastante representativo: em França, a comunidade portuguesa conta com mais de 600 mil emigrantes, nascidos em Portugal, e, na Alemanha, o seu número situa-se por volta dos 110 mil (Pires *et al.*, 2016).

Assim, ante a dimensão demográfica da comunidade emigrante portuguesa residente na Alemanha (e em outros países cuja língua de expressão não é a portuguesa) e dado que, juntamente com a população de um país, “emigra” também a sua cultura e a sua língua, torna-se fundamental procurar compreender como e em que circunstâncias milhares de crianças lusodescendentes adquirem a língua que é falada pelos seus pais e avós, mas que, contrariamente aos falantes monolíngues (FM) do país de origem, não é a língua de comunicação da sociedade maioritária em que nascem e crescem.

É, portanto, neste contexto de emigração que se enquadra, geoculturalmente, o objeto de estudo da presente dissertação – a competência lexical produtiva das crianças descendentes das primeiras e segundas gerações de emigrantes portugueses na Alemanha, no que à língua transmitida pelos seus pais e avós diz respeito, isto é, a sua língua de herança.

Dentro da área de aquisição de línguas, o presente estudo vem no seguimento das inúmeras investigações que têm procurado verificar se o desenvolvimento lexical de crianças bilingues simultâneas, expostas a duas línguas desde muito cedo na infância, apresenta padrões de desenvolvimento linguístico idênticos aos dos seus pares monolíngues, expostos a uma única língua desde o nascimento, ou se, pelo contrário, destes divergem (Cote e Bornstein, 2014; David e Wei, 2005; De Houwer *et al.*, 2014; Hoff *et al.*, 2012; Holowka *et al.*, 2002; Pearson *et al.*, 1993; Poulin-Dubois *et al.*, 2013; dos Santos e Kern, 2013; entre muitos outros).

Está comprovado que, quando ambas as suas línguas são tidas em consideração, as crianças bilingues percorrem os mesmos estádios de aquisição linguística que as correspondentes crianças monolíngues, em faixas etárias idênticas, não exibindo, portanto, nenhum atraso de desenvolvimento linguístico decorrente do facto de estarem a adquirir duas línguas em vez de uma (Holowka *et al.*, 2002; Pearson *et al.*, 1993; Pearson e Fernández, 1994; Petitto *et al.*, 2001). Contudo, os resultados de inúmeros estudos sobre o desenvolvimento lexical de crianças bilingues mostram que estas não desenvolvem, geralmente, ambas as suas línguas ao mesmo ritmo (Bialystok *et al.*, 2009; De Houwer, 2009; Pearson, 2009).

De facto, no domínio da aquisição lexical, vários estudos revelam que as crianças bilingues tendem a ter repertórios lexicais quantitativamente inferiores aos dos seus pares

monolingues quando apenas uma das suas línguas é avaliada (Bialystok *et al.*, 2010; Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; Hoff *et al.*, 2012; Pearson *et al.*, 1993; entre outros).

É comumente aceite que para a aquisição de uma língua, no geral, e do seu léxico, em particular, é necessário que uma criança receba *input* linguístico suficiente dessa mesma língua. Tendo em conta que as crianças que crescem expostas simultaneamente a mais do que uma língua têm de dividir o seu tempo entre duas (ou mais) línguas, isto significa que a quantidade de *input* linguístico que recebem de cada língua, para além de ser inferior à que os seus pares monolingues recebem, não está geralmente distribuída, de forma equitativa, entre as línguas em aquisição (Genesee, 2010; Montrul, 2008). Assim, muitos dos estudos sobre o desenvolvimento da competência lexical de crianças bilingues têm simultaneamente procurado avaliar o papel que a variação na quantidade (e qualidade) de exposição linguística desempenha na dimensão dos repertórios lexicais das mesmas (Gathercole, 2014; Montrul, 2016; Unsworth, 2016).

Os resultados de muitos destes estudos têm revelado que o desenvolvimento lexical das crianças bilingues está significativamente associado a fatores relacionados com a quantidade e a qualidade da exposição linguística a que estas crianças têm acesso. Neste âmbito, entre os fatores considerados preditivos da quantidade de *input* e, conseqüentemente, da competência lexical das crianças bilingues encontram-se: o número de línguas faladas em contexto doméstico, o número de progenitores falantes das línguas-alvo, bem como a (in)existência de irmãos mais velhos no agregado familiar e a quantidade de produção linguística das crianças bilingues (Bridges e Hoff, 2014; Cote e Bornstein, 2014; Gathercole e Thomas, 2009; Hoff *et al.*, 2014; Pearson *et al.*, 1997; Ribot *et al.*, 2017; entre outros). Por sua vez, variáveis como o contacto com diferentes fontes de *input* (mais próximas ou mais afastadas da língua padrão), a variedade de interlocutores falantes nativos das línguas-alvo e a diversidade de atividades de lazer em que as crianças bilingues participam (utilizando as línguas-alvo), entre outras, têm sido apresentadas como fatores de *input* qualitativos preditivos do desenvolvimento lexical bilingue (Patterson, 2002; Place e Hoff, 2011; Scheele *et al.*, 2010; entre outros).

Assim sendo, tendo em conta a conjuntura sociolinguística em que milhares de crianças bilingues lusodescendentes adquirem o PE, bem como a escassez de estudos realizados sobre o desenvolvimento lexical das mesmas na sua LH, o presente estudo procura, à semelhança da investigação desenvolvida na área de aquisição de línguas,

verificar se a competência lexical produtiva de crianças bilingues português-alemão apresenta, na LH, padrões de desenvolvimento lexical idênticos aos dos seus pares monolingues do PE ou se, pelo contrário, destes se afastam. Cumulativamente, a presente investigação procura ainda avaliar o papel que diferentes fatores relacionados com a quantidade e a qualidade da exposição linguística a que as crianças bilingues português-alemão têm acesso desempenham no processo de aquisição lexical da sua LH.

Deste modo, no Capítulo 1 – *Conceitos Preliminares*, problematizar-se-á um conjunto de conceitos-chave da área de aquisição de línguas, como os de *língua materna* (LM), *língua segunda* (L2), *língua estrangeira* (LE) e *língua de herança*, que se consideram essenciais para uma melhor leitura e compreensão da presente dissertação.

No Capítulo 2 – *A aquisição da língua materna*, começar-se-á pela abordagem das principais correntes teóricas que marcaram os estudos sobre a aquisição da linguagem nas últimas décadas, seguindo-se a apresentação dos resultados e conclusões da investigação realizada na área de aquisição de línguas maternas, no que se refere não só aos principais estádios de desenvolvimento linguístico de crianças monolingues, como também à composição do seu léxico inicial e à influência que a quantidade e a qualidade da exposição linguística exercem no desenvolvimento da competência lexical das mesmas.

No Capítulo 3 – *A aquisição de duas (ou mais) línguas durante a infância*, apresentar-se-ão, primeiramente, os resultados e conclusões da investigação centrada na identificação de diferenças e semelhanças entre a aquisição linguística monolíngue e a bilingue, seguindo-se uma parte dedicada aos estudos que têm investigado a influência que os fatores relacionados com a quantidade e a qualidade do *input* exercem sobre o desenvolvimento lexical de crianças inseridas em contextos sociolinguísticos bilingues.

No Capítulo 4 – *O presente estudo*, apresentar-se-ão detalhadamente os materiais e os procedimentos adotados na recolha de dados, bem como o perfil sociolinguístico, quer do grupo de crianças bilingues português-alemão, quer do grupo de controlo, constituído por crianças monolingues do PE, e as questões de investigação que orientam e fundamentam este estudo.

No Capítulo 5 – *Resultados*, apresentar-se-ão detalhadamente os resultados das análises estatísticas, a nível inter e intragrupal, dos dados obtidos de ambos os grupos de participantes, que permitiram não só avaliar o desenvolvimento da competência lexical produtiva das crianças bilingues na sua LH, comparando-a com a dos seus pares

monolíngues e entre si, como também aferir a influência que os diferentes fatores de *input* selecionados para este estudo têm no desenvolvimento da mesma.

No Capítulo 6 – *Discussão*, problematizar-se-ão os resultados obtidos do presente estudo, à luz das questões de investigação que o fundamentaram, apresentando-se simultaneamente algumas sugestões para futuras investigações.



## Capítulo 1 – Conceitos Preliminares

Na área de estudos da aquisição linguística, utilizam-se variados termos, como *língua materna* ou *primeira língua* (L1), *língua segunda*, *língua estrangeira* e, mais recentemente, *língua de herança*, para designar a(s) língua(s) falada(s) por um indivíduo, de acordo com as circunstâncias sociolinguísticas e político-culturais em que a sua aquisição/aprendizagem ocorre.

Este capítulo inicial procurará, então, caracterizar os diferentes conceitos supramencionados. Na primeira secção, problematizar-se-á o conceito de *língua materna*, na secção seguinte, abordar-se-ão, de uma perspetiva contrastiva, os termos *língua segunda* e *língua estrangeira* e, na última secção, apresentar-se-á o conceito de *língua de herança*.

### 1.1 Língua Materna

O termo *língua materna* é geralmente utilizado para designar a primeira língua adquirida por um indivíduo em contexto doméstico durante a infância (UNESCO, 2008), sendo esta frequentemente (mas nem sempre) a língua da sociedade em que este está inserido (Spinassé, 2006). Ao ser adquirida de forma natural e espontânea, durante os primeiros anos de vida, a LM é a língua nativa de um sujeito sobre a qual este vai construindo, intuitivamente, conhecimento linguístico relativamente à sua forma e uso (Xavier e Mateus, 1990). Ortega (2009) descreve língua(s) materna(s) como a(s) língua(s) que uma criança adquire de familiares próximos e cuidadores até por volta dos quatro anos de idade. Por sua vez, Leiria *et al.* (2006) definem LM como “aquela em que, aproximadamente até aos cinco anos de idade, a criança estabelece a sua primeira gramática, que depois vai reestruturando e desenvolvendo em direção à gramática dos adultos da comunidade linguística em que está inserida” (p.5). Contudo, apesar de ser comumente aceite que os alicerces de uma LM estão estabelecidos até por volta dos cinco anos de idade, sabe-se que nem todas as propriedades linguísticas de uma L1 são adquiridas durante a primeira infância (i.e. até aos cinco anos). Há, de facto, certas propriedades gramaticais, como a seleção do modo do conjuntivo no PE (Jesus, 2014) e no espanhol (Blake, 1983), ou o género gramatical no holandês (Blom *et al.*, 2008; De Houwer e Gillis, 1998) e no galês (Gathercole *et al.*, 2001; Thomas e Gathercole, 2007),

cujos processos de aquisição são mais morosos e que são, portanto, estabilizadas mais tardiamente.

Problematizando o conceito de língua materna, Pokorn (2005) apresenta quatro critérios que têm sido frequentemente adotados para a sua definição: (1) a *origem*, i.e. a(s) primeira(s) língua(s) que um indivíduo adquire; (2) a *competência*, i.e. a(s) língua(s) que este melhor domina; (3) a *função*, i.e. a(s) língua(s) que é mais usada no dia a dia; e (4) a *identificação*, i.e. a(s) língua(s) com a qual um indivíduo se identifica ou através da qual é identificado como falante nativo. No entanto, a autora defende que, apesar do seu uso extensivo, nenhum destes critérios é capaz de definir objetiva e completamente o conceito de LM.

Neste sentido, e atendendo ao conjunto de definições e critérios apresentados, o conceito de LM não representa um problema conceptual se for utilizado para definir a língua de indivíduos que nascem e crescem em contextos sociolinguísticos e político-culturais homogêneos. Contudo, se procurarmos definir o conceito de LM em contextos multilinguísticos e multiculturais, este poder-se-á complexificar.

De facto, em contextos de aquisição bilingue durante a infância, designadamente no caso específico de crianças bilingues simultâneas, isto é, que começam a ser expostas simultaneamente a duas (ou mais) línguas desde a nascença (ou até por volta dos três anos), um indivíduo pode adquirir não uma, mas duas (ou mais) línguas maternas (De Houwer, 1990, 1995, 2009; Genesee e Nicoladis, 2006; Klein, 1986; Meisel, 2006, 2011; Stern, 1991). Porém, não há consenso entre investigadores quanto ao critério cronológico que distingue crianças bilingues simultâneas de sequenciais (i.e. que começam a adquirir uma L2 numa fase posterior ao início de aquisição da L1). Se, por um lado, há autores, como De Houwer (1995), que propõem que, para uma criança adquirir duas línguas como L1, esta tem de iniciar o contacto com as duas línguas durante o primeiro mês de vida; por outro lado, autores como McLaughlin (1978) e Meisel (2008, 2011) sugerem que esse período cronológico poderá ser alargado até por volta dos três anos de idade. Na base desta divergência estão argumentos relacionados com a ideia de que certos domínios linguísticos, como os da fonologia e da morfossintaxe, parecem estar sujeitos a períodos críticos ou sensíveis de aquisição, após os quais um indivíduo já não será capaz de alcançar proficiência nativa nos mesmos (ver Hyltenstam e Abrahamsson (2003), Long (1990) e Meisel (2011) para uma discussão geral sobre a influência da idade do início de exposição a uma língua).

Deste modo, aceitando a existência de mais do que uma LM em falantes bilingues simultâneos, critérios como a *competência*, a *função* e/ou a *identificação* de Pokorn (2005) poderão não ser adequados para a definição de LM em contextos sociolinguísticos heterogêneos, uma vez que os falantes bilingues divergem não só quanto ao nível de proficiência que alcançam nas suas duas (ou mais) línguas, podendo estar mais próximo ou mais afastado dos padrões de desenvolvimento linguístico considerados nativos (De Houwer, 2011), como também quanto à frequência e aos contextos comunicacionais em que utilizam cada uma das suas línguas (Grosjean, 1982, 2010).

Assim sendo, embora aparentemente seja simples definir o conceito de LM, sobretudo em contextos de aquisição linguística monolingues, verifica-se que este pode assumir novos contornos em situações de bilinguismo ou de multilinguismo, não sendo possível aplicar linearmente todos os critérios que frequentemente são adotados para a sua definição.

## 1.2 Língua Segunda e Língua Estrangeira

Por oposição ao conceito de *língua materna*, surgem os de *língua segunda*, de *língua estrangeira* e, frequentemente, o de *língua de herança*. Nesta secção, abordar-se-ão os conceitos de L2 e de LE, sendo dedicada, de seguida, uma secção à problematização do conceito de LH.

Na área de investigação da aquisição de línguas segundas, o termo L2 tem sido genericamente utilizado para designar uma língua adicional que um indivíduo adquire/aprende após ter adquirido a sua língua materna (Ellis, 1985; Klein, 1986; Ortega, 2009; Stern, 1991). A ambiguidade decorrente desta definição de L2, baseada na ordem cronológica de aquisição, leva a que muitos autores distingam a aquisição adicional de uma língua em função do contexto em que esta ocorre e do estatuto que detém dentro de um território (ver discussão em Leiria (2004) e Stern (1991)).

Assim, apesar de muitos investigadores, principalmente os de formação generativista, defenderem que só há um modo de processamento e de assimilação de uma língua não nativa (Flores, 2013), tem-se estabelecido uma distinção entre uma língua não materna adquirida e utilizada dentro do país em que esta possui o estatuto de língua oficial (optando-se pelo termo *L2*) e uma língua não materna aprendida e utilizada fora do território em que é oficialmente reconhecida pelo Estado (designada de *LE*) (Leiria, 2004; Leiria *et al.*, 2006; Stern, 1991). Acresce ainda a estas definições o facto, nelas implícito,

de que uma L2 é adquirida em contexto naturalístico, ou seja, de imersão na comunidade linguística que é sua falante nativa, desempenhando, portanto, um papel institucional e social na comunidade de que é língua oficial, ao passo que a aprendizagem de uma LE ocorre, sobretudo, em contexto de exposição formal, não sendo esta a língua de socialização da comunidade linguística em que o seu falante aprendente está inserido (Ellis, 1994; Leiria, 2004).

Associados aos contextos de aquisição de uma língua adicional, surgem frequentemente os termos *aquisição* e *aprendizagem* (Ellis, 1985, 1994; Klein, 1986; Krashen, 1981, 1982; Madeira, 2008). De acordo com Ellis (1994), o termo *aquisição* tem sido utilizado para designar o processo subconsciente de aquisição de uma língua adicional em contexto naturalístico, enquanto que o termo *aprendizagem* é usado para referir o processo consciente de estudar uma língua não nativa. Extrapolando estes conceitos para a distinção entre L2 e LE, *aquisição* está relacionado com o processo de aquisição de uma L2 e *aprendizagem* com o de uma LE.

Contudo, apesar de vários investigadores procurarem diferenciar ambos os conceitos, fundamentando-se sobretudo nos critérios apresentados, não há, de facto, consenso na literatura quanto às realidades que os termos L2 e LE designam (Leiria, 2004). Tal como Ellis (1994) afirma, a aquisição de uma língua adicional “is a complex, multifaceted phenomenon and it is not surprising that it has come to mean different things to different people” (p. 15).

### **1.3 Língua de Herança**

O termo *língua de herança* (do inglês *heritage language*), que surgiu nos anos setenta no Canadá (Cummins, 2005), somente nas últimas décadas (e muito recentemente no contexto português) tem sido utilizado na área de aquisição de línguas para designar uma língua, cultural e etnolinguisticamente minoritária, que se desenvolve em contextos bilíngues, nos quais outra língua, maioritária a nível sociopolítico, é falada (Montrul, 2016). Para além de ser amplamente usado para referir as línguas faladas pelas segundas e terceiras gerações de emigrantes (Cummins, 2005; Flores, 2013; Flores e Barbosa, 2014; Flores e Melo-Pfeifer, 2014; Montrul, 2016; Valdés, 2001), o termo *língua de herança* tem sido igualmente associado às línguas minoritárias faladas por comunidades não emigrantes, como é o caso, por exemplo, das minorias indígenas nos EUA (Fishman, 2001, 2006; Montrul, 2016). Procurando estabelecer critérios que permitam classificar

uma língua como *de herança*, Rothman (2009) defende que esta tem de ser falada em casa, ou de estar, de alguma forma, fácil e imediatamente disponível à criança que a está a adquirir, sendo crucial que não seja a língua dominante da sociedade em que o seu falante está inserido.

Por sua vez, o termo *falante de herança* (*heritage speaker*) tem sido usado na literatura para designar este tipo de falantes bilingues precoces, que estão expostos, desde a infância, a duas (ou mais) línguas em contextos naturalísticos. Uma das definições de falante de herança (FH) mais utilizadas nas áreas de aquisição linguística bilingue e de ensino de LH, designadamente no que respeita ao contexto norte-americano, é a de Valdés (2000), que define FH como “a student who is raised in a home where a non-English language is spoken, who speaks or merely understands the heritage language, and who is to some degree bilingual in English and the heritage language” (p. 1). Por seu turno, Benmamoun *et al.* (2013) apresentam uma definição mais abrangente de FH. Para os autores, um FH é um bilingue precoce que cresce exposto à língua de herança (a sua L1) e, simultânea ou sequencialmente, à língua maioritária (L2) durante a primeira infância (i.e. até aos cinco anos) e para quem a língua maioritária se torna dominante por volta, ou após, o início da fase de escolarização. Enquanto a primeira definição, a de Valdés (2000), nos remete para o grau de proficiência linguística dos FH em cada uma das suas línguas, a última remete-nos não só para os conceitos de bilinguismo simultâneo e sequencial, como também para a alteração dos hábitos linguísticos destes falantes, decorrente do contacto mais intensivo com a língua da sociedade maioritária aquando da entrada das crianças para instituições escolares.

No que respeita à diferenciação entre falantes bilingues simultâneos e sequenciais, uma vez que a LH é adquirida desde o nascimento, sendo, portanto, uma L1, os FH são classificados como *bilingues simultâneos* ou *bilingues sequenciais* de acordo com a idade com que iniciam a sua exposição à língua maioritária (Flores *et al.*, 2016; Montrul, 2012, 2016; Pascual y Cabo e DeLaRosa-Prada, 2015; Rothman, 2009). Desta forma, se, por um lado, os FH são expostos simultaneamente às línguas minoritária e maioritária desde a nascença (ou até aos três anos – como foi discutido anteriormente), estamos perante o primeiro tipo de falantes bilingues (i.e. simultâneos); se, por outro lado, estes iniciam a exposição à língua maioritária após os alicerces da L1 estarem estabelecidos, então, estamos na presença do segundo tipo de falantes bilingues (i.e. sequenciais).

Uma das características dos falantes das LH, que está relacionada com a alteração dos seus hábitos linguísticos e com o nível de proficiência que estes podem alcançar nas

suas duas (ou mais línguas), é a sua exposição desequilibrada, quantitativa e qualitativamente, às línguas em aquisição.

No geral, os FH recebem *input* da LH de forma mais intensiva durante os primeiros anos de vida, porém, este padrão de exposição linguística rapidamente se altera por volta dos quatro/cinco anos de idade (Flores, 2015), com a entrada das crianças para instituições (pré-) escolares e com o conseqüente contacto massivo com a língua maioritária, que começa a exercer o seu papel dominante no processo de aquisição bilingue (Flores e Correia, 2016). Segundo Montrul (2008), a mudança nos hábitos linguísticos ocorre, geralmente, quando as crianças bilingues iniciam a escolarização na língua maioritária e sentem uma grande vontade de se integrarem na sociedade de acolhimento. A necessidade de integração na sociedade dominante é também um dos fatores, apontados por Rothman (2009), que está na origem da introdução gradual da língua maioritária no seio familiar.

Para além disso, a exposição linguística à LH pode ocorrer em contextos de comunicação muito restritos, estando, geralmente, limitada ao espaço familiar, a registos de língua pouco variados e a um número reduzido de interlocutores da língua-alvo (Flores, 2013, 2015; Flores e Barbosa, 2014; Montrul, 2012; Polinsky, 2015), cuja produção verbal pode, por sua vez, exibir efeitos, como o de erosão linguística, resultantes do contacto prolongado com a língua maioritária (Pascual y Cabo e Rothman, 2012). Polinsky (2015), por exemplo, descreve a exposição linguística à LH da seguinte forma:

“Very often the only exposure the heritage speaker has to his or her heritage language is through the speech of the same, small group of close relatives during childhood. This home speech is not representative of the speech of the entire native-speaking population, nor does it cover all the possible contexts in which a language can be used” (p. 10).

Em contrapartida, a exposição à língua maioritária ocorre em contextos sociolinguísticos multifacetados: é a língua de escolarização, dos *media* e de socialização fora do contexto doméstico (Flores, 2015), podendo também ser falada em casa em situações de bilinguismo simultâneo ou em lares em que a língua maioritária começa progressivamente a ser introduzida, principalmente quando há irmãos no agregado familiar (Rothman, 2009), uma vez que estes tendem a usar a língua maioritária para interagirem entre si (Flores e Melo-Pfeifer, 2014; Montrul, 2010).

De facto, ao ser adquirida em contextos muito restritos e, muitas vezes, sem exposição formal, o nível de proficiência dos falantes de LH varia muito de indivíduo para indivíduo, podendo estar mais próximo ou mais afastado dos padrões de desenvolvimento linguístico dos falantes monolíngues da língua-alvo (Benmamoun *et al.*, 2013; Montrul, 2016; Pascual y Cabo e Rothman, 2012; Pascual y Cabo e DeLaRosa-Prada, 2015; Polinsky, 2015; Polinsky e Kagan, 2007; Rothman, 2009; Valdés, 2001).

Procurando demonstrar a enorme variação existente entre os falantes bilíngues de herança quanto ao domínio das competências recetiva e produtiva na língua minoritária, Benmamoun *et al.* (2013) afirmam que, por um lado, pode haver falantes que detêm um domínio linguístico na LH equiparado ao de um falante nativo, ao nível da compreensão e da produção oral, bem como ao nível da leitura e da escrita, ao passo que, por outro lado, há falantes cujo saber linguístico na LH se limita ao domínio da competência recetiva. Os autores salientam ainda que o facto de os FH geralmente desenvolverem mais os domínios da compreensão e da produção oral do que os da expressão escrita e da leitura se deve principalmente à falta de escolarização na LH.

Desta forma, tendo em conta que, geralmente, os FH não alcançam níveis de proficiência linguística idênticos aos dos falantes monolíngues da LH, alguns autores, como Polinsky (2015)<sup>1</sup>, não consideram a maioria dos FH falantes nativos da língua-alvo. De facto, ao longo das últimas décadas, vários têm sido os termos associados aos falantes de LH que refletem não só a heterogeneidade de contextos sociolinguísticos em que ocorre a aquisição destas línguas, como também os resultados do desenvolvimento da sua competência linguística. Assim, *semi-speakers* (Dorian, 1981), *pseudo-bilinguals* (Baker e Jones, 1998), *incomplete acquirers* (Montrul, 2002), ou, inclusive, *speakers of “kitchen language* (Polinsky e Kagan, 2007), são algumas das designações, quase sempre depreciativas, atribuídas a este perfil de falante.

Refutando a ideia de que os FH não são falantes nativos da língua a que são expostos *desde a nascença em contexto naturalístico* (i.e. a LH/L1), Rothman e Treffers-Daller (2014) colocam, ironicamente, a seguinte questão: “if competence issues in adulthood lead us to deny them the label of native speaker of the heritage language and given that they are often child L2 learners of the majority language, then what are they native speakers of?” (p. 94). Ao argumentarem que não há uma, mas múltiplas bases de

---

<sup>1</sup> “Such differences distinguish heritage language from the language of monolinguals and the language of balanced bilinguals; based on this lack of full attainment (which can rise due to a number of reasons, from attrition to transfer), we cannot consider heritage language speakers to be native speakers” (Polinsky, 2015:17).

referência monolíngues que refletem a diversidade da expressão linguística, os autores sugerem que o estatuto de competência nativa (i.e. *nativeness*) pode e deve ser aplicado aos estados de conhecimento linguístico que divergem dos padrões monolíngues. Por sua vez, Pascual y Cabo e Rothman (2012) defendem que, por a LH ser adquirida em contextos sociolinguísticos bastante heterogêneos, o estado final de desenvolvimento linguístico dos seus falantes não pode ser interpretado como *aquisição incompleta* (de *incomplete acquisition*), como é frequentemente referido na literatura (Montrul, 2002, 2008; Polinsky, 2006; Polinsky e Kagan, 2007, entre outros). Segundo os autores, a aquisição das LH não é incompleta nem deficitária, simplesmente o processo de aquisição linguística dos FH se desenvolve de maneira diferente do dos falantes monolíngues no país de origem, devido a particularidades do *input* a que os FH têm acesso:

“HSs are exposed to input that has inevitably been affected to some degree by previous cross-generational attrition and/or other language contact consequences. Modifications in available input must contribute to HS competence differences. This, in turn, makes the complete/incomplete distinction rather weak if not unfounded” (p. 451).

É, portanto, esta multiplicidade de resultados individuais na LH, que é adquirida pelos seus falantes como L1, mas que apresenta, na grande maioria das vezes, um desenvolvimento linguístico diferente do da língua falada pelos seus pares monolíngues, que está na origem de inúmeros estudos que procuram compreender o processo de aquisição e de desenvolvimento linguístico deste grupo peculiar de falantes bilíngues (Au *et al.*, 2002; Flores e Barbosa, 2014; Flores *et al.*, 2016; Montrul, 2002, 2004; Montrul e Sánchez-Walker, 2013; O’Grady *et al.*, 2001; Pires e Rothman, 2009; Polinsky, 2005, 2006; Rinke e Flores, 2014; Rothman, 2007; entre muitos outros).

## Capítulo 2 – A aquisição da língua materna

A capacidade do ser humano para adquirir uma ou mais línguas maternas tem exercido fascínio ao longo dos tempos, originando variadas teorias sobre a aquisição da linguagem, que procuram desvendar e explicar este fenómeno comum a todos os seres humanos. O facto de todas as crianças saudáveis, independentemente da comunidade em que estão inseridas, serem capazes de se tornarem falantes nativas da(s) língua(s) a que são expostas durante a infância tem alimentado, desde o século passado, um intenso debate sobre a natureza do processo da aquisição linguística, frequentemente designado na literatura por debate *nature vs. nurture*. Se, por um lado, há teóricos que defendem o carácter inato do conhecimento linguístico (Chomsky, 1968, 1988; Wexler, 1998; Yang, 2002, entre muitos outros), por outro lado, há investigadores que encontram no meio em que a criança está inserida, ou na interação entre o meio e mecanismos cognitivos próprios do ser humano, os fundamentos para o processo de aquisição e desenvolvimento da linguagem (Bruner, 2007; Goldberg, 2006; O’Grady, 2008; Skinner, 1957; Tomasello, 2005, entre outros).

Sendo a linguagem humana um fenómeno complexo e multifacetado, esta somente pode ser compreendida se for observada sob diferentes perspetivas teóricas, não havendo, portanto, uma abordagem única capaz de explicar por si só o processo da aquisição linguística.

Deste modo, iniciar-se-á este capítulo pela apresentação de algumas ideias-chave dos pressupostos teóricos que marcaram os estudos sobre a aquisição da linguagem nas últimas décadas.

### 2.1 Pressupostos teóricos da aquisição da linguagem

Uma das correntes teóricas que procurou clarificar o processo de aquisição da linguagem é a behaviorista, da qual B. F. Skinner é um dos seus mais conhecidos precursores. Na sua obra *Verbal Behavior* (1957), severamente criticada por Noam Chomsky (1959), Skinner defendeu que a linguagem é um comportamento humano idêntico a tantos outros, sendo a sua aquisição o resultado da formação de um hábito que se adquire por imitação e repetição através do processo sequencial estímulo-resposta-reforço. Assim sendo, sob uma perspetiva behaviorista mais generalista no que respeita ao comportamento verbal, ante a presença de um determinado estímulo (i.e. objeto,

evento, produção verbal, etc.), a criança produz uma resposta que é reforçada positiva ou negativamente pela comunidade verbal. Se a resposta for acompanhada por um reforço positivo, a criança continuará a imitar esse comportamento até se constituir um hábito; se, por outro lado, a sua resposta for acompanhada por um reforço negativo (por exemplo, na presença de um desvio linguístico), então tal comportamento tornar-se-á menos frequente e extinguir-se-á. Deste modo, segundo os behavioristas, a faculdade da linguagem humana não advém de qualquer tipo de conhecimento inato, sendo o resultado da interação do ser humano com o ambiente que o rodeia e estando, portanto, dependente de estímulos externos.

Numa corrente oposta ao behaviorismo, surge a perspectiva generativista, ou inatista, da aquisição da linguagem preconizada por Noam Chomsky. De acordo com o autor (1968, 1981, 1986, 1988) e os seus seguidores, o ser humano está equipado geneticamente com um Dispositivo de Aquisição da Linguagem (DAL) que dispõe de uma Gramática Universal (GU), sendo esta ativada através do *input* linguístico da comunidade a que está exposto. A GU permite que uma criança adquira qualquer língua natural através de princípios biologicamente determinados, comuns a todas as línguas humanas, e da fixação de parâmetros concretos para cada língua, cujos valores são estabelecidos mediante o *input* linguístico da comunidade linguística a que está exposta. Deste modo, as crianças, quando nascem, estão geneticamente dotadas de um conhecimento gramatical abstrato que é ativado pela exposição a uma língua natural. Ao formular hipóteses de modo implícito sobre o que é permitido ou não na sua língua, a criança vai fixando parâmetros específicos da sua gramática materna, tornando a aquisição linguística um processo fácil, rápido, inconsciente e comum a todos os seres humanos, apesar das limitações, como a pobreza do estímulo, e das variações existentes nas diversas comunidades linguísticas.

Uma figura importante nos estudos sobre a aquisição linguística é a de Lenneberg e a sua conceção biológica da linguagem. Na sua obra *Biological Foundations of Language*, Lenneberg (1967) defende a existência de um período de maturação biológica após o qual o ser humano vê reduzida a sua capacidade para adquirir a língua materna com facilidade e sucesso. Segundo o autor, esse período crítico para a aquisição da linguagem inicia-se nos primeiros anos de vida da criança e termina por volta da puberdade quando o processo de lateralização cerebral das funções da linguagem se completa. Se, durante esse período, a criança for privada de exposição linguística, esta já não será capaz de adquirir a linguagem de forma natural e eficiente. Para formular a sua

*Hipótese do Período Crítico*, Lenneberg (1967) fundamentou-se em estudos de crianças cuja função da linguagem fora afetada por algum tipo de deficiência física ou mental, tendo verificado que, por exemplo, no caso de crianças com afasia, o grau de recuperação das funções da linguagem diminuía com o avançar da idade, podendo ser irreversível após o término do período crítico.

Outro contributo teórico que influenciou os estudos sobre a aquisição linguística foi o de Piaget e a sua conceção cognitivista da linguagem. Para Piaget, o processo de aquisição linguística depende do desenvolvimento cognitivo da criança, considerando o pensamento anterior à linguagem (Piaget, 1926). Desta forma, antes de desenvolver a linguagem, a criança tem que construir ativamente o seu próprio conhecimento do mundo através da sua interação com o meio (Piaget, 1967). Segundo Piaget e Inhelder (1993), a linguagem surge a partir do estágio sensório-motor com o desenvolvimento da função semiótica que “consiste em poder representar alguma coisa (um «significado» qualquer, objecto, acontecimento, esquema conceptual, etc.) por meio de um «significante» diferenciado” (p. 51) e que torna possível “a evocação representativa de objetos ou acontecimentos não percebidos actualmente” (p. 84). Acompanhando os diferentes estádios do desenvolvimento cognitivo, a aquisição linguística progride gradualmente até atingir o “nível das chamadas funções mentais superiores, de que o pensamento hipotético-dedutivo é o expoente máximo.” (Sim-Sim, 1998). Assim, na perspectiva piagetiana, a linguagem não aparece como um conhecimento independente e inato no cérebro humano, mas antes como um processo dependente do desenvolvimento intelectual de cada ser humano, que vai exercendo a sua própria ação sobre o meio.

De igual modo importante para a compreensão do processo de aquisição linguística foi o contributo de Jerome Bruner ao enfatizar o papel da interação natural entre a criança e o ambiente, designadamente os seus interlocutores, no desenvolvimento da linguagem. Segundo Bruner (2007), a criança dispõe de um conjunto próprio de capacidades de aprendizagem da linguagem, semelhante ao DAL de Noam Chomsky, que proporciona a aquisição linguística, sendo estas ativadas através das relações transacionais entre a criança e um adulto. São estas relações interativas, nas quais, inicialmente, o adulto tem um papel dominante, que fornecem um sistema de suporte à aquisição da linguagem (SSAL). Assim, “é a interação entre o DAL e o SSAL que torna possível à criança entrar na comunidade linguística – e, ao mesmo tempo, na cultura a que a linguagem dá acesso” (p. 15).

Outra corrente teórica que tem contado com a contribuição de muitos investigadores (Ellis, 1998; Elman *et al*, 1996; MacWhinney, 1998; O’Grady, 2008; entre outros) nos últimos anos é a emergentista. Segundo Montrul (2016) e Owens (2012), esta abordagem teórica define o fenómeno de aquisição linguística como sendo o resultado da interação de diferentes capacidades cognitivas entre si e com o ambiente linguístico. Para os emergentistas, as crianças não estão dotadas geneticamente de um conhecimento linguístico abstrato. A cognição, da qual a linguagem faz parte, é que é inata. Sob esta perspectiva, é do *input* e do ambiente linguístico a que estão expostas que as crianças, dotadas de “general learning principles and language processing mechanisms [...] extract inductive generalizations and statistical regularities” (Montrul, 2016:138), que lhes permitem adquirir e desenvolver a linguagem. Desta forma, a aquisição linguística depende da capacidade do ser humano em detetar implicitamente regularidades e padrões, não só a nível fonológico e morfológico, como também a nível sintático e semântico, no *input* a que tem acesso (Poll, 2011).

Como se verifica pelo conjunto de teorias apresentadas, o debate sobre a natureza da faculdade da linguagem tem contribuído positivamente para a compreensão do processo de aquisição linguística. Atualmente, a maioria dos investigadores da área reconhece que quer fatores biológicos, quer condicionantes ambientais, desempenham um papel importante no desenvolvimento da linguagem. De facto, embora o ser humano tenha uma capacidade natural para adquirir uma ou mais línguas, o seu desenvolvimento linguístico é significativamente condicionado pelo ambiente linguístico a que está exposto e pelas interações comunicacionais em que participa durante a infância. Independentemente do peso que cada corrente teórica atribui a fatores inatistas ou ambientalistas, na realidade, todas as crianças saudáveis adquirem, de forma automática e espontânea, a(s) língua(s) da(s) comunidade(s) linguística(s) a que estão expostas, apresentando, inclusivamente, trajetórias de desenvolvimento da linguagem idênticas.

Assim sendo, na secção seguinte, apresentar-se-ão os estádios iniciais de desenvolvimento linguístico característicos do processo de aquisição de uma língua materna, atribuindo-se um especial enfoque à aquisição e ao desenvolvimento da competência lexical.

## 2.2 Estádios iniciais de desenvolvimento linguístico

É amplamente reconhecido pela comunidade científica que crianças com um desenvolvimento normal da linguagem apresentam padrões universais no processo de aquisição linguística, independentemente da comunidade sociocultural e linguística em que estão inseridas ou da modalidade linguística (oral ou gestual) (Slobin, 2006).

Antes de proferirem a sua primeira palavra, por volta do primeiro ano de vida, todas as crianças passam por um período pré-linguístico, isto é, uma fase de preparação à etapa verbal que se segue (Ingram, 1989; Lima, 2000; Papalia *et al.*, 2009; Rigolet, 2000; Stoel-Gammon e Sosa, 2009). Durante o período pré-verbal, inicia-se o processo de interação comunicacional entre a criança e o ambiente que a rodeia, começando esta a produzir as suas primeiras vocalizações (i.e. choro, palreio e balbucio) e a desenvolver as capacidades de discriminação auditiva dos sons da fala (Sim-Sim, 1998). É também durante o primeiro ano de vida que as crianças, à medida que vão sendo expostas aos sons específicos da(s) sua(s) língua(s) nativa(s), perdem a sua capacidade natural para discriminar contrastes fonéticos universais (Best, 1995; Tsushima *et al.*, 1994; Werker e Tees, 1984).

A produção da primeira palavra, por volta dos 12 meses, marca a entrada da criança no período linguístico. Durante este período, que se prolonga por toda a vida dos falantes, as crianças passam por estádios idênticos de desenvolvimento linguístico: até aproximadamente aos 18 meses, estas produzem enunciados de apenas uma palavra; entre os 18 e os 24 meses, surgem as suas primeiras combinações de palavras; e, após este período (i.e. a partir dos 24 meses), começam a produzir frases simples e complexas (Ingram, 1989).

No que respeita à quantidade de itens lexicais produzidos e compreendidos pelas crianças nas fases de aquisição lexical inicial (i.e. no período pré-escolar) e tardia (i.e. a partir da fase de escolarização), existe uma assimetria constante entre a dimensão do léxico recetivo e produtivo, sendo que o primeiro é sempre superior ao último, inclusive durante a fase adulta (Sim-Sim, 1998).

Segundo Menyuk (1988), estudos sobre a dimensão do conhecimento lexical recetivo e produtivo de crianças nos primeiros anos de vida têm apresentado resultados consistentes. Relativamente à competência lexical recetiva, por volta dos 12 meses, as crianças compreendem cerca de dez palavras, aos 14 meses entendem aproximadamente 50 palavras e aos 17 meses apresentam um vocabulário recetivo de cerca de 100 palavras.

Já no que respeita à competência lexical produtiva, as crianças produzem cerca de dez palavras por volta dos 15 meses e 50 palavras aos 19 meses.

Após um período de aquisição de vocabulário a um ritmo lento, no qual as crianças adquirem entre uma a três palavras novas por semana (Barrett, 1997), verifica-se, por volta dos 18 meses, um aumento súbito no ritmo de aquisição do léxico produtivo da maioria das crianças, levando a que muitos autores designem este fenómeno como *explosão lexical* ou *de vocabulário* (Dromi, 1987; Li *et al.*, 2007; Nelson, 1973).

De acordo com Owens (2012), o vocabulário produtivo de uma criança de dois anos, que se situa entre os 200 e 300 itens lexicais, aumenta progressivamente até atingir o marco quantitativo das 2600 palavras aos seis anos, possuindo nesta faixa etária um vocabulário recetivo entre os 20000 e os 24000 itens lexicais. Segundo o autor, no final da infância (i.e. aos 12 anos), as crianças compreendem cerca de 50000 palavras e, no fim da adolescência, quando terminam o ensino secundário, os jovens adultos têm um vocabulário recetivo de aproximadamente 80000 palavras. Contudo, o seu léxico produtivo pode ser consideravelmente inferior.

Tendo em conta os valores apresentados, é importante salientar que, embora a maioria das crianças exiba padrões semelhantes nas trajetórias de crescimento lexical, há variações individuais no ritmo de aquisição de vocabulário, pelo que se devem perspetivar estes marcos cronológicos e quantitativos como valores estimativos das tendências gerais do desenvolvimento lexical. De facto, Castro e Gomes (2000) defendem que, embora a idade e a linguagem caminhem lado a lado nas fases iniciais de aquisição lexical, à medida que as crianças vão crescendo, já não é possível demarcar, de forma rígida, limites cronológicos entre os períodos de aquisição linguística.

Porém, independentemente das diferenças individuais que possam ocorrer no processo de aquisição lexical inicial, o período que antecede o início da escolarização, isto é, desde o nascimento até aos cinco anos, é marcado por uma rápida aquisição lexical e conceptual (Sim-Sim, 1998). Segundo Kuczaj (1999), durante este curto período de tempo, as crianças adquirem um quarto do vocabulário médio de um adulto, que pode rondar as 100000 palavras. Comparativamente a outros domínios linguísticos, o autor considera que a componente lexical é a que mais se desenvolve após a entrada para a escola, sugerindo que se, misteriosamente, o desenvolvimento linguístico de uma criança terminasse aos cinco anos, esta já dominaria grande parte dos sistemas morfosintático e fonológico da sua língua materna, mas o seu repertório lexical estaria ainda incompleto.

De facto, no final do período pré-escolar, as crianças já dominam um conjunto de processos cognitivos, linguísticos e sociais que lhes servirá de apoio ao desenvolvimento da competência lexical durante a fase de escolarização (Dockrell e Messer, 2004). Segundo Singleton (2016), com a entrada para escola e o desenvolvimento das competências de literacia, o leque de oportunidades para a aquisição lexical aumenta significativamente, quer por via ostensiva, quer por uso contextual. De acordo com Duarte *et al.* (2011):

“À medida que progridem na escolaridade, as necessidades de leitura e de compreensão oral que enfrentam exigem que as crianças e os jovens enriqueçam drasticamente o seu capital lexical não só com palavras de âmbito geral como, igualmente, com termos técnicos provenientes de várias áreas científicas e tecnológicas” (p. 20).

É também durante este período que ocorre um desenvolvimento mais acentuado ao nível da semântica, da pragmática e da metalinguagem (Owens, 2012). Segundo Menyuk (1988), apesar de já terem adquirido muitos itens lexicais e significados no final da primeira infância, as crianças ainda têm não só de adquirir os significados denotativos e conotativos para todos os itens lexicais que usam, como também de compreender as relações existentes entre orações marcadas por diferentes conectores, as relações hierárquicas e a linguagem figurativa.

Em suma, tendo em conta as fases de desenvolvimento linguístico apresentadas e os respetivos marcos lexicais alcançados pelas crianças, pode concluir-se, tal como Berman (2009) afirma, que “while becoming a *native* speaker is largely accomplished before school age, becoming a *proficient* speaker is a far more protracted process” (p. 347).

### **2.3 Composição do léxico inicial**

Na secção anterior, foi realçado que, no geral, as crianças tendem a seguir trajetórias de aquisição lexical idênticas, levando a que vários investigadores tenham estabelecido marcos cronológicos e quantitativos representativos do seu desenvolvimento lexical.

No que se refere ao conteúdo do vocabulário, diversos autores observam, igualmente, características comuns no repertório lexical das crianças, designadamente no período inicial de desenvolvimento linguístico.

As primeiras produções lexicais das crianças, durante o estágio de produção de palavras isoladas, têm um carácter holofrástico, ou seja, estas utilizam uma única unidade linguística para expressar a globalidade da sua intenção comunicativa (Barrett, 1982). De acordo com Tomasello (2006), as holófrases são utilizadas por crianças de todas as línguas do mundo para:

- “request or indicate the existence of objects (e.g., by naming them with a requestive or neutral intonation);
- request or describe the recurrence of objects or events (e.g., *More, Again, Another*);
- request or describe dynamic events involving objects (e.g., as described by *Up, Down, On, Off, In, Out, Open, Close*);
- request or describe the actions of people (e.g., *Eat, Kick, Ride, Draw*);
- comment on the location of objects and people (e.g., *Here, Outside*);
- ask some basic questions (e.g., *Whats-that?* or *Where-go?*);
- attribute a property to an object (e.g., *Pretty* or *Wet*); and
- use performatives to mark specific social events and situations (e.g., *Hi, Bye, Thank You, and No*)” (p. 261).

Estando relacionadas com o seu contexto vivencial, as primeiras dez produções lexicais mais frequentes das crianças referem-se geralmente a animais, alimentos e brinquedos (Owens, 2012). Quando atingem o marco das 50 palavras, por volta dos 18 meses, e começam a surgir as primeiras combinações de palavras, o repertório lexical produtivo das crianças é constituído maioritariamente (i.e. cerca de 65%) por nomes de pessoas, animais, peças de vestuário, brinquedos, meios de transporte e bens alimentares. Seguem, em menor número (i.e. aproximadamente 20%), itens lexicais que designam ação ou movimento e, em último lugar, não só palavras referentes a atributos, rotinas e partes do corpo, como também as que expressam recorrência (Sim-Sim, 1998).

Num estudo pioneiro sobre a aquisição lexical em fases iniciais de desenvolvimento linguístico, Nelson (1973) investigou o conteúdo e a distribuição lexical das primeiras 50 palavras produzidas por 18 crianças falantes monolíngues do inglês. Tendo agrupado o vocabulário dos dados recolhidos nas categorias lexicais *specific nominals, general nominals, action words, modifiers, personal-social words* e *function words*, a autora concluiu que cerca de dois terços do vocabulário adquirido pertencia a categorias nominais, sendo que os nomes gerais de objetos (31%) e os nomes específicos para pessoas (12%) constituíam as subcategorias lexicais mais frequentes.

Com uma amostra mais ampla de crianças falantes de inglês, Bates *et al.* (1994) analisam a distribuição e o desenvolvimento lexical de 1803 crianças, entre os oito e os 30 meses, agrupadas por níveis de dimensão de vocabulário. Os autores verificam que a composição do repertório lexical destes falantes apresenta padrões de crescimento distintos para as categorias nomes, verbos e adjetivos, bem como para itens lexicais pertencentes às classes fechadas de palavras, em função do tamanho lexical total. Assim, numa fase inicial, até ao marco das 200 palavras, os nomes comuns crescem a um ritmo acelerado, ocupando uma média de 55,2% da dimensão lexical total de crianças com um repertório entre as 101 e as 200 palavras. Após esta fronteira, o ritmo de crescimento nominal estabiliza. Por sua vez, a aquisição de predicados (i.e. os verbos e adjetivos) apresenta, inicialmente, um ritmo de crescimento lento, mas estável, constituindo 7,6% do vocabulário produtivo no período entre uma e 50 palavras e atingindo os 25,2% no último escalão lexical, entre as 601 e as 680 palavras. Já os itens lexicais pertencentes às classes fechadas de palavras revelam um comportamento diferente, representando uma proporção pequena e inalterável no vocabulário total de crianças com um repertório lexical de uma a 400 palavras, sendo que, após este marco quantitativo, o ritmo de aquisição das palavras funcionais acelera significativamente em relação às outras categorias analisadas. Os autores sugerem que, para a aquisição de itens lexicais das classes fechadas de palavras, as crianças necessitam de possuir uma certa “massa crítica” (de *critical mass*) de nomes e de verbos no seu repertório lexical. Em suma, os resultados deste estudo parecem sustentar a ideia de que as mudanças observadas na composição do vocabulário de crianças com idades compreendidas entre os oito e os 30 meses refletem “a shift in emphasis from reference, to predication, to grammar” (p. 116).

Uma das primeiras propostas teóricas sobre o pressuposto da *tendência nominal* (de *noun/nominal bias*) nos primeiros estádios da aquisição linguística foi a elaborada por Gentner (1982). Tendo analisado comparativamente a frequência de utilização de nomes e de verbos em seis línguas (inglês, japonês, alemão, kaluli, mandarim e turco), a autora defende que a inclinação verificada para o uso predominante de nomes em relação a outros itens lexicais, designadamente os verbos, é universal e que fatores de ordem semântica e perceptual fundamentam esta tendência. De acordo com a autora, o facto de os verbos, em comparação com os nomes concretos, possuírem estruturas semânticas mais complexas e de terem uma relação menos transparente relativamente ao mundo perceptual, leva a que seja mais fácil para uma criança identificar conceitos codificados por nomes do que por verbos.

Este uso predominante de nomes em detrimento de verbos e de outras categorias gramaticais, em fases iniciais de desenvolvimento linguístico, tem originado vários estudos que procuram verificar a hipótese da existência de uma *tendência nominal*, de carácter universal, no processo de aquisição lexical. Para além dos estudos desenvolvidos com crianças falantes do inglês (Bates *et al.*, 1994; Nelson, 1973), outras investigações têm documentado uma maior proporção de nomes, relativamente a outras categorias gramaticais, no repertório lexical inicial de crianças falantes de diversas línguas como: o italiano (Caselli *et al.*, 1995; D’Odorico e Fasolo, 2007), o espanhol (Jackson-Maldonado *et al.*, 1993), o hebreu (Maital *et al.*, 2000), o francês (Bassano, 2000), o holandês (Schlichting, 1996), o navajo (Gentner e Boroditsky, 2009), o português do Brasil (Scherer e Souza, 2011), entre outras. No entanto, estudos realizados com crianças falantes de línguas asiáticas, como o coreano (Au *et al.*, 1994; Choi e Gopnik, 1995; Kim *et al.*, 2000) e o mandarim (Gelman e Tardif, 1998; Tardif, 1996; Tardif *et al.*, 1997), apresentam resultados contraditórios.

Tendo analisado, comparativa e longitudinalmente, o desenvolvimento lexical de crianças falantes do inglês e do coreano, Choi e Gopnik (1995) observam que as crianças coreanas apresentam proporções semelhantes de nomes e de verbos no seu repertório lexical por volta do marco das 50 palavras, sendo que, contrariamente às crianças monolíngues de inglês, sete das nove crianças coreanas investigadas apresentam uma explosão verbal por volta dos 19 meses e que, para seis destas crianças, a explosão verbal ocorre antes da nominal. Contudo, estudos realizados por Au *et al.* (1994) e Kim *et al.* (2000) revelam que a tendência nominal, reportada em estádios iniciais de aquisição lexical de crianças falantes de diversas línguas, também ocorre em crianças falantes do coreano. Kim *et al.* (2000) verificam ainda que, embora a categoria dos nomes seja dominante no léxico inicial de crianças falantes do inglês e do coreano, o repertório lexical das últimas apresenta uma proporção maior de verbos do que o das crianças falantes do inglês. Por sua vez, em estudos realizados com crianças monolíngues do mandarim, Tardif (1996) e Tardif *et al.* (1997) verificam que estas produzem mais verbos do que nomes no seu discurso espontâneo, não apresentando indícios da existência de uma tendência nominal no processo de aquisição lexical inicial.

A presença de um maior número de verbos no léxico inicial de crianças falantes do coreano e do mandarim, de acordo com alguns estudos, tem sido associada não só a características do *input*, designadamente à saliência estrutural e perceptual dos verbos

nestas línguas<sup>23</sup> e à frequência com que as diferentes categorias gramaticais ocorrem no discurso direcionado às crianças (Choi, 2000; Kim *et al.*, 2000; Tardif *et al.*, 1997), bem como a aspetos pragmáticos da interação comunicacional adulto-criança (Choi, 2000; Choi e Gopnik, 1995; Gopnik *et al.*, 1996; Kim *et al.*, 2000; Tardif, 1996).

De facto, ao analisar o discurso de mães de crianças monolíngues de inglês e de coreano, Choi (2000) verifica que as mães das crianças falantes de inglês produzem mais nomes do que verbos no discurso dirigido aos seus filhos ao passo que as mães das crianças falantes do coreano produzem a mesma proporção de nomes e de verbos, sendo que, nos enunciados destas últimas, os verbos ocorrem mais frequentemente em posição final, ocupando, portanto, uma posição saliente no discurso. Ao nível da pragmática, a autora constata ainda que as mães falantes do inglês produzem enunciados mais direcionados para objetos do que para ações, enquanto que as mães falantes de coreano enfatizam equitativamente objetos e ações nos seus enunciados.

Por sua vez, Tardif *et al.* (1997), num estudo comparativo entre falantes de mandarim, de inglês e de italiano, verificam que os cuidadores de crianças falantes de mandarim utilizam com mais frequência verbos no discurso dirigido às crianças do que os de crianças falantes de inglês e de italiano. Para além disso, as autoras observam que, contrariamente a estes últimos grupos de falantes, os cuidadores de crianças chinesas tendem a colocar os verbos em posição final nos enunciados, ou seja, numa posição discursiva mais saliente.

Tendo em conta que os estudos sobre a aquisição e distribuição lexical analisam, geralmente, apenas uma língua (Bates *et al.*, 1994; Bassano, 2000; Gentner e Boroditsky, 2009; Goldfield, 2000; Scherer e Souza, 2011; entre outros), ou um par linguístico (Caselli *et al.*, 1995; Choi e Gopnik, 1995; Jackson-Maldonado *et al.*, 1993; Maital *et al.*, 2000; entre outros), e que estes variam frequentemente quanto à metodologia implementada (relatórios parentais, inventários lexicais, etc.), Bornstein *et al.* (2004) realizaram um estudo multilinguístico, no qual avaliam e investigam a composição do

---

<sup>2</sup> De acordo com Chang-Song e Pae (2009), "...Korean is a SOV language in which verbs are more likely to occur in the perceptually salient final position of an utterance. In addition, Korean allows for nominal ellipsis, so that the subject of a sentence can be left out. As a result, verbs are perceptually and structurally more salient in Korean language input to children" (p. 63).

<sup>3</sup> Segundo Gentner e Boroditsky (2009), o mandarim, tal como o inglês, tem a ordem SVO como canónica. No entanto, o facto de o mandarim ser uma língua que permite a omissão do sujeito em determinados contextos torna possível a ocorrência de enunciados que começam por verbos (VO) ou que são unicamente compostos por estes, fazendo com que esta categoria gramatical ocupe uma posição estruturalmente mais saliente no mandarim do que na língua inglesa, na qual os verbos ocupam sempre uma posição intermédia nos enunciados.

vocabulário produtivo de crianças monolíngues de diversas comunidades linguísticas (espanhola, holandesa, francesa, hebraica, italiana, coreana e norte-americana), utilizando uma metodologia uniforme. A análise da distribuição lexical incidiu nas categorias nomes, verbos, adjetivos e em itens lexicais pertencentes às classes fechadas de palavras. Tendo dividido a amostra por níveis de dimensão lexical aos 20 meses, os autores verificam que, com a exceção das crianças que detêm um repertório de um a 50 itens lexicais, a produção de nomes é sempre superior à das restantes categorias gramaticais, sendo que os participantes produzem mais nomes do que verbos, mais verbos do que adjetivos e mais adjetivos do que itens das classes fechadas de palavras. No caso específico das crianças com uma dimensão lexical de uma a 50 palavras, estas produzem mais nomes do que adjetivos ou itens da classe fechada de palavras, mas não se verifica uma diferença significativa entre a produção de nomes e a de verbos. Este resultado vai ao encontro do estudo realizado por Choi e Gopnik (1995), no qual crianças falantes do coreano, com uma dimensão lexical de cerca de 50 palavras, produzem nomes e verbos em quantidades idênticas. No entanto, o estudo de Bornstein *et al.* (2004) sugere que, após a fronteira das 50 palavras e a explosão de vocabulário, a classe dos nomes assume um papel predominante no repertório lexical inicial das crianças. Para além disso, resultados de correlações estatísticas entre as classes de palavras investigadas sugerem que, na maioria dos níveis de dimensão lexical, nomes, verbos e adjetivos se desenvolvem em paralelo, ou seja, o crescimento lexical numa destas categorias gramaticais está relacionado com o crescimento das outras classes de palavras.

Tendo investigado a composição do léxico de crianças falantes de línguas que variam quanto à sua constituição morfológica, à saliência estrutural e perceptual das diferentes categorias gramaticais, bem como no que respeita à frequência com que estas ocorrem no *output* direcionado às crianças e aos aspetos pragmáticos do mesmo, Bornstein e colegas (2004) sugerem que a prevalência de nomes no repertório lexical das crianças observadas advém de restrições (de *constraints*) que, teoricamente, são transversais a todas as línguas. De acordo com os pressupostos teóricos baseados em restrições, a aquisição lexical é simplificada pela aplicação de princípios lexicais que favorecem o mapeamento de nomes, ou seja:

“a child hearing a new word is biased to believe that the word is a conventional word in the language, that it is the name for an object, that the word names the whole object (and not just part of it, or a quality of it, or something that it does), that the same word also

names similar objects from the same category, that objects in different categories have different names, and that a new word is the name for a category of objects that does not already have a name” (p. 1119).

Deste modo, os autores defendem que esta perspectiva teórica prevê o predomínio de nomes relativamente às restantes categorias gramaticais e, conseqüentemente, fundamenta os resultados do estudo realizado, não desvalorizando, contudo, as justificações apresentadas em outros estudos sobre a composição do repertório lexical de crianças em estádios iniciais de desenvolvimento linguístico.

Apesar de as crianças exibirem trajetórias de aquisição e de desenvolvimento lexical idênticas, os estudos apresentados revelam que características intra e extralinguísticas do *input* a que estas têm acesso ditam a composição do seu léxico inicial, levando a que exista variação na ordem e no ritmo de aquisição das diferentes categorias gramaticais, não só dentro, como também entre as línguas investigadas. Para além de o *input* linguístico desempenhar um papel relevante na composição lexical infantil, vários estudos têm reportado a sua influência no desenvolvimento da competência lexical de crianças monolíngues (Harris *et al.*, 1988; Hart, 1991; Hart e Risley, 1995; Hoff-Ginsberg, 1998; Hoff e Naigles, 2002; Huttenlocher *et al.*, 1991; Huttenlocher *et al.*, 2010; Pan *et al.*, 2005; Pancsofar e Vernon-Feagans, 2006; Weizman e Snow, 2001). Desta forma, abordar-se-á, na secção seguinte, o papel que fatores de *input*, como a quantidade e a variedade e qualidade, desempenham na aquisição e no desenvolvimento lexical monolíngue.

#### **2.4 O papel do *input* em ambientes monolíngues**

Embora a aquisição de uma língua nativa seja bastante uniforme, a quantidade e a qualidade do *input* a que uma criança monolíngue está exposta é muito variável, estando na origem de variações individuais no ritmo de aquisição dos diferentes domínios linguísticos. De facto, nas diferentes fases de desenvolvimento linguístico, as crianças diferem entre si no que respeita não só à dimensão de vocabulário que dominam e à complexidade das estruturas que produzem, como também à mestria da competência comunicativa (Hoff, 2006). No domínio específico da aquisição lexical, as diferenças individuais observadas nos diversos estádios de desenvolvimento linguístico têm sido associadas à variação existente na quantidade e na qualidade da exposição linguística em

fases iniciais de aquisição lexical (i.e. pré-escolar), a qual, por sua vez, tem sido considerada uma variável preditiva do desenvolvimento lexical em estádios de aquisição lexical tardios (Hart e Risley, 1995; Weizman e Snow, 2001). Sendo que, na generalidade, todas as crianças monolíngues entram no período linguístico por volta do primeiro ano de vida (Ingram, 1989), o contexto familiar é o ambiente, por excelência, em que a aquisição linguística ocorre nos seus primeiros anos de vida, constituindo, deste modo, a principal fonte de *input* a que as crianças têm acesso.

Estudos sobre a aquisição lexical de crianças em estádios iniciais de desenvolvimento linguístico têm revelado que existe uma correlação significativa entre as primeiras palavras produzidas pelas crianças e a frequência com que estas ocorrem no *input* parental que recebem (Harris *et al.*, 1988; Hart, 1991; Huttenlocher *et al.*, 1991). De facto, ao analisarem a relação entre a quantidade de *input* parental e o ritmo de crescimento lexical de 22 crianças entre os 14 e os 26 meses, Huttenlocher *et al.* (1991) verificam não só que as diferenças individuais no processo de aquisição lexical estão associadas à variação na quantidade de exposição linguística, como também que a frequência com que diferentes itens lexicais ocorrem no discurso direcionado às crianças está significativamente correlacionada com a sua ordem de aquisição.

Num estudo longitudinal com 42 famílias norte-americanas, Hart e Risley (1995) evidenciam a influência que a variação existente na quantidade e na qualidade da exposição linguística em contexto doméstico exerce no desenvolvimento lexical de crianças nos primeiros três anos de vida e as assimetrias que provoca na competência linguística das mesmas durante a fase de escolarização. Ao recolherem dados das interações comunicativas entre crianças e adultos de famílias de diferentes classes socioeconómicas, durante uma hora por mês, num período de tempo superior a dois anos, os autores verificam a existência de correlações significativas entre a quantidade e a qualidade do *input* que as crianças recebem, bem como entre estas (i.e. quantidade e qualidade do *input*) e a dimensão e riqueza lexical das crianças observadas. Deste modo, se, por um lado, uma maior quantidade de exposição linguística aumenta o ritmo de aquisição lexical e, conseqüentemente, a dimensão do repertório lexical, por outro lado, a exposição a um maior número de itens lexicais origina um aumento de situações comunicativas em que podem ocorrer produções verbais mais ricas qualitativamente (i.e. lexical e morfossintaticamente mais complexas e diversificadas), resultando, assim, num repertório lexical mais variado. Para além disso, Hart e Risley observam que o estatuto socioeconómico é o fator familiar mais fortemente associado às diferenças na quantidade

e na qualidade de exposição linguística que as crianças recebem no seu dia a dia. Assim, no que respeita à quantidade de *input*, os resultados do estudo revelam que as crianças de famílias pertencentes à classe alta (i.e. *professional*) são expostas a uma média de 2153 palavras por hora, ao passo que as crianças da classe média-baixa (i.e. *working-class*) são expostas a cerca de 1251 palavras e as da classe baixa (i.e. *on welfare*) a 616 palavras. Por sua vez, relativamente à qualidade da exposição linguística, os dados mostram que, comparativamente às outras classes socioeconómicas (sobretudo à mais desfavorecida), as famílias da classe alta expõem os seus filhos não só a uma quantidade maior e mais diversificada de itens lexicais, como também a uma quantidade superior de frases complexas, de tempos verbais variados, de frases declarativas e de interrogativas. Com base na consistência dos valores longitudinalmente observados, os autores estimam que, aos três anos de idade, as crianças da classe socioeconómica mais elevada terão sido expostas a mais de 30 milhões de palavras, enquanto que as crianças da classe média-baixa terão sido expostas a cerca de 20 milhões e as da classe baixa a apenas 10 milhões. Apesar desta diferença de cerca de 20 milhões de palavras entre os extremos sociais, os autores salientam que, aos três anos de idade, todas as crianças tinham aprendido a falar e eram falantes competentes da sua língua nativa. Contudo, quando a competência linguística de 29 destas 42 crianças foi avaliada no terceiro ano de escolaridade, verificou-se que esta estava significativamente associada ao ritmo de aquisição e à produção lexical das mesmas aos três anos de idade, mostrando que a variação no grau de exposição linguística em fases iniciais de aquisição lexical é um fator preditivo da competência lexical em estádios subsequentes de desenvolvimento linguístico.

Por sua vez, Hoff e Naigles (2002) analisaram a relação entre propriedades do discurso materno e o vocabulário produtivo de 63 crianças por volta dos 24 meses. As autoras verificam que determinadas propriedades do *input*, como a quantidade de palavras produzidas (i.e. *tokens*), a riqueza lexical do vocabulário produzido (i.e. *types*) e a extensão média dos enunciados (i.e. *mean length of utterance*), estão positivamente correlacionadas com a dimensão dos repertórios lexicais das crianças investigadas. Já a quantidade de enunciados produzidos pelas mães, outra propriedade analisada, não apresenta correlações significativas com o vocabulário das crianças. Os resultados deste estudo mostram ainda que existem correlações significativas entre as três propriedades da exposição linguística preditivas do desenvolvimento lexical: a quantidade de itens lexicais produzidos está significativamente correlacionada com a quantidade de vocabulário diversificado, ou seja, tal como no estudo de Hart e Risley (1995), os

cuidadores que produzem uma maior quantidade de itens lexicais também produzem vocabulário mais variado; e a utilização de itens lexicais mais diversificados está, por sua vez, associada à produção de enunciados mais longos e, portanto, sintaticamente mais complexos.

Em contexto familiar, a presença de irmãos também parece influenciar a quantidade e a qualidade do *input* linguístico a que as crianças estão expostas e, conseqüentemente, o seu desenvolvimento lexical. Segundo Hoff (2006), a experiência sociolinguística das crianças primogénitas difere da das que nascem posteriormente. Sendo as únicas crianças no seio familiar, as primogénitas têm mais oportunidades de participar em interações comunicativas com adultos e de receber *input* exclusivamente dirigido a elas do que as crianças que já têm irmãos. Procurando investigar se o *input* maternal em interações diádicas (i.e. mãe e um filho) difere do de interações triádicas (i.e. mãe e dois filhos), Jones e Adamson (1987) constatam que, na presença de um irmão, a quantidade de exposição linguística que as crianças mais jovens recebem da mãe é inferior à quantidade de *input* a que estão expostas quando se encontram sozinhas com a mesma. Para além disso, o *input* proveniente dos irmãos parece não compensar quantitativamente essa diferença, visto que mais de 90% do seu discurso é dirigido à figura materna. As autoras verificam ainda que, em interações diádicas, as mães das crianças primogénitas colocam mais questões, procuram obter mais respostas e apoiam mais as tentativas de produção oral dos seus filhos do que as mães das crianças não primogénitas na mesma faixa etária (entre os 18 e os 23 meses). Por outro lado, embora o *input* proveniente de irmãos, em comparação com o de adultos, seja estruturalmente menos complexo e em menor quantidade (Hoff-Ginsberg e Krueger, 1991), a interação verbal com outras crianças parece favorecer o desenvolvimento da competência comunicativa das crianças mais novas (Hoff, 2006). Assim, se, por um lado, os primeiros filhos parecem estar em vantagem no que respeita ao desenvolvimento lexical e sintático em estádios iniciais de aquisição linguística, por outro lado, os filhos que se seguem parecem desenvolver melhor as competências conversacionais (Hoff-Ginsberg, 1998; Hoff, 2006).

No decorrer deste capítulo, após a abordagem de algumas das principais correntes teóricas que marcaram os estudos sobre a aquisição da linguagem desde o século XX, procedeu-se à apresentação dos resultados e conclusões da investigação realizada na área de aquisição de línguas maternas, não só no que respeita aos principais estádios de desenvolvimento linguístico que as crianças inseridas em contextos sociolinguísticos

monolingues percorrem, como também no que concerne à composição do seu léxico inicial e à influência que a quantidade e a qualidade da exposição linguística exerce no desenvolvimento da competência lexical das mesmas.

Desta forma, o enquadramento teórico realizado ao longo deste capítulo permite-nos apresentar, como síntese, as seguintes ideias-chave sobre o desenvolvimento lexical de crianças monolingues:

- (1) ainda que haja variações individuais no ritmo de aquisição lexical, todas as crianças, com um desenvolvimento normal da linguagem, percorrem os mesmos estádios de aquisição lexical em faixas etárias muito semelhantes;
- (2) em estádios iniciais de desenvolvimento linguístico, o léxico de crianças de muitas comunidades linguísticas é composto maioritariamente por itens lexicais pertencentes à categoria gramatical dos nomes;
- (3) a quantidade e a qualidade do *input* a que as crianças são expostas, em contexto doméstico, influencia significativamente o desenvolvimento da sua competência lexical.



### Capítulo 3 – A aquisição de duas (ou mais) línguas durante a infância

Em muitas partes do mundo, a aquisição linguística bilingue é a norma e não a exceção (Bialystok, 2009; Libben e Goral, 2015). Segundo Baker (2006) e Grosjean (2010), os falantes bilingues constituem mais de metade da população mundial, estimando-se que cerca de dois terços das crianças de todo o mundo cresçam expostas a duas línguas (Crystal, 1997). Contudo, apesar de o bilinguismo ser, de facto, um fenómeno cada vez mais comum no mundo globalizado em que vivemos, ainda persiste a ideia estereotipada de que a exposição precoce e simultânea a duas (ou mais) línguas prejudica a aquisição linguística das crianças, originando atrasos de desenvolvimento linguístico e cognitivo, bem como insucesso escolar (Macnamara, 1966).

A aquisição de duas (ou mais) línguas durante a infância pode ocorrer num conjunto variado de contextos, entre os quais se destacam: (1) o de crianças que crescem em territórios que reconhecem mais do que uma língua oficial, como, por exemplo, o País de Gales ou as comunidades autónomas espanholas Galiza e Catalunha; (2) o de crianças que crescem em territórios nos quais a língua oficial não corresponde às línguas faladas pelas comunidades nativas, como é o caso de muitas crianças que residem em alguns dos países-membros da Comunidade dos Países de Língua Portuguesa (CPLP), para quem a língua oficial (i.e. o português) pode não coincidir com a(s) língua(s) falada(s) em contexto familiar; e (3) o de crianças que, em situações de mobilidade migratória, crescem expostas à(s) língua(s) nativa(s) da família emigrante (i.e. língua(s) de herança) em contextos restritos e a uma outra língua falada e utilizada na sociedade de acolhimento.

Neste âmbito, Romaine (1999) apresenta seis contextos sociolinguísticos nos quais a aquisição bilingue, durante a infância, pode ocorrer:

- (1) *one-person-one-language*, no qual os progenitores não partilham a mesma língua materna, mas detêm algum grau de proficiência linguística na língua um do outro. Uma das línguas familiares é também a língua da sociedade maioritária e cada progenitor utiliza a sua L1 nas interações verbais com os filhos. Este é o contexto, por exemplo, de crianças luso-alemãs, residentes na Alemanha, que estão inseridas em agregados familiares em que um dos progenitores nasceu nesse país, tendo adquirido o alemão como LM, e o outro nasceu em Portugal, tendo adquirido o PE como LM. Contudo, o que faz com que estas crianças se enquadrem neste primeiro contexto sociolinguístico apresentado por Romaine é o facto de, apesar de os progenitores terem

desenvolvido competências comunicativas na língua um do outro, estes optam por usar cada um a sua LM na interação com os filhos, expondo-os, desta forma, quer à língua minoritária, quer à língua majoritária, em contexto doméstico;

- (2) *non-dominant home language*, no qual os progenitores têm línguas maternas diferentes, sendo que uma das línguas é a língua da sociedade majoritária. Contudo, ambos os progenitores interagem com os filhos na língua minoritária, sendo as crianças expostas à língua dominante somente “fora de casa”. Neste contexto, poderiam estar inseridas as crianças luso-alemãs apresentadas, como exemplo, no ponto anterior, se não fosse pelo facto de os seus progenitores terem adotado, na interação com os filhos, uma estratégia linguística diferente: o progenitor falante nativo da língua majoritária (i.e. o alemão) abdica de usar a sua LM na comunicação com os filhos e interage com estes apenas na língua minoritária (i.e. o português), de modo a reforçar a quantidade de exposição linguística à língua a que as crianças têm menos acesso. As crianças são, então, expostas majoritariamente à língua portuguesa em casa, recebendo *input* da língua alemã apenas fora do contexto doméstico;
- (3) *non-dominant home language without community support*, no qual ambos os progenitores partilham a mesma língua materna, não sendo esta a língua da sociedade majoritária. As crianças são expostas à língua dos progenitores em casa e à língua dominante fora do contexto doméstico. Este é o contexto, por exemplo, de crianças lusodescendentes, residentes em países cuja língua oficial não é a portuguesa, que estão inseridas em agregados familiares em que ambos os progenitores são falantes nativos do PE. Estas crianças são expostas majoritariamente ao português em casa e à língua majoritária fora do contexto doméstico;
- (4) *double non-dominant home language without community support*, no qual os progenitores têm línguas maternas diferentes, diferindo a língua da sociedade majoritária das mesmas. Os progenitores interagem com os filhos em cada uma das suas L1, podendo as crianças, neste contexto, desenvolver competências linguísticas em mais do que duas línguas. Este é o contexto, por exemplo, de crianças luso-alemãs, residentes em Inglaterra, que fazem parte de agregados familiares em que um dos progenitores é de origem portuguesa, tendo adquirido o PE como LM, e o outro é de origem alemã, tendo adquirido

o alemão como LM. Na comunicação com os filhos, cada progenitor usa a sua língua nativa. Estas crianças são expostas ao PE e ao alemão em casa e à língua maioritária (i.e. o inglês) fora do contexto doméstico;

(5) *non-native parents*, no qual os progenitores partilham a mesma língua materna, sendo esta também a língua da sociedade maioritária. No entanto, um dos progenitores interage com os filhos numa língua diferente da sua L1. Este é o contexto, por exemplo, de crianças portuguesas, residentes em Portugal, cujos progenitores são também portugueses e têm o PE como LM. Apesar de estarem reunidas todas as condições para que estas crianças sejam monolíngues, um dos progenitores, que desenvolveu competências comunicativas noutra língua, por via de formação académica e/ou por ter residido durante um período de tempo no estrangeiro, opta por interagir com os filhos na língua que adquiriu como LE e/ou L2, levando a que estes se tornem bilingues. Assim, estas crianças são expostas à língua maioritária (i.e. o PE) na sociedade dominante e também em casa através de um dos progenitores e a *input* não nativo de uma outra língua (i.e. LE) através do outro progenitor, também em contexto doméstico;

(6) *mixed languages*, no qual ambos os progenitores são bilingues, podendo alguns sectores da sociedade maioritária ser igualmente bilingues. Nas suas interações com os filhos, os pais alternam entre os códigos linguísticos e misturam as línguas. Este é o contexto, por exemplo, de algumas crianças espanholas residentes na Galiza, as quais, para além de estarem inseridas numa sociedade bilingue (i.e. espanhol-galego), fazem parte de agregados familiares em que ambos os progenitores são também bilingues. Estas crianças são expostas, quer em contexto doméstico, quer fora deste, ao espanhol e ao galego, sendo que o *input* que recebem é marcado pela alternância de códigos e interferências linguísticas.

Por outro lado, como foi realçado anteriormente (cf. secções 1.1 e 1.3), as crianças bilingues podem entrar em contacto com duas (ou mais) línguas em diferentes fases da vida, podendo o processo de aquisição bilingue ocorrer de forma simultânea, no caso de crianças que são expostas a duas línguas desde o nascimento (ou até aos três anos, cf. Meisel, 2011), ou de modo sequencial, no caso de iniciarem a exposição a uma outra língua numa fase posterior ao início de aquisição da L1. No caso específico de crianças

bilingues de herança, vimos que estas podem ser falantes bilingues simultâneas ou sequenciais, sendo que a língua de herança é adquirida como língua materna, isto é, desde a nascença e em contexto naturalístico. As situações de bilinguismo sequencial ocorrem nos casos em que a criança nasce no país de origem mas emigra com a família ainda em fase precoce do seu desenvolvimento, sendo exposta à língua de acolhimento apenas a partir desse momento. Podem também ocorrer nos casos em que a criança nasce já no país de acolhimento mas em comunidades imigrantes com forte presença na sociedade (por exemplo, turcos na Alemanha), sendo que a criança nos primeiros anos de vida interage pouco com falantes da língua maioritária. Nestes casos os FH são expostos única-/maioritariamente à LH nos primeiros anos de vida, adquirindo-a nos mesmos moldes que os falantes monolíngues da língua-alvo. Alterações à quantidade e à qualidade do *input* ocorrem, nestes casos, depois da imersão da criança no ambiente linguístico da língua maioritária. O contacto com a LH vai diminuindo e a exposição à língua dominante da sociedade vai aumentando. Nos casos da exposição simultânea à LH e à língua de acolhimento desde a nascença, a criança FH tem contacto com a LH desde a nascença no seio da família, mas é também exposta à língua dominante, por exemplo, quando os pais são eles próprios falantes de herança. Importa realçar que em ambos os casos estamos na presença de bilinguismo precoce. Assim sendo, resta-nos, então, procurar compreender, de um modo geral, como se desenvolve a aquisição precoce de duas (ou mais) línguas maternas.

Desta forma, na secção que se segue, apresentar-se-ão algumas das principais questões que têm motivado os estudos sobre o bilinguismo precoce, destacando-se as semelhanças e as diferenças observadas entre o desenvolvimento linguístico de crianças bilingues simultâneas e o de crianças monolíngues.

### **3.1 Diferenças e semelhanças entre a aquisição linguística monolíngue e a aquisição bilingue precoce**

As primeiras investigações sistemáticas centradas na aquisição linguística bilingue remontam à primeira metade do século XX, com a publicação dos estudos de Ronjat (1913) e de Leopold (1939-1949). Dos trabalhos pioneiros destes autores, que analisaram o desenvolvimento linguístico dos próprios filhos, expostos a duas línguas desde o nascimento em contexto naturalístico, salienta-se o facto de os mesmos não

perspetivarem a aquisição simultânea de duas línguas como algo negativo e prejudicial ao desenvolvimento linguístico dos sujeitos analisados.

No final dos anos setenta, inicia-se um intenso e prolongado debate que procura investigar se, em estádios iniciais de aquisição linguística, as crianças bilingues começam por desenvolver um sistema linguístico unitário ou se, pelo contrário, estas desenvolvem, desde muito cedo, dois sistemas linguísticos diferenciados. Apesar de, atualmente, a hipótese da existência de um único sistema linguístico inicial ser consensualmente rejeitada (De Houwer, 1990, 1995; Genesee, 1989, 2001; Genesee e Nicoladis, 2009; Meisel, 1989, 2001; Nicoladis e Genesee, 1997), os seus preconizadores defendem que as crianças bilingues começam por possuir um único sistema linguístico com elementos de ambas as línguas que vão sendo gradualmente diferenciados (Redlinger e Park, 1980; Volterra e Taeschner, 1978). A formulação da hipótese do sistema inicial único mais referida na literatura é a de Volterra e Taeschner (1978). Segundo as autoras, no processo de aquisição linguística bilingue inicial, a criança passa por três estádios de desenvolvimento distintos:

“In the first stage the child has one lexical system which includes words from both languages. A word in one language almost always does not have a corresponding word with the same meaning in the other language. In fact, in this stage the language development of the bilingual child seems to be like the language development of the monolingual child. [...] In the second stage the child distinguishes two different lexicons, but applies the same syntactic rules to both languages. [...] In the third stage the child speaks two languages differentiated both in lexicon and syntax, but each language is associated with the person using that language” (p. 312).

Fundamentando-se em estudos empíricos sobre a aquisição e o desenvolvimento bilingue de vários domínios linguísticos, sobretudo o lexical e o morfossintático, muitos autores têm desconstruído a hipótese do sistema unitário inicial (De Houwer, 1990, 1995; Genesee, 1989; Genesee *et al.*, 1995; Holowka *et al.*, 2002; Johnson e Lancaster, 1998; Köppe, 1996; Meisel, 1989; Petitto *et al.*, 2001; Quay, 1995; entre outros).

No domínio lexical, autores como Quay (1995), por exemplo, defendem que as crianças bilingues diferenciam, desde estádios iniciais de desenvolvimento linguístico, os sistemas lexicais das línguas que estão a adquirir. Tendo analisado o desenvolvimento linguístico de uma criança bilingue espanhol-inglês, entre o seu nascimento e os 22 meses, e observado a presença de pares de palavras com tradução equivalente (i.e. *translation*

*equivalents*) no seu léxico produtivo inicial, a autora conclui que esta já possui, a partir de fases iniciais de desenvolvimento linguístico, recursos lexicais que lhe permitem optar entre os dois sistemas lexicais em aquisição para interagir nas diferentes línguas faladas pelos seus interlocutores.

Tal como no domínio lexical, também vários estudos têm revelado que as crianças bilíngues simultâneas desenvolvem, desde muito cedo, dois sistemas morfossintáticos diferenciados (De Houwer, 1990; Meisel, 1989; Paradis e Genesee, 1996; Sinka e Schelletter, 1998; entre outros). Num estudo com dados recolhidos longitudinalmente, entre os 12 e os 48 meses, de duas crianças bilíngues alemão-francês, Meisel (1989) procura demonstrar que as crianças bilíngues distinguem, desde estádios iniciais de desenvolvimento linguístico, os sistemas sintáticos das línguas que estão a adquirir. Ao analisar dois fenómenos sintáticos cujas propriedades estruturais não são coincidentes nas línguas-alvo (i.e. a concordância sujeito-verbo e a ordem das palavras em determinados contextos sintáticos), o autor verifica que as crianças observadas diferenciam a ordem das palavras em ambas as línguas desde que começam a produzir enunciados multivocabulares. Além disso, flexionam corretamente os verbos para concordar com o sujeito, de acordo com as regras de cada língua, assim que a posição de sujeito começa a ser preenchida consistentemente nos seus enunciados (i.e. por volta dos 23 meses para um dos sujeitos e dos 32 meses para o outro). De acordo com Meisel (1989), os resultados deste estudo sugerem que uma criança exposta a duas línguas desde muito cedo é capaz de separar os dois sistemas gramaticais sem passar por uma fase de confusão temporária, como fora defendido por Volterra e Taeschner (1978).

Nas últimas décadas, principalmente a partir dos anos noventa, tem-se assistido à emergência de inúmeros estudos, na área da aquisição linguística bilingue, que têm como uma das principais questões de investigação verificar se as crianças bilíngues simultâneas percorrem trajetórias e estádios de aquisição linguística idênticos aos dos seus pares monolíngues nos vários domínios linguísticos, como os da morfossintaxe (De Houwer, 1990; Paradis e Genesee, 1996; Silva-Corvalán, 2014), do léxico (Holowka *et al.*, 2002; Pearson *et al.*, 1993; Pearson e Fernández, 1994; Petitto *et al.*, 2001) e da fonética e fonologia (Bosch e Sebastián-Gallés, 2003; Burns *et al.*, 2007; Fabiano-Smith e Goldstein, 2010; Ramírez *et al.*, 2016).

Embora ainda se encontrem leigos e mesmo especialistas (p. ex. professores ou terapeutas da fala) que acreditam, como foi mencionado anteriormente, que o bilinguismo é prejudicial ao desenvolvimento linguístico das crianças, os resultados da investigação

realizada na área de aquisição linguística bilingue comprovam que a mente humana está biologicamente preparada para a aquisição de duas (ou mais) línguas (De Houwer, 1995, 2009; Genesee, 2001; Meisel, 2001, 2011). O bilinguismo é, portanto, um estado normal que não exige esforços cognitivos adicionais às crianças.

De facto, uma criança que é exposta, regularmente, a duas (ou mais) línguas desde muito cedo adquire de forma espontânea e natural dois (ou mais) sistemas linguísticos, percorrendo os mesmos estádios de desenvolvimento linguístico que os seus pares monolíngues e alcançando os principais marcos de aquisição linguística em faixas etárias idênticas (Bialystok *et al.*, 2009; De Houwer, 2005, 2009; Grosjean, 2010; Holowka *et al.*, 2002; Meisel, 2006, 2011; Montrul, 2016; Nicoladis e Genesee, 1997; Pearson, 2009; Petitto *et al.*, 2001), quando ambas as línguas são tidas em consideração.

Assim, à semelhança das crianças monolíngues, também as crianças bilingues passam por uma fase pré-verbal que culmina com a produção da primeira palavra por volta do primeiro aniversário (Conboy e Montanari, 2016; Doyle *et al.*, 1978; Genesee e Nicoladis, 2009; Holowka *et al.*, 2002), prolongando-se o estádio de produção de enunciados holofráscos até ao aparecimento das primeiras combinações de palavras, entre os 18 e os 23 meses (Patterson, 1998), a partir das quais os enunciados se vão tornando progressivamente mais complexos (De Houwer, 1995, 2005).

No que respeita ao desenvolvimento do léxico inicial, as crianças expostas simultaneamente a mais do que uma língua também alcançam os mesmos marcos de aquisição de vocabulário dentro do intervalo cronológico de variação considerado normal em crianças monolíngues (Holowka *et al.*, 2002; Pearson *et al.*, 1993; Pearson e Fernández, 1994; Petitto *et al.*, 2001). Estudos realizados com crianças bilingues revelam que estas, geralmente, não só proferem a primeira palavra por volta dos 12 meses (Doyle *et al.*, 1978; Holowka *et al.*, 2002; Pearson e Fernández, 1994), como também apresentam um repertório lexical produtivo de dez palavras aos 14 meses (Pearson *et al.*, 1993) e alcançam o marco das 50 palavras aproximadamente aos 20 meses, quando os itens lexicais produzidos em ambas as línguas são contabilizados (Holowka *et al.*, 2002). Um estudo realizado por Pearson e Fernández (1994) mostra ainda que, tal como nos falantes monolíngues, após uma fase de aquisição lexical a um ritmo lento, a velocidade de crescimento do léxico produtivo aumenta significativamente entre os 18 meses e o fim do segundo ano de vida, tendo os autores observado, inclusivamente, o fenómeno da *explosão lexical* na maioria dos seus participantes bilingues.

Em paridade com as crianças monolíngues, também existe uma assimetria entre a competência lexical recetiva e a produtiva no desenvolvimento linguístico de crianças bilingues, sendo que os repertórios lexicais recetivos são sempre superiores aos produtivos (De Houwer, 2009; De Houwer *et al.*, 2014; Kan e Kohnert, 2005; Özcan *et al.*, 2016; Pearson *et al.*, 1993; Sheng *et al.*, 2011). De facto, os resultados de um estudo realizado por De Houwer *et al.* (2014), com 31 crianças bilingues francês-holandês e 30 crianças monolíngues holandês, confirmam que a compreensão oral é superior à produção verbal em ambos os grupos investigados. Assim sendo, por cada palavra produzida aos 13 meses, as crianças monolíngues compreendem em média cerca de 17 palavras, ao passo que as crianças bilingues compreendem aproximadamente 22 palavras em holandês e cerca de 32 quando ambas as suas línguas são tidas em consideração. Por sua vez, por cada item lexical produzido aos 20 meses, as crianças monolíngues compreendem em média cerca de quatro palavras, enquanto as crianças bilingues compreendem aproximadamente cinco palavras em holandês - os autores não apresentam resultados relativamente aos repertórios lexicais holandês e francês combinados.

Por sua vez, no que respeita à composição do léxico produtivo das crianças bilingues, os resultados de vários estudos indicam que, em estádios iniciais de desenvolvimento linguístico, os repertórios lexicais de falantes bilingues também se assemelham aos das crianças monolíngues das línguas-alvo, verificando-se, inclusive, o fenómeno da *tendência nominal* observado em falantes monolíngues de diversas comunidades linguísticas (Barnes e García, 2013; Conboy e Thal, 2006; David e Wei, 2005; Lucas e Bernardo, 2008; Nicoladis, 2001; Özcan *et al.*, 2016; dos Santos e Kern, 2013; Xuan e Dollaghan, 2013). Tal como os repertórios lexicais dos seus pares monolíngues, também os das crianças bilingues são constituídos inicialmente por um maior número de palavras sociais<sup>4</sup> (*social words*) que é rapidamente superado por itens lexicais pertencentes à categoria gramatical dos nomes, seguindo-se um aumento gradual na proporção de verbos, de adjetivos e de palavras funcionais (Conboy e Montanari, 2016).

Um dos estudos que procurou verificar se, em estádios iniciais de desenvolvimento linguístico, os repertórios lexicais de crianças bilingues apresentam padrões de distribuição e de desenvolvimento lexical idênticos aos dos seus pares monolíngues é o de David e Wei (2005). Os autores investigaram longitudinalmente o

---

<sup>4</sup> Palavras relacionadas com nomes de pessoas, efeitos sonoros, jogos e rotinas (Conboy e Thal, 2006; David e Wei, 2005).

léxico produtivo inicial de 13 crianças bilíngues francês-inglês, entre os 12 e os 36 meses de vida, comparando-o com os resultados de estudos desenvolvidos com crianças monolíngues das respectivas línguas-alvo. Desta forma, tendo analisado a amostra por níveis de dimensão lexical, em cada uma das suas línguas, os autores observam que, à exceção da primeira fase de produção linguística (i.e. até ao marco das 50 palavras), na qual as crianças bilíngues produzem maioritariamente palavras sociais, os repertórios lexicais das crianças bilíngues apresentam uma maior proporção de nomes relativamente às outras categorias gramaticais analisadas, isto é, às dos predicados (verbos e adjetivos) e dos itens lexicais pertencentes às classes fechadas de palavras, sendo que a produção de verbos e de adjetivos é superior à produção de palavras desta última categoria gramatical. David e Wei (2005) verificam também que a composição do repertório lexical dos falantes bilíngues apresenta padrões de crescimento idênticos aos das respectivas crianças monolíngues em função do tamanho de vocabulário, passando, inclusivamente, pelos mesmos quatro estádios de desenvolvimento lexical. Assim, no primeiro estádio de desenvolvimento, os repertórios lexicais das crianças bilíngues é dominado por palavras sociais; no segundo estádio, entre as 50 e as 200 palavras, observa-se um crescimento acelerado da classe nominal, apesar de este começar a estabilizar por volta do marco das 100 palavras; e, nos terceiro e quarto estádios, começam a surgir, com mais frequência, os itens lexicais pertencentes às categorias dos predicados e das classes fechadas de palavras, respetivamente. Para além de demonstrarem que as crianças bilíngues percorrem estádios de aquisição lexical idênticos aos dos seus pares monolíngues, os resultados deste estudo revelam ainda que as categorias gramaticais analisadas apresentam aproximadamente a mesma distribuição e se desenvolvem em paralelo em ambas as línguas-alvo.

Por sua vez, num dos raros estudos sobre o desenvolvimento lexical de falantes de herança do PE, dos Santos e Kern (2013) analisaram a composição dos repertórios lexicais de 29 crianças bilíngues português-francês, entre os 24 e os 36 meses, nas suas duas línguas, comparando-os entre si e com os de crianças monolíngues da língua maioritária (i.e. o francês). Tendo centrado a análise nas categorias gramaticais nomes e predicados, bem como em itens lexicais pertencentes às classes fechadas de palavras e palavras sociais, os autores verificam que, no que respeita à língua francesa, a composição dos repertórios lexicais das crianças bilíngues apresenta padrões de distribuição idênticos aos dos seus pares monolíngues, ou seja, estes são constituídos maioritariamente por nomes, seguindo-se, por ordem decrescente, os predicados, os itens lexicais das classes

fechadas de palavras e, por último, as palavras sociais. Já no que se refere à análise da composição lexical dos repertórios das crianças bilingues na língua minoritária, os autores constatam que esta diverge da distribuição lexical observada na língua maioritária, isto é, embora seja constituído maioritariamente por nomes e predicados, sendo a categoria nominal superior à dos predicados, o léxico produtivo das crianças bilingues na língua portuguesa apresenta uma maior proporção de palavras sociais comparativamente aos itens pertencentes às classes fechadas de palavras. Os autores sugerem que as diferenças observadas entre as línguas destas crianças bilingues se deve ao facto de a língua minoritária ainda não estar no mesmo nível de desenvolvimento linguístico em que a maioritária está. Este estudo será retomado na secção seguinte, na qual se fará uma abordagem ao conjunto de fatores preditivos do desenvolvimento lexical de crianças bilingues.

Contudo, embora percorram os mesmos estádios de aquisição linguística que os seus pares monolingues, em faixas etárias idênticas, os falantes bilingues não desenvolvem, geralmente, ambas as suas línguas ao mesmo ritmo (Bialystok *et al.*, 2009; De Houwer, 2009; Pearson, 2009). De facto, vários estudos têm reportado não só que as crianças bilingues tendem a ter repertórios lexicais recetivos e produtivos quantitativamente inferiores aos dos seus pares monolingues, quando apenas uma das suas línguas é avaliada (Bialystok *et al.*, 2010; Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; Hoff *et al.*, 2012; Marchman *et al.*, 2010; Pearson *et al.*, 1993; Poulin-Dubois *et al.*, 2013; Umbel *et al.*, 1992; entre outros), como também que estas podem demorar mais tempo a adquirir determinadas propriedades gramaticais das línguas-alvo (Flores e Barbosa, 2014; Flores *et al.*, 2016; Gathercole, 2002a, 2002b, 2002c; Gathercole e Thomas, 2009; Paradis, 2010; Thordardottir, 2015; entre outros).

No domínio específico da aquisição lexical, apesar de as crianças bilingues apresentarem dimensões de vocabulário inferiores aos das correspondentes crianças monolingues quando apenas uma das suas línguas é avaliada, muitos estudos comprovam que, quando ambas as línguas são tidas em consideração e se avaliam os seus repertórios lexicais no que respeita à dimensão do vocabulário total (i.e. itens lexicais contabilizados em ambas as línguas) e conceptual (i.e. subtraindo-se ao vocabulário total os equivalentes de tradução), os falantes bilingues têm um desempenho lexical idêntico, ou superior, ao dos falantes monolingues das línguas-alvo (De Houwer *et al.*, 2014; Hoff *et al.*, 2012; Junker e Stockman, 2002; Pearson *et al.*, 1993; Poulin-Dubois *et al.*, 2013; entre outros).

Segundo Oller *et al.* (2007), as diferenças observadas na dimensão do vocabulário de falantes bilingues, em cada uma das suas línguas, devem-se ao facto de o seu conhecimento lexical estar distribuído por duas línguas (i.e. “*distributed characteristic of bilingual lexical knowledge*”) e de ser adquirido, geralmente, em diferentes contextos comunicacionais, levando a que alguns conceitos estejam codificados numa língua-alvo, mas não na outra e vice-versa. Desta forma, é comum que as crianças bilingues possuam, por exemplo, um tamanho lexical de termos académicos maior na língua de escolarização (geralmente, a língua maioritária) do que na língua falada em contexto doméstico, assim como também é comum conhecerem mais vocabulário relacionado com o ambiente doméstico na língua que é falada neste contexto do que na língua em que são escolarizadas (Hoff, 2006; Hoff e Core, 2013).

No que respeita ao domínio da morfossintaxe, autores como Gathercole e Thomas (2009) recorrem à hipótese da “massa crítica de exposição” (*critical mass of input*) para justificar a aquisição tardia de determinadas propriedades morfossintáticas das línguas-alvo de crianças bilingues, como é o caso do género gramatical, por exemplo, no galês. Segundo esta hipótese, é necessário que uma criança seja exposta a uma quantidade mínima de *input* para acumular a “massa crítica” necessária para extrair generalizações dos padrões linguísticos que regem uma determinada propriedade gramatical. Assim sendo, uma criança que tenha exposição reduzida numa determinada língua-alvo demorará mais tempo a acumular a “massa crítica” indispensável para a aquisição de certas propriedades morfossintáticas dessa mesma língua. Tal como Gathercole (2007) afirma:

“even if the bilingual child shows initial delays in the acquisition of some structures, there will be *eventual neutralisation of differences across monolingual and bilingual groups as all children gradually gain the critical mass of linguistic information on which to base generalisations*” (pp. 229-230).

Devido às diferenças observadas entre o desenvolvimento linguístico de crianças bilingues e o de crianças monolíngues, bem como à heterogeneidade de contextos sociolinguísticos em que a aquisição de duas (ou mais) línguas ocorre, os estudos desenvolvidos nas últimas décadas têm procurado simultaneamente identificar fatores intervenientes no processo de aquisição linguística bilingue durante a infância, tendo sido apresentados, neste âmbito, fatores relacionados com a quantidade e a qualidade da

exposição linguística a que as crianças bilingues têm acesso (Gathercole, 2014; Montrul, 2016; Unsworth, 2016).

Desta forma, na secção seguinte, problematizar-se-á o papel que diferentes fatores de *input* desempenham no processo de aquisição de duas (ou mais) línguas durante a infância, principalmente no que respeita ao desenvolvimento da competência lexical de crianças bilingues.

### **3.2 O papel do *input* em ambientes bilingues**

A investigação realizada na área de aquisição de duas (ou mais) línguas revela que o desenvolvimento linguístico de crianças bilingues é caracterizado pela variação não só no que respeita à experiência sociolinguística de cada criança, isto é, ao contexto sociolinguístico em que a aquisição bilingue ocorre e à quantidade e qualidade da exposição linguística, bem como relativamente aos resultados individuais das crianças nas suas duas (ou mais) línguas, que podem estar mais próximos ou mais afastados dos padrões de desenvolvimento monolíngue das línguas-alvo (Gathercole, 2014; Montrul, 2016; Unsworth, 2016).

No domínio específico da aquisição lexical, os resultados dos inúmeros estudos realizados com crianças bilingues mostram que a variação observada na dimensão dos seus repertórios lexicais, numa ou nas suas duas línguas, está significativamente associada a fatores relacionados com a quantidade e a qualidade da exposição linguística a que estes falantes têm acesso (Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; David e Wei, 2008; Gathercole e Thomas, 2009; Hoff *et al.*, 2012; Hoff *et al.*, 2014; Marchman *et al.*, 2010; Pearson *et al.*, 1997; Place e Hoff, 2011; Smithson *et al.*, 2014; Thordardottir, 2011; entre muitos outros).

Neste âmbito, variáveis como o número de línguas faladas em contexto doméstico, o número de progenitores falantes das línguas-alvo, bem como a (in)existência de irmãos mais velhos no agregado familiar e a quantidade de produção linguística das crianças bilingues, entre outras, têm sido consideradas determinantes do grau de exposição linguística destas crianças às línguas-alvo e, conseqüentemente, preditivas do desenvolvimento lexical em cada uma das suas línguas. Cumulativamente, o contacto com diferentes fontes de *input* (mais próximas ou mais afastadas da língua padrão), a variedade de interlocutores falantes nativos das línguas-alvo e a diversidade de atividades de lazer em que as crianças bilingues participam (utilizando as línguas-alvo), entre outros,

têm sido igualmente apontados como fatores de *input* qualitativos preditivos do desenvolvimento lexical bilingue (Montrul, 2016; Unsworth, 2016).

Assim sendo, apresentar-se-ão, de seguida, os principais resultados da investigação centrada no papel da variação do *input* linguístico na aquisição de duas (ou mais) línguas durante a infância, sendo atribuído um especial enfoque ao papel que este desempenha no desenvolvimento da competência lexical de crianças bilingues.

### **3.2.1 A quantidade e a qualidade da exposição linguística**

Um contributo importante para a compreensão dos efeitos da exposição linguística sobre o desenvolvimento da competência lexical receptiva e produtiva de crianças bilingues foi o de Thordardottir (2011). A autora analisou a proficiência linguística de 49 crianças bilingues simultâneas francês-inglês, em cada uma das suas línguas, comparando-as entre si e com as crianças monolingues das línguas-alvo, da mesma faixa etária (i.e. cinco anos). Os resultados deste estudo revelam que a quantidade de exposição linguística a que as crianças bilingues estão expostas em cada uma das suas línguas está significativamente correlacionada com o seu desempenho lexical nas mesmas: as crianças com níveis de exposição linguística equilibrados entre ambas as línguas apresentam níveis de desempenho lexical idênticos nas suas duas línguas, ao passo que as crianças com padrões de exposição linguística assimétricos apresentam melhores resultados linguísticos na língua da qual recebem mais *input*. A autora constata ainda que a competência lexical produtiva é mais vulnerável a efeitos de exposição reduzida do que a receptiva. Desta forma, enquanto as crianças com exposição linguística equilibrada entre as duas línguas (i.e. 40-60% de *input*) obtêm resultados equiparados aos dos seus pares monolingues nos testes de vocabulário receptivo, somente as que têm níveis de exposição linguística superiores a 60% numa dada língua é que apresentam um desempenho lexical produtivo dentro dos padrões de desenvolvimento monolingue dessa mesma língua. No entanto, Thordardottir (2011) salienta que as crianças observadas, residentes em Montreal (Canadá), estão inseridas num contexto sociolinguístico favorável ao desenvolvimento linguístico bilingue, isto é, no qual ambas as línguas têm um estatuto sociopolítico equiparável e usufruem do mesmo prestígio, alertando para o facto de que os resultados de estudos com crianças inseridas em ambientes sociolinguísticos desfavoráveis à aquisição linguística bilingue podem apresentar conclusões diferentes.

É com um grupo de crianças bilingues inseridas num contexto sociolinguístico menos favorável ao desenvolvimento de competências comunicativas bilingues que surge a investigação de Sheng *et al.* (2011). As autoras analisaram a competência lexical de crianças bilingues mandarim-inglês residentes nos EUA, onde o inglês tem o estatuto de língua oficial, sendo a língua utilizada pela sociedade maioritária, e o mandarim o de língua minoritária, sendo utilizada em contextos limitados às comunidades emigrantes asiáticas. Para a avaliação do seu conhecimento lexical, as crianças bilingues analisadas, agrupadas em função da faixa etária (3-5 anos; 6-8 anos), participaram em tarefas de identificação (competência recetiva) e de nomeação (competência produtiva) de imagens em ambas as línguas. Quanto à experiência sociolinguística das crianças bilingues, estas eram expostas predominantemente ao mandarim em contexto doméstico, tendo iniciado a exposição linguística à língua maioritária aquando da entrada para a creche ou para o infantário. Os resultados deste estudo revelam que a dimensão do vocabulário conceptual total (i.e. itens lexicais de ambas as línguas, contabilizando apenas uma vez as palavras referentes a um único conceito) das crianças mais novas é superior ao tamanho do léxico contabilizado em cada uma das suas línguas, contudo, o tamanho do vocabulário total do grupo das crianças mais velhas é quase equivalente ao léxico contabilizado na língua maioritária. Os resultados mostram ainda que o nível de acurácia é maior nos testes de identificação do que nos de nomeação e que a diferença é maior na língua minoritária do que na maioritária, sendo que, no caso do inglês, esta vai diminuindo com a idade e com o crescimento lexical. Tendo em conta os resultados, as autoras concluem que, de uma faixa etária para a outra, o ritmo de aquisição lexical na língua maioritária aumenta significativamente, sugerindo, simultaneamente, o aparecimento de um estado de estagnação precoce na aquisição lexical e no desenvolvimento das competências produtiva e recetiva na língua minoritária. De acordo com Sheng *et al.* (2011), embora estas crianças sejam expostas maioritariamente ao mandarim em contexto doméstico, o facto de não terem contacto com outras fontes de *input* da língua minoritária na sociedade dominante origina este desequilíbrio lexical entre as duas línguas em aquisição, podendo resultar num declínio das competências comunicativas na língua minoritária.

O efeito da exposição linguística reduzida sobre o desenvolvimento da competência lexical produtiva de crianças bilingues em contextos menos favoráveis à aquisição de duas línguas é também comprovado por dos Santos e Kern (2013) e O'Toole *et al.* (2017). Neste estudo desenvolvido com 29 crianças bilingues português-francês (entre os 24 e os 36 meses), os autores avaliaram, através de relatórios parentais

padronizados, a dimensão do léxico produtivo das crianças bilíngues em ambas as suas línguas, comparando-o com o dos seus pares monolíngues da língua maioritária (i.e. do francês). Os resultados dos relatórios parentais revelam que o repertório lexical das crianças bilíngues é maior na língua maioritária do que na minoritária e que este não diverge dos padrões de desenvolvimento lexical monolíngue no que respeita à língua francesa. Para além da dimensão lexical, os autores avaliaram, através de questionários sociolinguísticos, o grau de exposição linguística dos falantes bilíngues a cada uma das suas línguas, tendo verificado que, em contexto doméstico, a maioria destes falantes recebe mais *input* da língua maioritária do que da minoritária. Os resultados desta investigação sugerem que o grau de exposição linguística às línguas-alvo, em contexto doméstico, é um fator preditivo do desenvolvimento lexical de crianças bilíngues em cada uma das suas línguas. Assim, uma maior exposição linguística ao francês e menor ao português resulta num conhecimento lexical superior na língua maioritária, comparável ao das correspondentes crianças monolíngues, e inferior na minoritária.

Num outro estudo centrado no efeito que diferentes padrões de exposição linguística exercem sobre o ritmo de aquisição lexical de falantes bilíngues, Hoff *et al.* (2014) analisaram, comparativa e longitudinalmente, as trajetórias de desenvolvimento lexical produtivo de 26 crianças bilíngues espanhol-inglês e de 31 crianças monolíngues de inglês, entre os 22 e os 48 meses, residentes nos EUA. No que respeita ao contexto familiar das crianças bilíngues, onze são provenientes de lares em que ambos os progenitores são falantes nativos de espanhol (i.e. língua minoritária) e 15 crianças são oriundas de famílias em que um dos progenitores é falante nativo de espanhol e o outro de inglês (i.e. língua maioritária). Quer as crianças bilíngues, quer as monolíngues, já haviam participado, entre os 22 e os 30 meses, em estudos sobre o desenvolvimento linguístico bilíngue e monolíngue (Hoff *et al.*, 2012; Place e Hoff, 2011), tendo os dados recolhidos anteriormente servido de ponto de partida para o estudo em questão. Os resultados de Hoff *et al.* (2014) revelam que as crianças bilíngues provenientes de famílias em que ambos os pais são falantes nativos de espanhol apresentam um ritmo de crescimento lexical mais acelerado do que as crianças dos outros grupos no que respeita à dimensão do vocabulário total (i.e. a soma do vocabulário de ambas as línguas), aproximando-se de um bilinguismo equilibrado aos 48 meses. Por sua vez, as crianças que estão mais expostas à língua da sociedade maioritária, ou seja, as que têm um progenitor falante nativo de espanhol e o outro de inglês, apresentam trajetórias de declínio no desenvolvimento da competência lexical da língua minoritária. As autoras

constatam ainda que, na língua maioritária, a proficiência linguística das crianças bilingues é equiparável à das correspondentes monolíngues. No entanto, as crianças que são expostas exclusivamente ao espanhol em casa apresentam resultados inferiores. Em suma, os resultados deste estudo longitudinal comprovam que o grau de exposição linguística em contexto doméstico é um fator preditivo da competência lexical de crianças bilingues e influencia significativamente as suas trajetórias de desenvolvimento linguístico, sendo que a língua minoritária é a mais suscetível a efeitos de exposição reduzida.

Um dos fatores relacionados com a quantidade de exposição linguística que uma criança bilingue recebe em cada uma das suas línguas é a presença de irmãos mais velhos no agregado familiar. De acordo com vários autores, a existência de irmãos mais velhos no seio familiar conduz, geralmente, a uma mudança nos hábitos linguísticos de famílias bilingues, uma vez que estes tendem a introduzir a língua da sociedade maioritária no ambiente doméstico (Barron-Hauwaert, 2011; Bridges e Hoff, 2014; Kopeliovich, 2013; Schwartz, 2010; Spolsky, 2009), podendo comprometer, desta forma, a transmissão intergeracional da língua minoritária.

Procurando avaliar a influência de irmãos mais velhos no grau de exposição e de desenvolvimento linguístico de crianças bilingues, Bridges e Hoff (2014) realizaram dois estudos com crianças bilingues simultâneas residentes nos EUA. No primeiro estudo, participaram 60 crianças bilingues, entre os 16 e os 30 meses, agrupadas de acordo com a existência, ou inexistência, de irmãos mais velhos no agregado familiar. No segundo estudo, os autores investigaram 27 crianças bilingues, em dois momentos temporais (aos 22 e aos 30 meses), tendo sido constituídos dois grupos distintos: crianças com irmãos mais velhos em idade escolar e crianças sem irmãos a frequentarem instituições de ensino. Os resultados de ambos os estudos revelam não só que as crianças mais novas são expostas mais frequentemente à língua maioritária (i.e. ao inglês) por meio de irmãos mais velhos do que através de outros membros do agregado familiar, como também que as crianças com irmãos mais velhos apresentam níveis de proficiência mais avançados na língua maioritária do que as outras crianças investigadas. Os autores constatam ainda que a presença de irmãos em idade escolar no seio familiar origina uma mudança nos hábitos linguísticos da figura maternal, ou seja, as mães de crianças a frequentarem instituições escolares utilizam com mais frequência a língua maioritária nas interações verbais com os seus filhos do que as mães cujos filhos ainda não estão em idade escolar. Para além disso, os resultados mostram que, comparativamente às crianças com irmãos a

frequentarem a escola, as crianças sem irmãos em idade escolar são mais proficientes na língua minoritária do que na maioritária, ocorrendo o processo inverso em crianças com irmãos a frequentarem instituições de ensino.

Outro fator que tem sido apontado como preditivo do desenvolvimento das competências comunicativas de falantes bilingues, numa ou nas suas duas (ou mais) línguas, é a produção linguística (i.e. *output*) dos mesmos (Bedore *et al.*, 2012; Bohman *et al.*, 2010; Cohen, 2016; Hammer *et al.*, 2012; Ribot *et al.*, 2017; Unsworth, 2012). Segundo Pearson (2007), a produção linguística numa língua-alvo provoca mais *input* nessa mesma língua, gerando um ciclo linguístico no qual: “A greater amount of input leads to greater proficiency, which leads to more use, which invites more input and the cycle starts again” (p. 400). Assim, se uma criança não usa verbalmente, por exemplo, uma língua de herança, então este processo pode inverter-se, levando a que esta receba menos exposição linguística da língua minoritária e a que se torne, conseqüentemente, menos proficiente na mesma. Bohman *et al.* (2010), por exemplo, ao avaliarem fatores preditivos do desenvolvimento linguístico de crianças bilingues espanhol-inglês, verificam que a quantidade de produção linguística dos seus participantes está positivamente correlacionada com o seu desempenho linguístico, em ambas as línguas, em testes nas áreas da morfossintaxe e da semântica. Bedore *et al.* (2012), por seu turno, constatam que, para além da quantidade de produção linguística estar associada à competência linguística de crianças bilingues espanhol-inglês em cada uma das suas línguas, esta tem um papel mais determinante na variação observada nos resultados individuais da língua inglesa do que a idade de início de exposição a esta língua. Recentemente, Ribot *et al.* (2017) investigaram a influência do fator *output* no desenvolvimento lexical recetivo e produtivo de crianças bilingues espanhol-inglês, residentes nos EUA, na língua maioritária (i.e. inglês). Os autores dividiram a amostra em dois grupos: um grupo constituído por crianças cuja produção linguística em inglês é superior à exposição linguística na mesma língua, isto é, crianças que, mesmo quando o discurso que lhes é dirigido é em espanhol, se exprimem frequentemente em inglês; e outro grupo constituído por crianças cujo *output* em inglês é inferior ao *input*, ou seja, crianças que usam a língua espanhola quando os seus interlocutores a utilizam e que optam por usar, por vezes, o espanhol, quando estes interagem com elas em inglês. Os resultados do estudo mostram que, na língua analisada, a competência lexical produtiva das crianças do primeiro grupo (i.e. *output* superior ao *input*) se desenvolve a um ritmo mais acelerado do que a das crianças do segundo grupo, cujo *output* é inferior ao *input*,

revelando que a quantidade de produção linguística é uma variável preditiva do desenvolvimento lexical produtivo de crianças bilingues. Contudo, a quantidade de *output* não apresenta correlações significativas com o desenvolvimento da competência lexical receptiva das crianças investigadas.

Embora a maioria dos estudos sobre os efeitos da quantidade de *input* no desenvolvimento linguístico bilingue se centre na exposição linguística que as crianças bilingues recebem no momento da recolha dos dados (i.e. exposição corrente), alguns autores têm procurado investigar o papel que o *tempo de exposição acumulada* (*cumulative length of exposure*) desempenha no processo de aquisição linguística destes falantes (Cohen, 2016; Gutiérrez-Clellen e Kreiter, 2003; Maas, 2014; Unsworth, 2012, 2013; Unsworth *et al.*, 2014). Unsworth (2012) define *tempo de exposição acumulada* como “the total amount of exposure from birth to the child’s current age, taking into account any variation over time” (p.16). Segundo a autora, em estudos que comparam a proficiência linguística de crianças bilingues com a de crianças monolíngues, o *tempo de exposição acumulada* (doravante também *TEA*) à L2 de crianças bilingues sequenciais é tradicionalmente calculado subtraindo-se à idade cronológica do falante bilingue a idade de início da sua exposição à L2 (Unsworth, 2011), ao passo que o *TEA* de crianças bilingues simultâneas tem sido tradicionalmente medido a partir do início de exposição linguística às suas L1, sendo, portanto, equivalente à idade cronológica das mesmas (Unsworth, 2013). No entanto, Unsworth (2013) defende que o cálculo da variável *TEA* pelo método tradicional não fornece valores precisos sobre a quantidade de exposição linguística que uma criança bilingue recebe ao longo da sua vida, uma vez que, contrariamente à quantidade de *input* a que uma criança monolíngue está exposta, a quantidade de exposição linguística que uma criança bilingue recebe, para além de estar distribuída por duas línguas, é muito variável em cada língua-alvo. Desta forma, tendo em consideração que “one year of “bilingual” language exposure is not the same as one year of “monolingual” language exposure, and the amount of exposure varies among bilinguals” (Unsworth, 2013:91), uma criança bilingue de cinco anos de idade pode ter um *tempo de exposição acumulada* de, por exemplo, quatro anos a uma língua-alvo, ao passo que o de uma outra criança bilingue da mesma idade pode ser de apenas dois anos, não podendo nunca corresponder ao de uma criança monolíngue de idade equivalente, o qual seria de cinco anos.

Assim sendo, Unsworth (2011) desenvolveu um instrumento de avaliação da exposição linguística – o *Utrecht Bilingual Language Exposure Calculator* (UBiLEC),

que, com base na análise detalhada de variáveis relacionadas com a quantidade e a qualidade do *input*, permite calcular o valor médio de exposição acumulada ao longo da vida<sup>5</sup>, bem como o valor médio de exposição corrente de falantes bilingues, entre os dois e os 18 anos de vida. Com base neste instrumento, alguns estudos têm verificado que o *TEA* é uma variável preditiva não só do desenvolvimento lexical bilingue (Cohen, 2016; Maas, 2014; Unsworth, 2012), como também da aquisição de determinadas propriedades morfossintáticas, como é o caso, por exemplo, do género gramatical no holandês (Unsworth, 2013; Unsworth *et al.*, 2014). Assim, quanto maior for o *tempo de exposição acumulada* numa língua-alvo, melhor será o desempenho linguístico das crianças bilingues na mesma.

Para além de variar na quantidade, também a exposição linguística bilingue varia na qualidade, sendo múltiplos os fatores que podem contribuir para as diferenças na qualidade do *input* a que as crianças bilingues têm acesso. Entre os vários fatores que parecem influenciar a riqueza da exposição linguística das crianças bilingues e, conseqüentemente, o seu desenvolvimento linguístico encontram-se: a diversidade de interlocutores e o contacto com falantes nativos e não nativos (Cornips e Hulk, 2008; Driessen *et al.*, 2002; Place e Hoff, 2011), bem como a variedade de diferentes fontes de *input*, como os *media* e a literatura, e a pluralidade de atividades em que as línguas-alvo são utilizadas (Cha e Goldenberg, 2015; Jia e Aaronson, 2003; Jia e Fuse, 2007; Leseman *et al.*, 2009; Patterson, 2002; Scheele *et al.*, 2010).

De facto, num estudo com 29 crianças bilingues espanhol-inglês de 25 meses, residentes no estado da Florida (EUA), no qual existe uma forte comunidade de imigrantes falantes de espanhol, Place e Hoff (2011) verificam que, para além da quantidade de *input*, em cada uma das suas línguas, ser um fator preditivo do desenvolvimento linguístico bilingue das crianças investigadas, propriedades específicas da exposição linguística ao inglês, designadamente o número de parceiros conversacionais com quem as crianças interagem exclusivamente em inglês, a variedade de falantes que fornecem *input* nesta língua e a percentagem de exposição linguística a falantes nativos, também apresentam correlações positivas com a competência lexical dos

---

<sup>5</sup> De acordo com Unsworth (2011), para o cálculo do *tempo de exposição acumulada*, devem ser tidos em conta: (1) a quantidade de *input* de cada língua-alvo que a criança recebeu de cada progenitor e de cada adulto a viver na mesma casa, por cada período de um ou de dois anos antes do momento da recolha de dados; (2) a frequência de creches ou de instituições (pré-) escolares durante esses períodos (de um ou de dois anos) antecedentes à recolha de dados, bem como a língua utilizada nesses locais; e (3) a(s) língua(s) utilizada(s) nos períodos de férias.

falantes bilingues nesta língua. Contudo, no que respeita ao espanhol, as autoras apenas encontram uma correlação negativa, que consideram ininterpretável, entre o número de contextos em que esta língua é utilizada em exclusivo e a dimensão do vocabulário espanhol. Este estudo de Place e Hoff (2011) revela também que a quantidade de exposição linguística a cada uma das línguas varia de acordo com as diferentes constelações familiares das crianças bilingues. Assim, as crianças provenientes de famílias em que um dos progenitores é falante nativo de inglês recebem mais *input* dessa língua, levando a que os seus repertórios lexicais de inglês sejam maiores do que os de espanhol, principalmente nas famílias em que a figura materna é a falante nativa de inglês. Por outro lado, as crianças de famílias em que ambos os progenitores são falantes nativos de espanhol recebem mais exposição linguística desta língua, sendo os seus repertórios lexicais de espanhol superiores aos de inglês.

Por sua vez, Ishizawa (2004) desenvolveu um estudo com 5962 crianças bilingues, de famílias imigrantes nos EUA, no qual procura investigar se a presença de avós, falantes não nativos da língua dominante (i.e. inglês), em agregados familiares multigeracionais influencia positivamente a manutenção de uma língua minoritária. Os resultados do estudo mostram que as crianças bilingues que coabitam com pais e avós falantes de uma língua minoritária têm mais probabilidade de falar essa língua em contexto doméstico do que as crianças que vivem em agregados menores (i.e. apenas com os pais ou com os avós).

Em último lugar, no que respeita à influência da exposição linguística através de atividades comunicativas relacionadas com os *media* e a literatura no desenvolvimento linguístico bilingue, Scheele *et al.* (2010) verificam que, por exemplo, a quantidade de *input* proveniente da atividade de contar histórias (mas não da de leitura conjunta) na L1 de crianças bilingues, de origem marroquina e turca a residirem na Holanda, está associada à competência lexical recetiva destas crianças nas suas L1. Para além disso, apesar de o *input* de atividades relacionadas com a música e a televisão não exercer influência no desenvolvimento lexical das mesmas, a atividade de cantar na L2 (i.e. holandês) está correlacionada com o vocabulário recetivo destas crianças na língua maioritária. Por sua vez, Patterson (2002) observa que a competência lexical produtiva de crianças bilingues espanhol-inglês é, em cada uma das suas línguas, significativa e positivamente influenciada pelo *input* proveniente dos pais em atividades de leitura, mas não pela quantidade de exposição linguística a programas televisivos. Já Jia e colegas (2003, 2007), tendo analisado, quantitativa e qualitativamente, o ambiente linguístico de

crianças e adolescentes bilingues sequenciais, constatam que o número de horas de exposição linguística a programas televisivos, a quantidade de *input* proveniente da leitura de livros, bem como o número de amigos falantes de cada língua-alvo e a quantidade de utilização de cada língua em contexto doméstico, são variáveis ambientais preditivas do desenvolvimento linguístico dos sujeitos analisados, em cada uma das línguas. Desta forma, os falantes bilingues inseridos em ambientes linguísticos nos quais a L2 predomina desenvolvem mais a competência linguística nesta língua, ao passo que os falantes bilingues inseridos em contextos linguísticos em que a L1 é predominantemente usada tendem a ser menos proficientes na L2 e a preservar a língua minoritária. Jia e colegas (2003, 2007) verificam ainda que a idade de chegada ao país de acolhimento está associada ao predomínio de uma das línguas em relação à outra nos ambientes linguísticos destes falantes bilingues, isto é, a L2 predomina nos contextos sociolinguísticos dos falantes mais jovens enquanto que a L1 é mais utilizada nos ambientes sociolinguísticos dos falantes mais velhos.

### **3.2.2 Sumário**

Os resultados dos estudos apresentados revelam que há uma característica comum e transversal ao desenvolvimento linguístico de todas as crianças bilingues – este está significativamente dependente do grau de exposição linguística a cada uma das línguas em aquisição.

Para além disso, os estudos mostram que a quantidade e a qualidade de exposição linguística estão estritamente correlacionadas com o contexto sociolinguístico e a constelação familiar de cada criança bilingue, sendo que estes variam muito de falante para falante. Por um lado, as crianças podem estar inseridas em contextos sociopolíticos em que as duas línguas são apoiadas pela sociedade e partilham o mesmo nível de prestígio. Nestes casos, as crianças estão expostas, em cada uma das suas línguas, a múltiplas fontes de *input*, a registos de língua diversificados e ao ensino formal, tendo ainda a oportunidade de as utilizar em diversos contextos comunicacionais. Por outro lado, as crianças podem estar inseridas em comunidades bilingues, geralmente imigrantes, em que uma das línguas é minoritária, sendo esta menos apoiada pela sociedade maioritária e não usufruindo do mesmo prestígio que a língua dominante. Nestes casos, as crianças são expostas a cada uma das suas línguas em contextos comunicacionais diferenciados. A língua maioritária é utilizada em contextos

multifacetados, enquanto que a língua minoritária é falada em contextos muito restritos, geralmente circunscrita ao ambiente familiar ou à comunidade imigrante, cuja representação demográfica no país de acolhimento pode ser muito significativa ou, então, muito reduzida.

No que respeita ao contexto familiar, este é, tal como na aquisição linguística monolíngue, a principal fonte de *input* das crianças bilingues, sobretudo em estádios iniciais de desenvolvimento linguístico. Como comprovam os estudos apresentados, a quantidade de exposição linguística que uma criança bilingue recebe de cada língua varia de acordo com a constituição da família nuclear e com a língua utilizada por cada um dos seus membros. A exposição linguística a ambas as línguas pode ser, por um lado, bastante equilibrada ou, por outro lado, bastante assimétrica. As crianças bilingues podem receber *input* de uma língua de ambos os progenitores ou podem ser expostas a uma língua através de um progenitor e à outra através do outro, sendo também comum os dois progenitores utilizarem ambas as línguas nas interações verbais com os filhos. Por outro lado, a presença de irmãos no agregado familiar, sobretudo se forem mais velhos, parece igualmente influenciar a quantidade de exposição linguística a que as crianças bilingues têm acesso no seio familiar. Apesar de estes constituírem mais uma fonte de *input* para os mais novos, os irmãos mais velhos estão, geralmente, associados à alteração dos hábitos linguísticos familiares, sendo, na maioria das vezes, os responsáveis pela introdução da língua maioritária no seio de famílias em que ambos os pais são falantes de uma língua minoritária e esta é a única língua de comunicação em casa. A par dos pais e dos irmãos, também outros membros familiares, como os avós, parecem ter um papel determinante e positivo na quantidade de *input* que as crianças recebem de uma língua minoritária, contribuindo para a transmissão intergeracional e manutenção da mesma.

Os estudos revistos mostram ainda que a própria produção linguística das crianças bilingues é uma variável preditiva do seu desenvolvimento linguístico. As crianças que se exprimem com mais frequência numa língua-alvo têm um nível de desempenho linguístico superior ao de crianças que se exprimem com menos frequência nessa mesma língua. Tal como Bohman *et al.* (2010) sugerem, “using a language (i.e. output) forces the learner to process the language in a way that only hearing it (i.e. input) does not” (p. 339).

Para além da quantidade, também a qualidade da exposição linguística é apresentada como sendo uma variável preditiva do desempenho linguístico de crianças bilingues. Assim, se, por um lado, a interação com falantes nativos e o acesso a fontes de

*input* mais próximas da norma padrão das línguas-alvo (o ensino formal, os *media*, a literatura, entre outras) influenciam positivamente o desenvolvimento linguístico das crianças bilíngues; por outro lado, o contacto com falantes não nativos e com registos de língua pouco diversificados podem exercer efeitos negativos na competência linguística das mesmas. De facto, no seu quotidiano, as crianças bilíngues contactam não só com falantes não nativos das línguas-alvo, cujo nível de proficiência pode variar do muito baixo ao muito elevado, como também com diferentes gerações de imigrantes que, por seu turno, podem apresentar efeitos de erosão linguística na língua nativa e um fraco domínio da língua da sociedade de acolhimento, fazendo com que o *input* a que as crianças bilíngues têm acesso não seja, muitas vezes, representativo da variedade linguística falada pelas comunidades monolíngues das línguas-alvo.

Por fim, uma vez que o conhecimento lexical de crianças bilíngues tende a estar distribuído por diferentes contextos de comunicação, de acordo com a língua falada pelos seus interlocutores, a interação com diferentes falantes das línguas em aquisição, em contextos comunicacionais diversificados, pode contribuir positivamente para a variedade e a riqueza lexical das crianças bilíngues, tornando-se fundamental para um desenvolvimento lexical bilingue equilibrado.



## Capítulo 4 – O presente estudo

No presente estudo, analisou-se a competência lexical de 44 crianças com idades compreendidas entre os seis e os 11 anos, agrupadas em dois grupos distintos: um grupo experimental constituído por 23 falantes de herança do PE a frequentarem aulas de Português Língua de Herança (PLH) na Alemanha e um grupo de controlo composto por 21 falantes monolíngues do PE, tendo todos os participantes deste grupo participado num estudo anterior (projeto *EFFE – Escreves como falas – falas como escreves?*<sup>6</sup>), do qual se extraíram e analisaram os dados. Para além de se procurar verificar se as crianças bilingues lusodescendentes têm um desempenho linguístico idêntico ao das crianças monolíngues do PE, procurou-se ainda identificar fatores preditivos do desenvolvimento da competência lexical dos falantes de herança.

Deste modo, neste capítulo, apresentar-se-ão detalhadamente não só os materiais e procedimentos adotados na recolha de dados, como também o perfil sociolinguístico dos dois grupos de participantes e as questões de investigação que fundamentam este estudo.

### 4.1 Recolha de dados

#### 4.1.1 Autorização

Todos os Encarregados de Educação do grupo de falantes de herança assinaram uma declaração na qual consentem a participação dos seus educandos na investigação e autorizam a utilização dos dados recolhidos para efeitos de investigação, desde que o anonimato dos participantes fosse garantido (ver Anexo A).

#### 4.1.2 Tarefa de produção oral semi-espontânea

De modo a se constituir um *corpus* de produções orais semi-espontâneas para as análises comparativas do desempenho lexical entre ambos os grupos em observação, o grupo de falantes de herança participou numa tarefa de produção oral semi-espontânea

---

<sup>6</sup> O projeto *EFFE* é um *corpus* de dados de escrita e fala de crianças portuguesas nos primeiros anos de escolaridade, que começou a ser desenvolvido em 2012 no Centro de Linguística da Universidade de Lisboa (CLUL) e que está disponível no sítio de internet: <http://alfclul.clul.ul.pt/teitok/effe/index.php>.

semelhante à realizada pelos investigadores do projeto *EFFE* com os participantes do grupo de controlo.

Assim, para a tarefa experimental, utilizou-se a imagem-cenário *Sala* (ver Anexo B), uma das cinco imagens temáticas usadas no projeto. Originalmente desenvolvidas por Yavas *et al.* (1991) para a avaliação fonológica de crianças falantes do português do Brasil (PB) com idade superior a três anos, estas imagens temáticas, *Casa de Banho*, *Cidade*, *Cozinha*, *Floresta* e *Sala*, foram adaptadas para o PE por Guerreiro (2007) e selecionadas pelo grupo de investigação do projeto *EFFE* para as tarefas de produção escrita e oral aplicadas a crianças monolíngues do PE nos primeiros anos de escolarização. A seleção destas imagens por estes investigadores deve-se ao facto de estas incluírem figuras-estímulo para a nomeação de palavras-alvo e descrição de cenários temáticos (Lourenço-Gomes *et al.*, 2016) relevantes para a constituição do *corpus* de dados de escrita e de fala de crianças portuguesas nos primeiros anos de escolaridade. Por sua vez, a opção pela imagem-cenário *Sala*, para o presente estudo, prende-se com o facto de esta incluir figuras-estímulo propícias à produção de vocabulário relacionado com o contexto doméstico, ambiente no qual a aquisição do português como LH, geralmente, ocorre.

Após a seleção da imagem-cenário a ser utilizada na tarefa experimental, analisou-se a metodologia utilizada pelo grupo de investigação do projeto *EFFE* na recolha dos dados orais do grupo de controlo de modo a ser adaptada e aplicada ao grupo dos falantes de herança.

Os participantes realizaram a tarefa em sessões individuais, num único momento temporal e sem limite de tempo, em salas anexas aos locais onde decorriam as aulas de PLH, nas cidades de Esslingen am Neckar, Heilbronn, Schwäbisch Gmünd e Weilheim an der Teck, na Alemanha. Os dados foram registados em formato áudio (.wav) através de um gravador digital com microfone interno (Edirol R-09HR). No início de cada sessão, cada criança recebeu a imagem-cenário *Sala* em suporte papel (formato A4), tendo-lhe sido apresentada, como estímulo à tarefa descritivo-narrativa, uma situação fictícia, na qual esta teria que imaginar que chegava a casa e encontrava a cena representada na imagem. Era-lhe, então, solicitado que a descrevesse, indicando ao mesmo tempo possíveis razões para o cenário encontrado. No decorrer da tarefa, cada criança foi sendo incentivada a nomear e a descrever elementos presentes na imagem, através de perguntas como *O que vês aqui?*, *O que está o pai a fazer?*, *Como está a sala?*, entre outras questões, de modo a que os participantes produzissem o maior número de itens lexicais possíveis. Todos os dados orais foram recolhidos pela autora deste estudo, falante nativa do PE,

tendo todos os participantes sido informados de que a mesma não falava alemão, de forma a se evitar que estes recorressem à língua alemã durante a tarefa experimental.

#### 4.1.3 Constituição do *corpus*

Para a constituição do *corpus* de produções orais semi-espontâneas, transcreveram-se ortograficamente as gravações áudio do grupo de FH e do grupo de controlo para formato *.docx*. De seguida, selecionaram-se os itens lexicais pertencentes às categorias gramaticais nomes, verbos e adjetivos, tendo estes sido simultaneamente distribuídos por bases de dados, em formato *.xlsx*, correspondentes a cada uma das suas categorias. No decorrer deste processo, lematizaram-se as formas flexionadas dos nomes, verbos e adjetivos, de modo a constituírem uma única entrada nas bases de dados. Assim, no que respeita aos verbos, todas as suas formas flexionadas surgem sob a forma do infinitivo; quanto aos nomes, estes aparecem na base de dados na sua forma do singular, sendo que os nomes biformes, como *gato/gata*, surgem na forma do masculino; por sua vez, tal como na categoria nomes, as formas dos adjetivos no plural foram reduzidas ao singular e, no caso dos adjetivos biformes, estes surgem na sua forma masculina.

Para além de ir ao encontro de outros estudos sobre a avaliação da competência lexical produtiva, nos quais o lema é considerado a unidade de medição lexical mais apropriada (ver Nation (2007) e Schmitt (2010) para uma discussão geral), a opção pelo processo de lematização do vocabulário proferido pelos participantes deste estudo prende-se com o facto de se estar a avaliar o seu conhecimento lexical e não o morfossintático. Portanto, não é relevante, para este estudo, a aferição do conhecimento das diferentes formas flexionadas das palavras lematizadas.

Deste modo, independentemente de um participante ter utilizado no seu discurso diferentes formas flexionadas de um item lexical, estas foram contabilizadas apenas uma vez no cálculo do seu tamanho lexical, ou seja, a um participante que tenha produzido três formas flexionadas do verbo *cantar* (por exemplo, *canta*, *cantou* e *cantava*) apenas lhe foi contabilizado uma vez o verbo *cantar* para o cálculo do tamanho do seu vocabulário.

De salientar ainda que palavras utilizadas metonimicamente, como *Playstation* por *consola de jogos* ou *Jogo da Glória* por *jogo de tabuleiro*, bem como estrangeirismos, como *bordeaux* e *puzzle*, constituem entradas nas bases de dados. Já no que respeita a unidades multilexicais, como *dar conta de*, estas constituem uma única entrada no *corpus*

por constituírem um bloco semântico. Por sua vez, todos os itens lexicais produzidos em língua estrangeira (i.e. alemão e inglês), que não são aceites como estrangeirismos na língua portuguesa, foram excluídos da constituição do *corpus*.

Assim sendo, constituíram-se inicialmente dois *corpora*, um do grupo de FH e outro do grupo de FM, compostos por três *subcorpora* correspondentes às categorias gramaticais nomes, verbos e adjetivos, que permitiram calcular a dimensão lexical de cada grupo em análise. Posteriormente, integraram-se ambos os *corpora* num único *corpus* que serviu de base às análises comparativas do desempenho lexical inter e intragrupal.

#### **4.1.4 Questionário sociolinguístico**

Uma vez que o contexto sociolinguístico em que ocorre a aquisição de línguas de herança é o que determina o perfil dos seus falantes, os questionários sociolinguísticos são considerados um instrumento fundamental nos estudos sobre os falantes de herança (Montrul, 2016) e têm sido utilizados num leque variado de investigações sobre a aquisição linguística bilingue (Cha e Goldenberg, 2015; De Houwer, 2009; Flores *et al.*, 2016; Gutiérrez-Clellen e Kreiter, 2003; Jia e Fuse, 2007; Paradis, 2011; Unsworth, 2013; entre outros). De facto, através deste tipo de questionários, geralmente aplicados aos pais dos participantes nos estudos, é possível a recolha não só de dados biográficos sobre os sujeitos em análise, como também de informações sobre as suas experiências sociolinguísticas dentro e fora do contexto doméstico, que se revelam essenciais para se determinar, por exemplo, a quantidade de *input* linguístico que os falantes bilingues recebem em cada uma das suas línguas, bem como a variedade e a qualidade das fontes de exposição linguística a que têm acesso (i.e. interlocutores nativos/não nativos; contexto doméstico/escolar; etc.), que são tidos como fatores intervenientes no processo de aquisição bilingue.

Assim sendo, elaborou-se um questionário sociolinguístico detalhado, adaptado dos questionários sociolinguísticos de Gutiérrez-Clellen e Kreiter (2003) e Unsworth (2013), composto por três secções: as duas primeiras centradas em informações biográficas e sociolinguísticas sobre os pais dos FH e uma última parte centrada no FH (ver Anexo C). O questionário sociolinguístico foi, então, aplicado a um dos pais de cada falante de herança.

No que respeita às secções sobre os dados parentais, o questionário inclui questões sobre a idade, o país de nascimento, a profissão e o nível de escolaridade. Para além disso, o questionário integra também questões mais específicas de acordo com o país de nascimento dos pais, ou seja, no caso de os pais não terem nascido no país de acolhimento (i.e. na Alemanha), solicita-se que estes mencionem o número de anos a residirem nesse país. Por sua vez, no caso dos pais que já nasceram na Alemanha, questiona-se sobre o país de origem dos seus progenitores. Em ambos os contextos, isto é, terem nascido em Portugal ou no país de acolhimento, é pedido que informem sobre a região portuguesa de origem.

Em relação à secção dedicada às crianças, falantes de herança, para além de informações sobre a idade, o país de nascimento, o número de anos a residirem na Alemanha (no caso de não terem nascido no país de acolhimento), a existência de irmãos e a idade dos mesmos, solicita-se no questionário não só que os pais informem sobre a idade e o contexto (i.e. doméstico/familiar ou escolar) em que os FH foram expostos pela primeira vez a cada uma das suas línguas, como também que indiquem as pessoas com quem as crianças bilingues contactam regularmente e a frequência com que estas pessoas interagem em português com as mesmas e vice-versa. Para além disso, os pais forneceram informações sobre a regularidade com que os FH passam férias em Portugal e a frequência com que utilizam a língua portuguesa neste contexto, bem como sobre a(s) língua(s) utilizada(s) na comunicação com amigos falantes de português em contexto de emigração, o número de anos a frequentar aulas de PLH e o acesso regular a outras fontes de *input* em português (i.e. televisão, rádio, livros, etc.).

Deste modo, o questionário permitiu recolher informações essenciais quer para a caracterização sociolinguística dos falantes de herança, quer também para a quantificação dos dados das variáveis preditivas do desenvolvimento linguístico bilingue, de forma a poderem ser correlacionadas estatisticamente com os resultados do desempenho linguístico destes falantes na LH. Assim sendo, com base no questionário apresentado, calculou-se: (1) a quantidade de exposição linguística à LH em contexto doméstico, através da soma da quantidade de *input* recebido de cada progenitor, cujo valor pode variar entre o mínimo de zero e o máximo de 100, resultando numa escala de zero a 200; (2) a quantidade de produção linguística na LH em contexto doméstico, através da soma da quantidade de *output* para cada progenitor, utilizando-se uma escala de medição idêntica à anterior; (3) a quantidade de progenitores falantes do PE, através da soma do número de progenitores que têm o PE como língua materna, variando os resultados

dicotomicamente entre *ambos* e *mãe ou pai*; (4) a quantidade de parceiros conversacionais falantes nativos do PE em contexto de emigração (i.e. avós e tios), através da soma do número de locutores nativos que contactam frequentemente em português com os FH; e (5) a quantidade de atividades de lazer (i.e. ver televisão, ouvir música, navegar na internet, entre outras) em que a LH é utilizada, através da soma do número de atividades lúdicas em que os FH participam com regularidade, usando a língua portuguesa. Através dos dados fornecidos no questionário foi também possível verificar se variáveis como (6) a existência de irmãos mais velhos e (7) a idade dos FH, no momento da recolha de dados, têm influência no desempenho lexical das crianças bilingues na sua LH.

## **4.2 Participantes**

### **4.2.1 Falantes de herança**

A amostra dos falantes de herança foi selecionada de um conjunto de alunos a frequentar, em regime de ensino paralelo e com a periodicidade de 90 minutos semanais, aulas de língua e cultura portuguesa, organizadas e apoiadas pela rede de Ensino de Português no Estrangeiro (EPE) do Instituto Camões. Como critérios de seleção dos participantes, foram definidos os seguintes parâmetros: (1) ter pais (um ou ambos) portugueses ou de descendência portuguesa; (2) ter nascido ou emigrado até aos cinco anos de idade para o país de acolhimento; (3) ter entre os seis e os 12 anos de idade; e (4) ter a autorização do Encarregado de Educação para participar no estudo.

Deste modo, selecionaram-se 23 crianças lusodescendentes residentes na Alemanha (região de Estugarda), 12 do sexo masculino (52,2%) e 11 do feminino (47,8%), com idades compreendidas entre os seis e os 11 anos (Média ( $M$ ) = 8,48; Desvio-Padrão ( $DP$ ) = 1,44).

Uma vez que a heterogeneidade e a variabilidade são características dos contextos em que a aquisição linguística bilingue ocorre, o perfil sociolinguístico destes falantes foi traçado através do questionário parental detalhado apresentado na secção anterior.

Todos os participantes cresceram e vivem num ambiente bilingue, no qual o PE é língua de herança, usada em contexto doméstico, e o alemão é a língua maioritária, ou seja, a língua de escolarização e da sociedade. De salientar ainda que dois dos participantes, o FH\_4 e o FH\_20, partilham juntamente com o PE, o italiano e o romeno, respetivamente, como línguas de herança. Das 23 crianças apenas duas nasceram em

Portugal, tendo, porém, emigrado para a Alemanha ainda no primeiro ano de vida. Quanto ao início de exposição linguística a cada uma das suas línguas, todas as crianças foram expostas ao português desde a nascença, tendo 16 crianças sido também expostas ao alemão antes dos três anos de idade. Por sua vez, sete crianças começaram apenas a receber exposição linguística mais intensiva à língua maioritária entre os três e os cinco anos, quando começaram a frequentar instituições de ensino pré-escolares.

A maioria dos falantes ( $n = 19$ ; 82,6%) é de famílias em que ambos os pais têm o português como língua materna, sendo que, destas 19 famílias, em oito, os dois progenitores nasceram em Portugal; em nove, apenas um progenitor nasceu em Portugal, tendo o outro nascido já na Alemanha; e, em duas, ambos os progenitores nasceram na Alemanha. De salientar ainda que, das oito famílias em que ambos os pais nasceram em Portugal, em cinco, quer o pai, quer a mãe, emigraram já em fase adulta para a Alemanha; em duas, um dos progenitores emigrou durante a adolescência (i.e. entre os 12 e os 18 anos) e o outro em fase adulta; e, em uma família, ambos os pais emigraram ainda adolescentes. Por sua vez, das nove famílias em que apenas um dos pais nasceu em Portugal, em seis, o progenitor emigrou em fase adulta e, em três, durante a adolescência. As restantes crianças ( $n = 4$ ; 17,4%) vivem em famílias em que apenas um dos pais é português. Destas quatro famílias, em duas, dois dos progenitores emigraram para a Alemanha no primeiro ano de vida e, nas outras duas, um emigrou durante a adolescência e o outro em fase adulta. Ainda relativamente à constituição familiar destas crianças bilingues, a maioria ( $n = 21$ ; 91,3%) tem irmãos, sendo que apenas dez têm irmãos mais velhos.

De acordo com as constelações familiares apresentadas, verificou-se que, em contexto doméstico, a quantidade de exposição linguística à LH varia de acordo com o *background* sociolinguístico da família, ou seja, as crianças cujos pais têm ambos o português como língua materna recebem mais *input* linguístico nesta língua ( $M = 164,47$ ;  $DP = 37,56$ ) do que as crianças com apenas um pai falante do português ( $M = 68,75$ ;  $DP = 31,46$ ).

Relativamente ao nível de escolaridade dos pais das crianças bilingues, este varia entre o 1.º Ciclo do Ensino Básico e o Ensino Superior, sendo que a maioria se situa entre o 2.º Ciclo do Ensino Básico e o Ensino Secundário.

Uma fonte de exposição linguística comum a todas as crianças bilingues desta amostra é o ensino formal do português. Contudo, devido ao seu carácter extracurricular, o número de anos de escolarização em português destes falantes varia frequentemente,

inclusive, dentro da mesma faixa etária, o que revela que nem todos os FH começam a frequentar as aulas de português na idade em que, geralmente, as crianças monolíngues portuguesas iniciam a sua escolarização (i.e. aos seis anos). Assim, das 23 crianças sob investigação, 22 começaram a ter aulas de PLH entre os seis e os sete anos de idade e apenas uma criança de 11 anos começou a frequentar as aulas de português aos nove anos de idade.

Para além da exposição linguística em contexto doméstico e nas aulas extracurriculares de PLH, 15 das 23 crianças bilingues contactam frequentemente com outros falantes do PE, geralmente, familiares próximos (i.e. avós e tios) que se encontram igualmente em contexto de emigração. Quanto à(s) língua(s) de comunicação com amigos falantes do PE também em contexto de emigração, a maioria dos FH ( $n = 14$ ; 60,9%) utiliza mais o alemão do que o português. Relativamente ao restante da amostra, uma criança utiliza apenas o português; duas interagem mais em português do que em alemão; outras duas usam quer o português, quer o alemão, em quantidades idênticas; e quatro comunicam apenas em alemão.

Outras fontes de exposição linguística, como os *media* e a literatura, i.e. fontes mais próximas da língua padrão, foram também mencionadas nos questionários, sendo que a maioria das crianças ( $n = 21$ ; 91,3%) recebe *input* linguístico destas fontes em português.

De salientar ainda que 11 das 23 crianças bilingues passam sempre férias em Portugal, nove passam quase sempre e três apenas às vezes. Quando se encontram de férias em Portugal, a maioria dos FH ( $n = 13$ ; 56,5%) utiliza sempre a língua portuguesa como meio de comunicação, quatro utilizam-na frequentemente, três somente às vezes interagem em português e uma raramente utiliza a LH. Quanto às duas crianças restantes, os pais não forneceram essa informação no questionário.

Os dados sociolinguísticos de cada participante (i.e. idade, sexo, número de progenitores falantes do PE, região portuguesa de origem, idade de emigração dos progenitores, presença de irmãos mais velhos no agregado familiar, quantidade de *input* e de *output* em contexto doméstico, número de parceiros conversacionais falantes do PE, e outras fontes de *input* mais próximas da norma padrão do PE) podem ser consultados na Tabela D1, no Anexo D.

#### **4.2.2 Falantes monolíngues**

O grupo de controlo foi seleccionado de um conjunto de 92 crianças do 2.º ano de escolaridade que participou no projeto *EFFE – Escreves como falas – falas como escreves?* (ver Alves *et al.*, 2015; Lourenço-Gomes *et al.*, 2016), desenvolvido pelo Centro de Linguística da Universidade de Lisboa (CLUL). Todas as crianças seleccionadas participaram na tarefa experimental de descrição oral da imagem-cenário *Sala* (Anexo B), tendo sido este o critério de selecção. Deste modo, constituiu-se um grupo de controlo com 21 falantes monolíngues do PE, 11 do sexo masculino (52,38%) e 10 do feminino (47,62%), com idades compreendidas entre os sete e os 10 anos ( $M = 7,20$ ;  $DP = 0,70$ ), residentes nas áreas geográficas de Lisboa e do Porto.

Uma vez que o grupo de controlo apresenta diferenças socioculturais de acordo com o local de recolha dos dados, optou-se por se delinear o perfil destes falantes consoante a sua área geográfica.

Assim, dez das 21 crianças monolíngues, seis do sexo masculino e quatro do feminino, estavam matriculadas no 2.º ano de escolaridade em escolas da rede pública da cidade do Porto. A idade deste subgrupo de falantes situa-se entre os sete e os 10 anos e o nível de escolaridade dos seus Encarregados de Educação varia entre o 1.º Ciclo do Ensino Básico e o Ensino Superior, havendo um caso em que ambos os pais são analfabetos.

As restantes 11 crianças monolíngues, cinco do sexo masculino e seis do feminino, estavam matriculadas no 2.º ano de escolaridade de uma escola de ensino privado da cidade de Lisboa. Todas as crianças deste subgrupo estão na faixa etária dos sete anos e o nível de escolaridade dos seus Encarregados de Educação varia entre o 3.º Ciclo do Ensino Básico e o Ensino Superior, sendo que a grande maioria se situa neste último grau de escolarização.

#### **4.3 Questões de investigação**

Tendo como objetivos centrais verificar, empiricamente, se crianças bilingues lusodescendentes têm um desempenho lexical semelhante ao dos seus pares monolíngues do PE e, cumulativamente, identificar fatores preditivos do desenvolvimento lexical daquelas, o presente estudo fundamenta-se nas seguintes questões de investigação:

**Q 1:** As crianças bilíngues lusodescendentes com idades compreendidas entre os seis e os 11 anos têm um desempenho linguístico lexical idêntico ao de crianças monolíngues do PE a frequentarem o 1.º Ciclo do Ensino Básico?

**Q 1.1:** Isto é, as crianças bilíngues utilizam uma quantidade de itens lexicais diferentes (i.e. lemas) idêntica à dos seus pares monolíngues na descrição de imagens semelhantes?

**Q 1.2:** É possível identificar semelhanças/diferenças intra e intergrupais quanto à distribuição quantitativa dos itens lexicais produzidos nas categorias gramaticais nomes, verbos e adjetivos?

**Q 2:** Qual é o papel que diferentes fatores de *input*, relacionados com a quantidade e a qualidade da exposição linguística, desempenham no desenvolvimento da competência lexical dos falantes de herança?

**Q 2.1:** Será que as crianças que têm uma exposição maior e mais diversificada ao PE têm um desempenho lexical superior ao de crianças com exposição mais reduzida e menos diversificada?

**Q 2.2:** Se existem diferenças entre as crianças bilíngues relacionadas com o tipo de exposição, quais as variáveis que mais influenciam o desenvolvimento lexical, quer relativamente ao repertório lexical total, quer em relação ao desempenho por categoria gramatical?

No que respeita à primeira questão de investigação (Q 1) e respetivas subquestões (Q 1.1 e Q 1.2), estas centram-se no desenvolvimento lexical das crianças bilíngues português-alemão em comparação com o do grupo de crianças monolíngues do PE, através da análise, a nível intra e intergrupais, dos dados obtidos na tarefa de produção oral semi-espontânea.

Tendo em conta os resultados dos inúmeros estudos centrados no desenvolvimento lexical de crianças bilíngues, os quais revelam que estas tendem a apresentar repertórios lexicais quantitativamente inferiores aos dos seus pares monolíngues, quando apenas uma das suas línguas é avaliada (Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; Hoff *et al.*, 2012; Pearson *et al.*, 1993; Poulin-Dubois *et al.*, 2013; entre outros), prevê-se que, na tarefa experimental, os falantes de herança do

PE produzam uma quantidade de itens lexicais diferentes (i.e. lemas) inferior à das crianças monolíngues da língua-alvo, não só no que respeita à produção lexical total (i.e. *corpus* total), como também no que concerne à produção lexical por categoria gramatical (i.e. *subcorpora* nomes, verbos e adjetivos).

Por outro lado, com base nos estudos apresentados sobre a composição do léxico inicial de crianças bilingues (David e Wei, 2005; dos Santos e Kern, 2013), os quais demonstram que estas apresentam padrões de distribuição lexical idênticos aos das correspondentes crianças monolíngues, prevê-se que os repertórios lexicais das crianças bilingues não difiram significativamente dos das crianças monolíngues do PE no que diz respeito à proporção de itens lexicais pertencentes às categorias gramaticais dos nomes e dos predicados, designadamente verbos e adjetivos, pressupondo-se igualmente que, à semelhança de estudos com crianças bilingues e monolíngues de diversas comunidades linguísticas (Bates *et al.*, 1994; Bornstein *et al.*, 2004; Conboy e Thal, 2006; David e Wei, 2005; Jackson-Maldonado *et al.*, 1993; Özcan *et al.*, 2016; dos Santos e Kern, 2013; entre outros), ambos os grupos em análise produzam mais nomes do que verbos e adjetivos.

Por sua vez, no que respeita à segunda questão de investigação (Q 2) e respetivas subquestões (Q 2.1 e Q 2.2), estas centram-se no papel que os diferentes fatores de *input*, relacionados com a quantidade e a qualidade da exposição linguística, desempenham no desenvolvimento lexical do grupo de falantes de herança do PE.

Partindo da predição de que o conhecimento lexical das crianças bilingues, avaliado na tarefa experimental, será marcado pela variação a nível individual, havendo crianças que se aproximarão dos padrões de desenvolvimento lexical monolíngue e outras que destes se afastarão, prevê-se que fatores extralinguísticos relacionados com a quantidade e a qualidade da exposição linguística tenham um papel determinante no desenvolvimento da sua competência lexical.

Assim, tal como os estudos apresentados na secção 3.2 revelam, espera-se que a competência lexical dos falantes de herança do PE seja significativamente influenciada pelos seguintes fatores:

(1) *quantidade de input do PE em contexto doméstico*: prevê-se que as crianças que estão linguisticamente mais expostas ao PE em casa (através dos progenitores) tenham um desempenho lexical superior ao das crianças que recebem menos exposição linguística desta língua, neste contexto;

(2) *quantidade de output em PE em contexto doméstico*: prevê-se que as crianças que se expressam com mais frequência em PE em casa (com os progenitores) tenham um

desempenho lexical superior ao das crianças que se exprimem com menos frequência nesta língua, neste contexto;

(3) *quantidade de progenitores falantes do PE*: prevê-se que as crianças cujos progenitores são ambos falantes do PE tenham um desempenho lexical superior ao das crianças que possuem apenas um progenitor falante do PE;

(4) *quantidade de parceiros conversacionais falantes do PE em contexto de emigração*: prevê-se que as crianças que contactam regularmente com um maior número de interlocutores falantes nativos do PE (i.e. tios e avós) em contexto de emigração tenham um desempenho lexical superior ao das crianças cujo contacto com estas fontes de *input* do PE é mais reduzido ou, inclusivamente, inexistente;

(5) *quantidade de atividades de lazer em que o PE é utilizado*: prevê-se que as crianças que participam regularmente num maior número de atividades lúdicas (i.e. ver televisão, navegar na internet, leitura de livros, entre outras) em que o PE é utilizado tenham um desempenho lexical superior ao das crianças cujo contacto com estas fontes de *input* mais próximas da norma padrão do PE é mais reduzido ou inexistente;

(6) *presença de irmãos mais velhos no agregado familiar*: prevê-se que as crianças que não têm irmãos mais velhos apresentem um desempenho lexical superior ao das crianças que os têm;

(7) *idade dos falantes de herança*: prevê-se que as crianças mais velhas tenham um desempenho lexical superior ao das crianças mais novas.

Apresentados os participantes, os instrumentos utilizados e a metodologia implementada na recolha dos dados, bem como as questões de investigação que fundamentam o presente estudo, apresentar-se-ão, no capítulo seguinte, os resultados que derivam da análise dos dados recolhidos.

## Capítulo 5 – Resultados

Neste capítulo, apresentar-se-ão os resultados das análises descritivas e estatísticas dos dados obtidos, quer na tarefa de produção oral semi-espontânea, quer no questionário sociolinguístico, que não só permitirão avaliar o desenvolvimento da competência lexical produtiva dos falantes de herança do PE, como também possibilitarão aferir a influência que os diferentes fatores de *input*, selecionados para este estudo, têm no desenvolvimento da mesma.

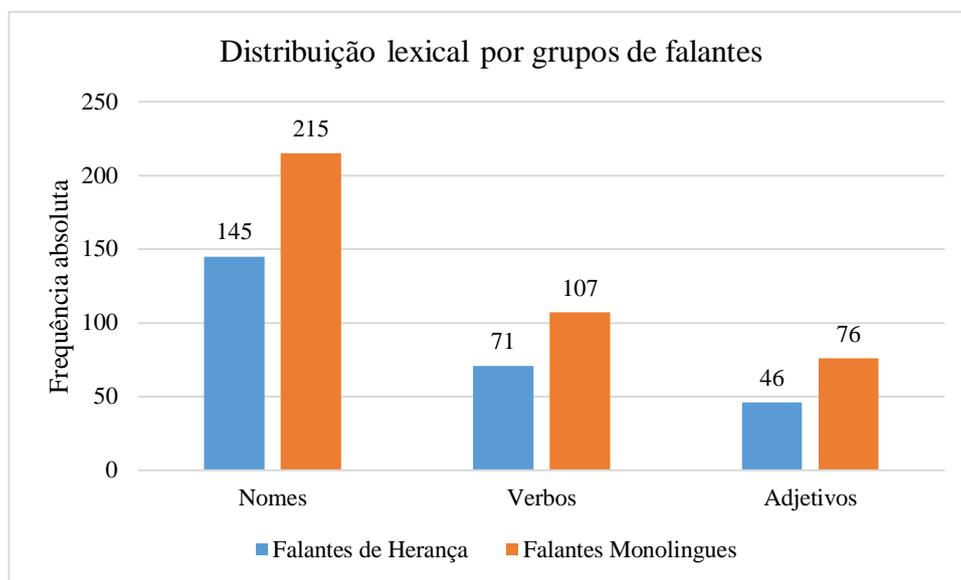
Assim, na primeira secção, com base nos dados recolhidos do grupo dos FH e do grupo dos FM do PE na tarefa experimental, começar-se-á pela apresentação comparativa da quantidade de itens lexicais, agrupados nas categorias gramaticais nomes, verbos e adjetivos, produzidos por ambos os grupos de participantes. De seguida, apresentar-se-ão os resultados da análise descritiva sobre a variação intra e intergrupar não só relativamente ao *corpus* total, como também por *subcorpora* (i.e. categorias gramaticais nomes, verbos e adjetivos). Por fim, a primeira secção terminará com a apresentação dos resultados da análise inferencial dos dados, realizada por meio de testes estatísticos de diferenças a nível inter e intragrupal, que permitirão concluir se as diferenças previamente observadas na análise descritiva são estatisticamente significativas ou não.

Por sua vez, na última secção, referente à relação entre os fatores de *input* e o desempenho lexical dos FH, iniciar-se-á pela apresentação da análise descritiva dos dados obtidos no questionário sociolinguístico, seguindo-se a apresentação dos resultados dos testes estatísticos de associação entre os fatores de *input* e o desempenho lexical dos FH na tarefa experimental.

### 5.1 Tarefa de produção oral semi-espontânea

O Gráfico 1 apresenta o número total (i.e. a frequência absoluta) das produções semi-espontâneas dos dois grupos de participantes por categoria gramatical. Uma análise detalhada do mesmo permite verificar que o *corpus* de produções orais semi-espontâneas do grupo de FH é constituído por 145 nomes, 71 verbos e 46 adjetivos num total de 262 itens lexicais. Por sua vez, o *corpus* referente ao grupo de FM é formado por 215 nomes, 107 verbos e 76 adjetivos num total de 398 itens lexicais. A combinação dos dois *corpora*, nos quais 91 nomes, 51 verbos e 31 adjetivos são comuns a ambos os grupos, resulta num *corpus* total de 487 itens lexicais composto por 269 nomes, 127 verbos e 91 adjetivos. A distribuição de itens lexicais por grupo de falantes, representada graficamente, permite

ainda concluir, de um modo geral, que o grupo de FM produz, em todas as categorias, mais itens lexicais do que o grupo de FH e que ambos os grupos produzem mais nomes do que verbos e mais verbos do que adjetivos.

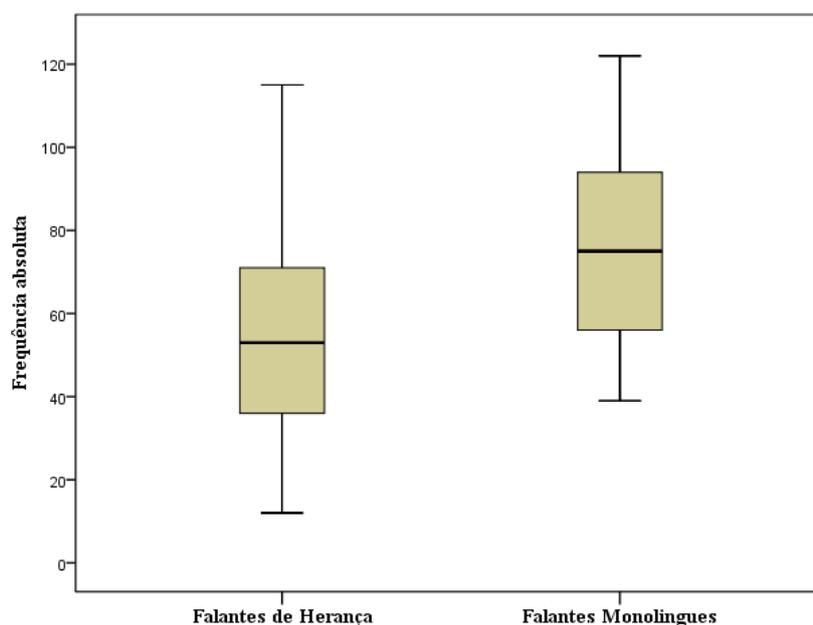


**Gráfico 1-** Quantidade de itens lexicais produzidos por grupo de falantes em cada categoria gramatical.

Quanto às médias de itens lexicais produzidos por cada um dos grupos de falantes, uma análise das suas frequências absolutas mostra que os FH apresentam uma média de 53,74 ( $DP = 25,21$ ) itens lexicais, ao passo que os FM têm uma média de 77,52 ( $DP = 23,41$ ). Já no que respeita às médias das frequências absolutas de itens lexicais por categoria gramatical, os FH produzem em média 28,74 ( $DP = 13,36$ ) nomes, 14,22 ( $DP = 7,21$ ) verbos e 10,78 ( $DP = 5,68$ ) adjetivos; por sua vez, os FM utilizam em média 44,57 ( $DP = 14,24$ ) nomes, 19,76 ( $DP = 6,49$ ) verbos e 13,19 ( $DP = 6,29$ ) adjetivos.

A distribuição dos dados, representada no Gráfico 2, permite realizar uma análise descritiva sobre a variação intra e intergrupar relativamente ao *corpus* total. Assim, pode verificar-se que, no que respeita aos FH, o centro da amostra (i.e. a mediana) se situa nos 53 itens lexicais, estando os valores da parte central dos dados (i.e. intervalo interquartil), correspondentes a 50% da amostra total, compreendidos entre os 32 e os 72 itens lexicais. A amostra apresenta ainda um valor mínimo de 12 e um máximo de 115 itens lexicais. Por sua vez, a mediana dos FM situa-se nos 75 itens lexicais, variando o intervalo interquartil (IIQ) entre os 55,5 e os 96 itens lexicais. Quanto aos valores menor e maior da amostra, estes correspondem a 39 e a 122 itens lexicais, respetivamente.

Deste modo, pode afirmar-se que, em todos os dados supramencionados, os valores relativos aos FM são superiores aos dos FH, apontando para diferenças no desenvolvimento lexical entre ambos os grupos. De salientar ainda que, embora ambos os grupos apresentem valores aproximadamente simétricos relativamente ao ponto central da amostra (i.e. a mediana) no intervalo interquartílico, quando os valores extremos (i.e. o máximo e o mínimo) são tidos em consideração, verifica-se que os valores dos FH apresentam um enviesamento para a parte superior, revelando uma maior dispersão dos dados na parte superior do que na inferior, o que significa que os valores correspondentes aos itens lexicais produzidos por estes falantes variam menos abaixo da mediana.



**Gráfico 2** - Caixa de Bigodes: distribuição e variação lexical por grupo de falantes relativamente ao *corpus* total (nomes, verbos e adjetivos).

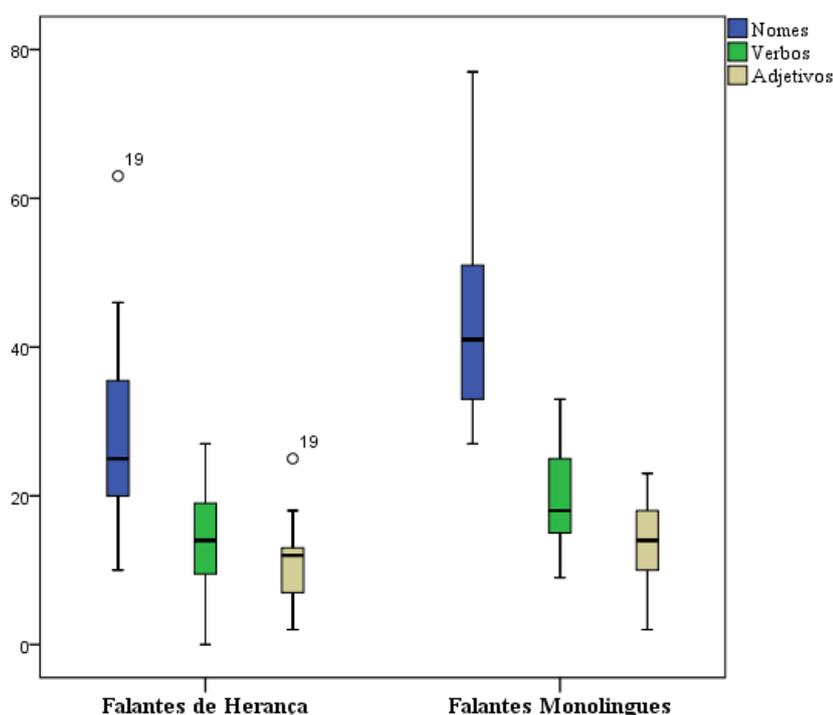
Por sua vez, a distribuição dos dados representada no Gráfico 3 e na Tabela 1 permite discriminar a variação previamente descrita no que respeita às categorias gramaticais nomes, verbos e adjetivos. Relativamente aos nomes e aos verbos, os FH apresentam valores inferiores aos FM em todas as medidas (i.e. mediana, mínimo, máximo e intervalo interquartil), embora as diferenças sejam mais salientes na categoria nomes. No entanto, verifica-se um comportamento diferente na categoria adjetivos, na qual as diferenças entre grupos são mais ténues, tendo os FH apresentado resultados ligeiramente superiores aos FM no que respeita ao valor máximo e idênticos

no valor mínimo. Uma informação importante que ressalta da análise do Gráfico 3 é a existência de um *outlier* (i.e. um valor atípico que se afasta do padrão geral dos dados) na distribuição dos dados referentes às categorias gramaticais nomes e adjetivos. Uma vez que a presença de um *outlier* influencia os resultados dos cálculos estatísticos, no processo de inferência estatística, realizar-se-ão, nas categorias nomes e adjetivos, testes com e sem este valor atípico, de modo a se avaliar a sua influência na análise e na interpretação dos resultados.

**Tabela 1** – Distribuição e variação lexical por grupo de falantes relativamente aos *subcorpora*.

	Nomes		Verbos		Adjetivos	
	FH	FM	FH	FM	FH	FM
<b>Mediana</b>	25	41	14	18	12	14
<b>Mínimo</b>	10	27	0	9	2	2
<b>Máximo*</b>	46/63	77	27	33	18/25	23
<b>IIQ</b>	20 - 36	32,5 - 55	9 - 20	15 - 25,5	6 - 13	8 - 18,5

*Nota.* \* Devido à existência de um *outlier*, nas categorias nomes e adjetivos, no grupo de FH, optou-se pela apresentação dos primeiro (correspondente ao *outlier*) e segundo valores mais elevados.



**Gráfico 3** – Caixa de Bigodes: distribuição e variação lexical por grupo de falantes relativamente aos *subcorpora*.

Antes de se proceder à análise estatística dos dados observados, realizou-se uma análise exploratória dos dados no programa de análise estatística *IBM SPSS*<sup>7</sup> para se averiguar se estes cumpriam os pressupostos subjacentes à utilização de testes paramétricos<sup>8</sup>. Assim, para se verificar a normalidade da distribuição dos dados, aplicaram-se os testes Kolmogorov-Smirnov e Shapiro-Wilk a todas as modalidades de análise (i.e. *corpus* total e *subcorpora*, com e sem *outlier*). Os resultados de ambos os testes revelam que há uma distribuição aproximadamente normal dos dados em todos os contextos de análise (ver resultados nas Tabelas E1 e E2, no Anexo E). Por sua vez, a homogeneidade das variâncias dos dois grupos em comparação é também confirmada através do teste de Levene para todos os contextos (ver resultados nas Tabelas E3 e E4, no Anexo E).

Deste modo, cumpridas as condições obrigatórias para a utilização de testes paramétricos, procedeu-se à análise estatística dos dados para se apurarem as diferenças de desempenho lexical a nível intra e intergrupar.

Assim, para se verificar se existem diferenças significativas entre os dois grupos no que respeita à quantidade de itens lexicais utilizados na tarefa experimental, aplicaram-se vários testes t para amostras independentes, não só ao *corpus* total, como também aos três *subcorpora* (i.e. nomes, verbos e adjetivos). Os resultados mostram que há diferenças significativas entre os FH e os FM relativamente ao *corpus* total ( $t(42) = -3.23, p = .002$ ), ou seja, o grupo de FH utiliza significativamente menos itens lexicais do que o grupo de FM. Por sua vez, no que respeita às diferenças entre grupos por categoria gramatical, os testes t revelam que existem diferenças estatisticamente relevantes nas categorias nomes ( $t(42) = -3.81, p < .001$ ) e verbos ( $t(42) = -2.67, p = .011$ ), sendo que os FH utilizam menos nomes e verbos do que os seus pares monolíngues. Contudo, no que se refere à categoria adjetivos, a análise estatística não revela diferenças significativas entre ambos os grupos ( $t(42) = -1.33, p = .189$ ). Uma vez que fora detetada a presença de um *outlier* nas categorias nomes e adjetivos, correram-se novamente testes t, desta vez sem este valor atípico, de modo a se verificar se a sua presença influenciava os resultados da análise estatística. Os resultados demonstram que, mesmo com a ausência do *outlier*, continuam a existir diferenças significativas entre os grupos relativamente aos nomes ( $t(41) = -4.44,$

---

<sup>7</sup> Versão 23.

<sup>8</sup> De acordo com Martins (2011), os pressupostos para a utilização de testes paramétricos são: (1) normalidade da distribuição dos dados, (2) homogeneidade das variâncias e (3) carácter intervalar da variável.

$p < .001$ ) e que, no que respeita aos adjetivos, ambos os grupos não diferem significativamente entre si ( $t(41) = -1.79, p = .082$ ).

Após a análise estatística das diferenças intergrupais relativamente à quantidade de itens lexicais utilizados na descrição da mesma imagem, a análise centrou-se na verificação de diferenças/semelhanças intragrupo, ou seja, investigou-se se há diferenças significativas dentro de cada grupo relativamente à utilização das três categorias gramaticais. Assim, realizou-se uma análise de variância (ANOVA) para medidas repetidas, quer para o grupo de FH, quer para o de FM. Uma vez que o pressuposto da esfericidade, um pressuposto subjacente a este teste estatístico, não foi cumprido por nenhum dos grupos em análise ( $p < 0.05$ ), considerou-se o valor da correção de Greenhouse-Geisser na análise dos dados (Field, 2009; Martins, 2011). Deste modo, verificou-se que há diferenças significativas não só dentro do grupo dos FH ( $F(1.204, 26.489) = 82.18, p < .001$ ), como também dentro do grupo dos FM ( $F(1.389, 27.788) = 113.24, p < .001$ ), na utilização de nomes, verbos e adjetivos. Visto que este teste não especifica onde se encontram as diferenças, comparações por método *pairwise* de Bonferroni revelaram que, em ambos os grupos, a utilização de nomes é significativamente superior à utilização de verbos, sendo o uso destes também estatisticamente superior ao uso de adjetivos (ver resultados nas Tabelas F1 e F2, no Anexo F).

## **5.2 Relação entre os fatores de *input* e o desempenho lexical dos falantes de herança**

Para se avaliar a influência dos diferentes fatores preditivos do desenvolvimento linguístico bilingue no desempenho lexical dos FH na tarefa de produção oral semi-espontânea, começar-se-á pela apresentação dos dados recolhidos no questionário sociolinguístico relacionados com as variáveis selecionadas para o processo de inferência estatística. De modo a se evitarem resultados falaciosos nos testes de correlação, optou-se por se retirar o *outlier* da amostra, que passou a ser constituída por 22 participantes.

Assim sendo, relativamente à quantidade de exposição linguística à LH em contexto doméstico (doravante também *input\_pais*), verificou-se que esta varia entre o valor mínimo de 25 e o máximo de 200 ( $M = 151,14; DP = 50,28$ ), sendo que o valor da mediana é de 150, isto é, os valores de uma metade da amostra são iguais ou superiores a 150 e os da outra metade são menores ou iguais a esse valor. Por sua vez, no que respeita à quantidade de produção linguística na LH em contexto doméstico (doravante também

*output\_pais*), a amostra apresenta igualmente um valor mínimo de 25 e um máximo de 200 ( $M = 112,50$ ;  $DP = 63,03$ ). Contudo, a distribuição dos dados da amostra é diferente, sendo o valor da mediana de 100, ou seja, 50% dos dados da amostra são menores ou iguais a 100 e os outros 50% são maiores ou iguais a esse valor. No que se refere à quantidade de parceiros conversacionais falantes nativos do PE em contexto de emigração (doravante também *par\_conv\_PE*), o número de locutores nativos com quem os FH contactam frequentemente em português varia entre o mínimo de zero e o máximo de quatro ( $M = 1,64$ ;  $DP = 1,33$ ), apresentando a seguinte distribuição: sete dos participantes (31,8%) não têm contacto frequente com este tipo de falantes; um (4,5%) contacta apenas com um falante; nove (40,9%) contactam com dois falantes; três (13,6%) contactam com três falantes; e dois (9,1%) contactam com quatro falantes. Já no que diz respeito à quantidade de atividades de lazer em que a LH é utilizada (doravante também *ativ\_PE*), o número de atividades em que os FH participam regularmente varia entre o mínimo de zero e o máximo de cinco ( $M = 2,45$ ;  $DP = 1,22$ ), sendo que dois dos participantes da amostra (9,1%) não utilizam a LH em nenhuma atividade de lazer; um (4,5%) utiliza-a em apenas uma atividade; oito (36,4%) utilizam-na em duas; nove (40,9%) utilizam-na em três; e dois (9,1%) utilizam-na em cinco. Quanto à quantidade de progenitores falantes do PE (doravante também *pais\_PE*), em 86,4% da amostra ( $n = 19$ ) o PE é a língua materna de ambos os pais e em 13,6% ( $n = 3$ ) apenas um dos pais tem o PE como língua materna. Em relação à presença de irmãos mais velhos no agregado familiar (doravante também *irmãos*), 45,5% dos participantes ( $n = 10$ ) têm irmãos em faixas etárias superiores. Por fim, no que respeita à idade dos participantes (doravante também *idade*), esta varia entre os seis e os 11 anos ( $M = 8,36$ ;  $DP = 1,36$ ), apresentando a seguinte distribuição: dois dos participantes (9,1%) têm seis anos; quatro (18,2%) têm sete; cinco (22,7%) têm oito; oito (36,4%) têm nove; um (4,5%) tem dez; e dois (9,1%) têm onze. Os dados dos participantes, que permitiram quantificar as variáveis relativas aos fatores de *input*, encontram-se discriminados na Tabela D1, no Anexo D.

Antes de se analisar estatisticamente a influência dos sete fatores de *input*, selecionados para este estudo, no desempenho lexical dos FH, também se procedeu a uma análise exploratória dos dados destas variáveis para se apurar se, no caso das variáveis intervalares (i.e. *input\_pais*; *output\_pais*; *par\_conv\_PE*; *ativ\_PE* e *idade*), estes cumpriam os pressupostos subjacentes à utilização de testes paramétricos. Uma vez que os testes estatísticos serão a nível intragrupal, apenas foi necessário verificar o pressuposto referente à normalidade da distribuição dos dados através dos testes

Kolmogorov-Smirnov e Shapiro-Wilk. Os resultados destes testes revelam que, das cinco variáveis intervalares preditivas do desenvolvimento lexical bilingue, apenas os dados relativos à variável *idade* apresentam uma distribuição normal (ver resultados na Tabela E5, no Anexo E). Assim sendo, somente nas análises estatísticas com a variável *idade* se utilizarão testes de associação paramétricos, designadamente o Coeficiente de Correlação de Pearson ( $r$ ). Com as restantes variáveis intervalares que não apresentam uma distribuição normal dos dados será utilizado o teste de associação não paramétrico Coeficiente de Correlação de Spearman ( $r_s$ ) (de acordo com Field (2009)). Por sua vez, com as variáveis nominais (i.e. *pais\_PE* e *irmãos*) será utilizado o teste estatístico de associação Coeficiente de Correlação Ponto-Bisserial ( $r_{pb}$ ).

Desta forma, aplicados os testes de associação entre os diferentes fatores de *input* e o desempenho lexical dos FH quer em relação ao *corpus* total, quer relativamente aos três *subcorpora* (i.e. nomes, verbos e adjetivos), obtiveram-se os resultados representados na Tabela 2.

**Tabela 2** – Correlações entre os fatores de *input* e o desempenho lexical dos FH

Coeficiente de correlação	Fatores	<i>Corpus</i> total	<i>Subcorpora</i>		
			Nomes	Verbos	Adjetivos
$r$	<i>idade</i>	.40	.34	.43	.41
	<i>input_pais</i>	<b>.63**</b>	<b>.56**</b>	<b>.65**</b>	<b>.67**</b>
$r_s$	<i>output_pais</i>	<b>.60**</b>	<b>.55**</b>	<b>.67**</b>	<b>.54**</b>
	<i>par_conv_PE</i>	-.38	-.44	-.36	-.26
	<i>ativ_PE</i>	.39	.40	.42	.28
$r_{pb}$	<i>pais_PE</i>	<b>.46*</b>	.38	<b>.48*</b>	<b>.51*</b>
	<i>irmãos</i>	-.33	-.36	-.24	-.31

Nota.  $r$  = Pearson;  $r_s$  = Spearman;  $r_{pb}$  = Ponto-Bisserial; níveis de significância: \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Uma análise detalhada dos coeficientes de correlação e dos respetivos níveis de significância estatística revela que as variáveis *input\_pais* e *output\_pais* estão positivamente correlacionadas com o tamanho lexical dos FH (no *corpus* total e nos três *subcorpora*), ou seja, os FH que estão mais expostos linguisticamente ao PE e que mais interagem verbalmente na LH, em contexto doméstico, apresentam um domínio lexical significativamente maior do que os FH que têm uma exposição mais reduzida e que

comunicam menos na LH neste contexto. De igual modo, verifica-se a existência de correlações positivas significativas entre a variável *pais\_PE* e os resultados lexicais dos FH relativamente ao *corpus* total e aos *subcorpora* verbos e adjetivos (mas não no *subcorpus* nomes), isto é, os FH cujos pais têm ambos o PE como língua materna apresentam melhores resultados lexicais no *corpus* total e nos *subcorpora* verbos e adjetivos do que os FH que têm apenas um dos pais de origem portuguesa. Esta associação sugere que os FH cujos progenitores são ambos falantes do PE recebem mais *input* linguístico e produzem mais *output* na LH, tendo, conseqüentemente, um melhor desempenho lexical. De facto, um teste de associação entre a quantidade de exposição linguística à LH em contexto doméstico e o número de progenitores falantes do PE mostra que as duas variáveis estão positivamente correlacionadas ( $r_{pb} = .68, p < .001$ ). Por sua vez, um teste de associação entre a quantidade de produção linguística na LH em contexto doméstico e a quantidade de progenitores falantes do PE revela que há também uma correlação positiva significativa entre ambas as variáveis ( $r_{pb} = .46, p < .05$ ), embora mais fraca do que a anterior (i.e. *input\_pais/pais\_PE*). Por último, no que respeita às variáveis *idade*, *par\_conv\_PE*, *ativ\_PE* e *irmãos*, estas não apresentam correlações significativas com a competência lexical dos FH do PE.



## Capítulo 6 - Discussão

No capítulo anterior, apresentaram-se, num primeiro momento, os resultados da análise descritiva e do processamento estatístico dos dados orais, extraídos da tarefa de produção oral semi-espontânea, dos dois grupos de crianças falantes do PE. Ambas as análises, centradas na comparação, a nível inter e intragrupal, do desempenho lexical dos grupos sob investigação na tarefa experimental, permitiram avaliar a competência lexical produtiva das crianças bilingues na sua LH, isto é, no PE. Num segundo momento, apresentaram-se, inicialmente, os resultados da análise descritiva dos dados sociolinguísticos dos FH, extraídos do questionário parental detalhado, relativos aos fatores relacionados com a quantidade e a qualidade da exposição linguística a que os FH têm acesso, tendo estes sido, de seguida, correlacionados estatisticamente com o desempenho lexical dos FH do PE na tarefa experimental. Os resultados da análise estatística entre os fatores de *input* selecionados para este estudo e o desempenho lexical das crianças bilingues na LH revelaram que, por um lado, há fatores que estão significativamente associados à competência lexical produtiva dos FH, havendo também, por outro lado, fatores que com esta não apresentam correlações significativas.

Assim, tendo em conta os resultados apresentados no capítulo anterior, proceder-se-á, neste capítulo que se inicia, à discussão dos mesmos à luz das questões de investigação que orientaram e fundamentaram o presente estudo.

No que respeita à primeira questão de investigação (Q 1) e subquestão Q 1.1, os resultados das análises estatísticas, derivados de testes de diferenças a nível intergrupar, revelam que o desempenho lexical de crianças bilingues lusodescendentes com idades compreendidas entre os seis e os 11 anos difere do dos seus pares monolingues do PE a frequentarem o primeiro ciclo do ensino básico no país de origem, ou seja, o grupo de crianças falantes de herança do PE produz significativamente menos itens lexicais (i.e. lemas) do que o grupo de crianças falantes monolingues do PE na tarefa descritivo-narrativa da mesma imagem-cenário. De facto, a diferença inicialmente observada entre as médias dos itens lexicais totais produzidos pelo grupo de FH ( $M = 53,74$ ;  $DP = 25,21$ ) e pelo grupo de FM ( $M = 77,52$ ;  $DP = 23,41$ ) revelou ser estatisticamente significativa ( $t(42) = -3.23$ ,  $p = .002$ ). Assim sendo, tendo em consideração as questões de investigação supramencionadas, os resultados do presente estudo permitem concluir que, à semelhança de outros estudos sobre a dimensão lexical de crianças bilingues de diversos pares linguísticos (Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; Hoff *et al.*, 2012;

Pearson *et al.*, 1993; Poulin-Dubois *et al.*, 2013; entre outros), também as crianças bilingues português-alemão apresentam repertórios lexicais quantitativamente inferiores aos das correspondentes crianças monolingues quando apenas uma das suas línguas é avaliada, neste caso, o PE (i.e. a LH).

Por sua vez, e ainda no âmbito da primeira questão de investigação, concretamente no da subquestão Q 1.2, os resultados das análises estatísticas centradas na distribuição quantitativa dos itens lexicais produzidos por categoria gramatical (i.e. *subcorpora* nomes, verbos e adjetivos) revelam que as diferenças intergrupais observadas na dimensão total dos repertórios lexicais se manifestam sobretudo nas categorias nomes e verbos, mas não na dos adjetivos. Embora as médias dos itens lexicais produzidos por categoria gramatical apontassem para diferenças entre os grupos em comparação (com os FH a produzirem em média 28,74 ( $DP = 13,36$ ) nomes, 14,22 ( $DP = 7,21$ ) verbos e 10,78 ( $DP = 5,68$ ) adjetivos, e os FM a produzirem em média 44,57 ( $DP = 14,24$ ) nomes, 19,76 ( $DP = 6,49$ ) verbos e 13,19 ( $DP = 6,29$ ) adjetivos), o tratamento estatístico dos dados revelou que as mesmas apenas se verificam na categoria dos nomes ( $t(42) = -3.81, p < .001$ ) e dos verbos ( $t(42) = -2.67, p = .011$ ). Não são, portanto, estatisticamente relevantes no que se refere à categoria dos adjetivos ( $t(42) = -1.33, p = .189$ ). Assim, constatou-se que as crianças bilingues produzem significativamente menos nomes e menos verbos do que as crianças monolingues, mas utilizam quantidades idênticas de adjetivos. Não indo ao encontro do que se previra na secção 4.3, isto é, que as crianças bilingues produziriam quantidades de itens lexicais inferiores aos das crianças monolingues em cada uma das categorias gramaticais, a divergência observada, designadamente no que se refere à categoria gramatical dos adjetivos, poderá estar relacionada com o facto de que, nos estudos sobre a composição do léxico de crianças bilingues (David e Wei, 2005; dos Santos e Kern, 2013), nos quais algumas das predições desta investigação se fundamentam, os itens lexicais pertencentes às categorias gramaticais verbos e adjetivos se encontram agrupados sob a categoria predicados, não sendo possível discriminar, especificamente, a proporção de verbos e de adjetivos nos seus repertórios lexicais. No entanto, se atentarmos novamente aos valores quantitativos, supramencionados, de itens lexicais produzidos por cada grupo de falantes do PE, verifica-se que, comparativamente às outras categorias gramaticais, quer o grupo de FH, quer o grupo de FM, produzem poucos itens lexicais da categoria adjetivos, na tarefa experimental. De facto, os dados apresentados no Gráfico 3 e na Tabela 1 mostram que 75% dos FH (i.e. os três primeiros quartis) utilizam entre dois e 13 adjetivos, sendo que 75% dos FM produzem uma

quantidade de adjetivos igualmente baixa, isto é, entre dois e 18,5. Veja-se que, no que respeita aos nomes, estes valores situam-se entre os 10 e os 36 nomes, para os FH, e entre os 27 e os 55, para os FM. Já na categoria dos verbos, 75% das crianças bilingues produzem entre zero e 20 verbos e as crianças monolingues entre nove e 25,5. Assim, o facto de a maioria dos participantes, bilingues e monolingues, utilizar uma quantidade baixa de adjetivos na tarefa descritivo-narrativa poderá ser uma das razões para não se terem verificado diferenças significativas entre os grupos sob investigação.

Contudo, apesar das diferenças observadas entre o grupo de FH e o grupo de controlo relativamente à produção lexical total, nomeadamente nas categorias gramaticais nomes e verbos, os resultados das análises estatísticas a nível intragrupal mostram que os repertórios lexicais das crianças falantes do PE, sejam estas bilingues ou monolingues, apresentam padrões de distribuição lexical idênticos no que respeita à proporção de itens lexicais pertencentes às categorias gramaticais nomes, verbos e adjetivos, ou seja, ambos os grupos em comparação produzem significativamente mais nomes do que verbos e mais verbos do que adjetivos na tarefa de produção oral semi-espontânea. Desta forma, confirmam-se para as crianças falantes do PE os resultados de estudos com crianças bilingues e monolingues de outras populações linguísticas (Bates *et al.*, 1994; Bornstein *et al.*, 2004; Conboy e Thal, 2006; David e Wei, 2005; Jackson-Maldonado *et al.*, 1993; Özcan *et al.*, 2016; dos Santos e Kern, 2013; entre outros), segundo os quais os repertórios lexicais das crianças investigadas são compostos maioritariamente por itens lexicais pertencentes à categoria gramatical dos nomes. Estes resultados sugerem ainda que, no que se refere à composição lexical, o fenómeno da *tendência nominal* observado em estádios iniciais de aquisição lexical de crianças monolingues e bilingues de várias populações linguísticas se prolonga em fases subsequentes de desenvolvimento linguístico nas crianças falantes do PE.

Por outro lado, em consonância com as conclusões de autores como Gathercole (2014), Montrul (2016) e Unsworth (2016), retiradas da análise de inúmeros estudos sobre crianças bilingues de diversas línguas-alvo, a análise descritiva dos dados recolhidos na tarefa experimental demonstra que o conhecimento lexical das crianças bilingues português-alemão é caracterizado pela variação a nível individual, ou seja, há crianças falantes de herança do PE que se aproximam dos padrões de desenvolvimento lexical das crianças monolingues da língua-alvo, existindo também, por outro lado, crianças que destes se afastam. Esta conclusão pode ser confirmada pela reanálise dos Gráficos 2 e 3, representativos da distribuição e variação lexical por grupo de falantes relativamente à

produção lexical total e por categoria gramatical, nos quais se pode verificar que há valores dos dois grupos de falantes do PE que se sobrepõem, havendo também outros que não são coincidentes. Vejamos, então, por exemplo, a Caixa de Bigodes referente à produção lexical total (i.e. *corpus* total), ou seja, o Gráfico 2. Nesse gráfico, é possível observar que os repertórios lexicais totais dos FH variam entre o mínimo de 12 itens lexicais e o máximo de 115. Por sua vez, os repertórios lexicais totais dos FM situam-se entre o mínimo de 39 itens lexicais e o máximo de 122. Ora, considerando-se esta amplitude lexical como representativa dos padrões de desenvolvimento lexical monolíngue e tendo em conta que o valor da mediana dos dados do grupo bilingue se situa nos 53 itens lexicais, então podemos concluir que, no mínimo, 50% das crianças bilingues possuem repertórios lexicais dentro dos padrões de desenvolvimento linguístico dos seus pares monolíngues do PE. Assim, tendo-se confirmado esta predição, de que o conhecimento lexical dos FH seria marcado pela variação a nível individual, e estando comprovado empiricamente que a variação existente na dimensão dos repertórios lexicais de crianças bilingues está significativamente associada a fatores relacionados com a quantidade e a qualidade de exposição linguística a que estas crianças têm acesso (Bridges e Hoff, 2014; Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; David e Wei, 2008; Gathercole e Thomas, 2009; Hoff *et al.*, 2014; Patterson, 2002; Pearson *et al.*, 1997; Place e Hoff, 2011; Ribot *et al.*, 2017; entre outros), tornou-se fundamental determinar o papel que estes desempenham no desenvolvimento da competência lexical dos falantes de herança do PE.

Neste âmbito, e no que à segunda questão de investigação (Q 2) e respetivas subquestões (Q 2.1 e Q 2.2) diz respeito, os resultados das análises estatísticas, derivados de testes de associação entre o desempenho lexical dos FH e cada um dos sete fatores de *input* selecionados para este estudo, revelam que somente os fatores *input\_pais*, *output\_pais* e *pais\_PE* estão significativa e positivamente correlacionados com o desempenho lexical dos FH na tarefa de produção oral semi-espontânea. Os restantes fatores (i.e. *par\_conv\_PE*, *ativ\_PE*, *irmãos* e *idade*) não apresentam correlações significativas com a competência lexical das crianças bilingues português-alemão, comprovando, desta forma, que apenas uma parte dos fatores de *input* considerados preditivos do desenvolvimento lexical bilingue tem influência sobre o desempenho lexical dos FH do PE.

Assim, à semelhança de outros estudos que investigaram o efeito dos fatores *quantidade de input* e *quantidade de output* sobre a competência lexical produtiva de

crianças bilíngues de outras línguas-alvo (Cohen, 2016; Cote e Bornstein, 2014; Hoff *et al.*, 2014; Place e Hoff, 2011; Ribot *et al.*, 2017; Thordardottir, 2011; entre outros), os resultados estatísticos do presente estudo revelam que o conhecimento lexical dos falantes de herança do PE, avaliado na tarefa experimental, é significativamente influenciado pela quantidade de exposição e de produção linguística que advêm das interações comunicacionais com os progenitores em contexto doméstico. O facto de se ter confirmado que os FH que recebem mais *input* e que mais interagem verbalmente na LH têm, em todas as modalidades de análise (produção lexical total e por categoria gramatical), um desempenho lexical superior ao dos que têm uma exposição linguística mais reduzida e que recorrem a ambas as línguas, i.e. ao português e ao alemão, na comunicação com os progenitores, remete-nos para a importância de que, para se desenvolverem competências comunicativas na LH e para se assegurar a transmissão intergeracional da mesma, é estritamente necessário que a língua minoritária assuma um papel de relevo no quotidiano familiar das crianças bilíngues. É importante salientar ainda que, em ambientes sociolinguísticos bilíngues, em que uma das línguas da criança é falada pela sociedade dominante em contextos comunicacionais diversificados (i.e. a língua maioritária) e a outra é, na grande maioria das vezes, utilizada em contextos muito restritos, geralmente circunscritos ao ambiente familiar e nos quais os pais são as principais fontes de exposição linguística, a língua minoritária, neste caso o PE, é mais vulnerável a efeitos de *input* reduzido do que a maioritária, isto é, o alemão.

Para além de a quantidade de *input* e de *output* em contexto doméstico estarem significativa e positivamente correlacionadas com a dimensão dos repertórios lexicais dos FH do PE, os resultados mostram ainda que o número de progenitores falantes do PE está positivamente associado ao repertório lexical total de cada FH, concretamente às categorias gramaticais verbos e adjetivos, mas não à dos nomes, ou seja, os FH cujos progenitores são ambos falantes do PE produzem mais verbos e mais adjetivos do que os FH que têm apenas um progenitor português. Ora, tendo-se verificado uma correlação estatisticamente positiva entre a quantidade de progenitores falantes do PE e a quantidade de *input* e de *output* em contexto doméstico, poder-se-á concluir que os FH necessitam de mais exposição e interação linguística na LH para a aquisição de verbos e de adjetivos do que para a aquisição de nomes. O facto de as crianças investigadas, aparentemente, precisarem de menos exposição e de menos interação linguística na LH para a aquisição de itens lexicais pertencentes à categoria gramatical nomes do que para aquisição de itens das categorias verbos e adjetivos poderá estar relacionado com o pressuposto teórico

defendido por Gentner (1982) de que é mais fácil para uma criança identificar conceitos codificados por nomes do que por verbos e outros predicados, visto que estes possuem estruturas semanticamente mais complexas e têm uma relação menos transparente relativamente ao mundo perceptual. Assim, a conjugação dos resultados provenientes dos testes de associação entre o desempenho lexical dos FH e as variáveis *input\_pais* e *output\_pais*, por um lado, e entre estas duas variáveis e a variável *pais\_PE*, por outro lado, parece sugerir que, para os FH estabelecerem as relações semântico-perceptuais complexas que a aquisição de itens lexicais pertencentes aos predicados exige, estes precisam de receber mais exposição linguística do PE e de interagir com mais frequência nesta língua.

É de salientar ainda que a existência de correlações positivas entre a competência lexical dos FH e a quantidade de progenitores falantes do PE, bem como entre esta última variável e a quantidade de exposição linguística ao PE em contexto doméstico, coloca o presente estudo em linha com a investigação desenvolvida por Place e Hoff (2011), na qual as autoras constataam que a competência lexical de crianças falantes de herança de espanhol é significativamente influenciada pela quantidade de exposição linguística à língua minoritária, sendo as crianças que têm dois progenitores falantes de espanhol mais expostas a esta língua do que as que têm apenas um progenitor falante da mesma.

Contudo, embora se tenha previsto, com base em alguns dos estudos apresentados na secção 3.2 (Ishizawa, 2004; Jia e Aaronson, 2003; Jia e Fuse, 2007; Patterson, 2002; Place e Hoff, 2011; Scheele *et al.*, 2010), que o desempenho linguístico dos FH seria significativamente influenciado pela qualidade da exposição linguística não só através do contacto regular com um maior número de interlocutores falantes nativos do PE em contexto de emigração (i.e. tios e avós), como também através da quantidade de atividades de lazer em que os FH participam regularmente, utilizando a LH, os resultados estatísticos não apresentam correlações significativas entre o conhecimento lexical dos FH e as variáveis *par\_conv\_PE* e *ativ\_PE*. Assim sendo, tendo em conta os resultados das variáveis analisadas até este ponto da discussão, os dados estatísticos parecem sugerir que, com o grupo de falantes bilingues sob investigação, a quantidade de *input* e a quantidade de *output* linguístico em contexto doméstico (i.e. interação filhos-pais) desempenham um papel mais significativo no desenvolvimento da competência lexical produtiva na LH do que o contacto com outros falantes nativos do PE e com outras fontes de exposição linguística mais próximas da norma padrão do PE.

Por sua vez, e também contrariamente à predição elaborada na secção 4.3, a análise estatística entre o desempenho lexical dos FH e o fator *irmãos* revela que a presença de irmãos mais velhos no agregado familiar não tem influência sobre o desenvolvimento da competência lexical produtiva dos FH do PE. Uma vez que, no presente estudo, não se analisou a quantidade de *input* do PE proveniente de irmãos mais velhos, nem a quantidade de *output* em PE dirigida a estes, considera-se que seria importante, em futuras investigações, incluir estas duas variáveis no modelo de análise estatística, de modo a se ter uma perspectiva mais esclarecedora sobre o efeito que a existência de irmãos mais velhos no seio familiar exerce sobre o ambiente linguístico de crianças bilingues e, conseqüentemente, sobre a sua competência linguística na LH.

Por último, no que respeita ao fator *idade*, apesar de se ter partido do pressuposto de que a idade dos FH, expostos ao PE desde o nascimento, refletiria a quantidade de exposição linguística a esta língua ao longo das suas vidas, isto é, de que as crianças bilingues mais velhas teriam sido expostas a uma maior quantidade de *input* da LH do que as crianças mais novas, apresentando os falantes mais velhos, conseqüentemente, melhores resultados do que os mais novos na tarefa de produção, a análise inferencial dos dados revela que a idade cronológica não é um fator preditivo do desenvolvimento lexical deste grupo de falantes de herança do PE. O facto de a idade dos FH não apresentar correlações significativas com o desempenho lexical dos mesmos corrobora o argumento defendido por Unsworth (2013) de que na avaliação do *tempo de exposição acumulada* se deve ter em conta não só que a quantidade de *input* que as crianças bilingues recebem de cada língua-alvo é inferior à que as correspondentes crianças monolingues recebem, como também que o grau de exposição a cada língua-alvo pode variar muito ao longo da vida de cada falante bilingue. Desta forma, a ausência de correlações significativas entre a idade cronológica dos FH e o seu desempenho lexical na tarefa experimental pode dever-se ao facto de a idade das crianças bilingues, na faixa etária testada (i.e. dos seis aos 11 anos), não ser uma medida representativa da quantidade de exposição linguística que estas efetivamente receberam ao longo das suas vidas. De facto, se se tiverem em conta as variáveis propostas por Unsworth (2011) para o cálculo do *TEA*, ou seja, (1) a quantidade de *input* de cada língua-alvo proveniente de cada progenitor e adulto a residir na mesma casa, por cada período de um ou de dois anos, antes do momento da recolha de dados, (2) a quantidade de *input* de cada língua-alvo a que as crianças são expostas em creches ou instituições (pré-) escolares durante os períodos anuais ou bianuais antecedentes à recolha de dados, bem como (3) a quantidade de exposição às línguas-alvo

durante o período de férias, então é muito provável que, no grupo sob investigação, duas crianças bilingues que partilhem a mesma idade cronológica tenham *TEA* diferentes à LH. Por exemplo, na subsecção 4.2.1, dedicada à caracterização sociolinguística dos FH, verificou-se que, das 23 crianças bilingues expostas ao PE desde a nascença, 16 foram expostas à língua alemã antes dos três anos de idade e que sete crianças somente começaram a receber exposição linguística mais intensiva a esta língua entre os três e os cinco anos. Ora, focando apenas esta variação na quantidade de exposição linguística que cada criança bilingue teve a ambas as línguas durante os primeiros anos de vida, pode concluir-se, *a priori*, que uma criança de oito anos, cuja idade de início de exposição ao alemão foi aos cinco anos de idade, tem um *TEA* à LH superior ao de uma criança de oito anos que, desde o nascimento, teve de dividir o seu tempo de exposição linguística entre o PE e o alemão. Assim, sugere-se que, futuramente, em alternativa à idade cronológica, se adotem os critérios apresentados por Unsworth (2011, 2013) para o cálculo da exposição linguística que as crianças bilingues recebem ao longo das suas vidas.

Para concluir este capítulo, é ainda relevante salientar que, apesar de se crer que os resultados desta investigação contribuíram para uma melhor compreensão do processo de aquisição lexical das crianças falantes de herança do PE, o campo de estudos da aquisição bilingue teria beneficiado se, para além da competência lexical produtiva, se tivesse analisado também a competência lexical recetiva deste grupo de falantes bilingues. Essa análise comparativa entre o domínio da produção e o da compreensão oral permitiria verificar se, à semelhança de vários estudos com crianças bilingues de outros pares linguísticos (De Houwer *et al.*, 2014; Pearson *et al.*, 1993; Sheng *et al.*, 2011; entre outros), a competência lexical recetiva das crianças bilingues falantes do PE é superior à produtiva e também se, como alguns estudos indicam (Sheng *et al.*, 2011; Thordardottir, 2011), esta última é mais vulnerável a efeitos de *input* reduzido do que a primeira (i.e. a recetiva). Para além disso, considera-se que a vasta literatura sobre o bilinguismo sairia enriquecida se se analisasse, concomitantemente, não só a produção, mas também a compreensão oral, em ambas as línguas dos participantes investigados, ou seja, no PE e no alemão. Uma análise lexical contrastiva entre a língua minoritária e a maioritária das crianças bilingues português-alemão permitiria avaliar se estas, tal como outros falantes bilingues de outros contextos sociolinguísticos (Hoff *et al.*, 2012; Junker e Stockman, 2002; Poulin-Dubois *et al.*, 2013; entre outros), tendem a apresentar repertórios lexicais produtivos e recetivos quantitativamente inferiores aos das correspondentes crianças monolingues, quando apenas uma das suas línguas é avaliada (geralmente, a língua não

dominante), bem como se, quando ambas as suas línguas são tidas em consideração, as crianças bilíngues português-alemão detêm uma dimensão de vocabulário total e conceptual, ao nível da produção e da compreensão, semelhante ao das crianças monolíngues das respetivas comunidades linguísticas.



## Conclusão

Ante a dimensão demográfica da comunidade de emigrantes portugueses e de lusodescendentes a residirem na Alemanha e dada a escassez de estudos sobre a aquisição e o desenvolvimento lexical de falantes de herança do PE, o presente estudo analisou o desempenho lexical de 23 crianças bilingues português-alemão, falantes de herança do PE, numa tarefa de produção oral semi-espontânea, comparando-as entre si e com um grupo de controlo composto por 21 crianças monolingues do PE. À semelhança de inúmeras investigações empíricas sobre a aquisição lexical de crianças falantes de herança de diversas comunidades linguísticas, segundo as quais os falantes bilingues tendem a ter repertórios lexicais quantitativamente inferiores aos das crianças monolingues das respetivas línguas-alvo, quando apenas uma das suas línguas é avaliada (Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; Hoff *et al.*, 2012; Pearson *et al.*, 1993; Poulin-Dubois *et al.*, 2013; entre outros), este estudo procurou investigar se a competência lexical das crianças bilingues lusodescendentes apresenta padrões de desenvolvimento lexical idênticos aos dos seus pares monolingues do PE ou se, pelo contrário, destes se afastam. Cumulativamente, e também no âmbito da investigação internacional realizada na área da aquisição linguística bilingue, de acordo com a qual o desenvolvimento lexical de crianças expostas simultaneamente, desde muito cedo, a mais do que uma língua é fortemente influenciado por fatores relacionados com a quantidade e a qualidade da exposição linguística a que estas têm acesso (Bridges e Hoff, 2014; Cobo-Lewis *et al.*, 2002a, 2002b; Cote e Bornstein, 2014; David e Wei, 2008; Gathercole e Thomas, 2009; Hoff *et al.*, 2014; Patterson, 2002; Pearson *et al.*, 1997; Place e Hoff, 2011; Ribot *et al.*, 2017; entre outros), procurou-se avaliar a influência que os mesmos têm no processo de aquisição lexical das crianças bilingues português-alemão, designadamente no que à LH diz respeito.

Os resultados da análise comparativa, a nível intra e intergrupar, das produções orais de ambos os grupos de falantes do PE, extraídas da tarefa experimental, por um lado, bem como os resultados da avaliação do desempenho lexical das crianças bilingues conjugados com as informações, fornecidas no questionário sociolinguístico, relativas à quantidade e à qualidade da exposição linguística à LH a que as mesmas têm acesso no seu quotidiano, por outro, permitem-nos tecer algumas conclusões, que se apresentarão de seguida, sobre o desenvolvimento da competência lexical deste tipo de falantes do PE, ainda pouco estudado, e contribuir, desta forma, para o enriquecimento da literatura

especializada na área de aquisição de línguas de herança, neste caso, do português europeu como LH.

Assim, da primeira questão de investigação (Q 1) e respetivas subquestões (Q 1.1 e Q 1.2), centradas na avaliação do conhecimento lexical das crianças falantes de herança do PE em comparação com o das crianças monolíngues da língua-alvo, advém o primeiro conjunto de conclusões:

- (1) Como grupo, os falantes de herança do PE apresentam repertórios lexicais quantitativamente inferiores aos das correspondentes crianças monolíngues do PE;
- (2) A nível individual, o conhecimento lexical das crianças falantes de herança do PE é marcado pela variação. Por um lado, há crianças bilingues cuja dimensão dos seus repertórios lexicais na LH é semelhante à das crianças monolíngues do PE. Por outro lado, existem também crianças cujo tamanho dos seus repertórios lexicais na LH não se enquadra nos padrões de desenvolvimento lexical apresentados pelos seus pares monolíngues do PE;
- (3) As diferenças observadas entre a dimensão lexical do grupo de FH do PE e a do grupo de FM do PE manifestam-se nas categorias gramaticais nomes e verbos, mas não na dos adjetivos, categoria da qual as crianças bilingues e monolíngues produzem uma quantidade de itens lexicais igualmente baixa;
- (4) Porém, os repertórios lexicais de ambos os grupos de falantes do PE apresentam padrões de distribuição lexical idênticos no que se refere à proporção de itens lexicais pertencentes às categorias gramaticais nomes, verbos e adjetivos – quer as crianças falantes de herança do PE, quer os seus pares monolíngues, produzem significativamente mais nomes do que verbos e mais verbos do que adjetivos na tarefa descritivo-narrativa da mesma imagem-cenário;
- (5) Por fim, a presença de uma maior quantidade de itens lexicais pertencentes à categoria gramatical nomes nos repertórios lexicais dos dois grupos de participantes sugere que o fenómeno da *tendência nominal*, observado em estádios iniciais de desenvolvimento lexical de crianças bilingues e monolíngues de diversas populações linguísticas, se mantém em fases subsequentes de desenvolvimento linguístico nas crianças falantes do PE.

Por sua vez, da segunda questão de investigação (Q 2) e correspondentes subquestões (Q 2.1 e Q 2.2), centradas no papel que os diferentes fatores de *input* selecionados para este estudo (i.e. *input\_pais*, *output\_pais*, *pais\_PE par\_conv\_PE*, *ativ\_PE*, *irmãos* e *idade*) desempenham no desenvolvimento lexical das crianças falantes de herança do PE, surge o segundo grupo de conclusões:

- (1) As crianças bilingues português-alemão que recebem mais *input* linguístico do PE e que comunicam com mais frequência nesta língua com os progenitores, em contexto doméstico, apresentam um conhecimento lexical na LH superior, em todas as modalidades de análise (i.e. produção lexical total e por categoria gramatical), ao das crianças que são menos expostas ao PE através das figuras parentais, em contexto doméstico, e que utilizam menos o PE (e mais o alemão) nas interações verbais com os progenitores no mesmo contexto comunicacional;
- (2) As crianças cujos progenitores são ambos falantes do PE, para além de estarem mais expostas e de comunicarem mais na LH no seio familiar, têm repertórios lexicais totais quantitativamente superiores, concretamente nas categorias gramaticais verbos e adjetivos, ao das crianças que têm apenas um progenitor português e que, conseqüentemente, recebem menos *input* da LH e produzem menos *output* na mesma;
- (3) A aquisição de itens lexicais pertencentes às categorias gramaticais verbos e adjetivos parece ser mais vulnerável a efeitos de *input* reduzido do que a de itens lexicais da categoria nominal. Aparentemente, para estabelecerem as relações semântico-perceptuais complexas que a aquisição de itens lexicais pertencentes aos predicados exige (Gentner, 1982), as crianças falantes de herança do PE necessitam de ser mais expostas e de interagir mais na LH, em contexto doméstico, com os progenitores;
- (4) Não se tendo verificado correlações significativas entre a competência lexical na LH das crianças bilingues e outros fatores de *input*, mais qualitativos, do PE, designadamente o contacto regular com um maior número de interlocutores falantes nativos do PE em contexto de emigração (i.e. tios e avós) e a pluralidade de atividades de lazer em que os FH participam regularmente utilizando a língua portuguesa (i.e. ver televisão, ler livros, etc.), os progenitores das crianças bilingues assumem-se, assim, como os principais intervenientes, quer através da quantidade de *input* que fornecem, quer através da quantidade de *output* que

provocam enquanto parceiros conversacionais, no processo de aquisição do PE como língua de herança;

- (5) Apesar de alguns estudos, como o de Bridges e Hoff (2014), terem constatado que a existência de irmãos mais velhos no agregado familiar de crianças bilingues tem um efeito negativo no desenvolvimento de competências comunicativas na LH, a presença destes no seio familiar das crianças falantes de herança do PE não revelou ter influência sobre o desenvolvimento da competência lexical produtiva das mesmas na sua LH;
- (6) Por último, a idade cronológica dos falantes de herança, tida, inicialmente, como uma medida representativa da quantidade de *input* do PE a que as crianças bilingues teriam sido expostas ao longo das suas vidas, revelou também não ser um fator de *input* preditivo da competência lexical produtiva das mesmas na LH. O facto de nem todas as crianças bilingues mais velhas apresentarem repertórios lexicais quantitativamente superiores aos das crianças bilingues mais novas sugere que, para uma avaliação exata da quantidade de exposição linguística à LH ao longo das suas vidas, se deve ter em consideração toda e qualquer variação que possa ter ocorrido no grau de exposição linguística a que estas tiveram acesso nos anos antecedentes ao momento da recolha dos dados.

Assim, tendo em consideração as conclusões apresentadas, o que ressalta do presente estudo é o facto de se ter observado que, apesar de estarem inseridas num contexto sociolinguístico em que são expostas à língua de herança (i.e. PE) em ambientes comunicacionais muito restritos e à língua maioritária (i.e. alemão) em contextos interacionais multifacetados, mais de metade das crianças bilingues português-alemão investigadas apresentam, na LH, padrões de desenvolvimento lexical idênticos aos dos seus pares monolingues do PE. Contudo, o facto de pertencerem a um agregado familiar em que pelo menos um dos progenitores é falante do PE não dá como garantida a aquisição da língua portuguesa dentro dos padrões de desenvolvimento linguístico dos falantes monolingues do PE no país de origem. Tal deve-se, como constatado, à variação no grau de exposição linguística ao PE e de produção linguística nesta língua procedente das interações comunicacionais entre as crianças bilingues e os seus progenitores em contexto doméstico, o que nos leva a reiterar a ideia de que, enquanto fontes primárias de *input* da língua minoritária e principais parceiros conversacionais das crianças bilingues

lusodescendentes, os pais são, efetivamente, os principais intervenientes no processo de aquisição linguística do PE como LH.

Embora não se tenha avaliado a competência lexical dos FH no alemão, a língua da sociedade dominante, sabe-se que as crianças que crescem expostas a duas línguas invariavelmente adquirem a língua maioritária, sendo a língua minoritária, mais vulnerável a efeitos de *input* reduzido, aquela que está em risco de não ser adquirida (De Houwer, 2007).

Assim sendo, os pais dos falantes de herança não devem ter o receio, infundado, de que os seus filhos, ao serem expostos à LH em contexto doméstico, não venham a desenvolver proficiência linguística na língua maioritária. Contudo, se desejarem que os seus descendentes “herdem” a língua dos pais e avós, então é crucial que, no processo de aculturação à sociedade de acolhimento, estes não abdicuem de a usar nas interações quotidianas com os seus filhos.



## Referências

- Alves, I., Costa, P., Lourenço-Gomes, M. C., e Rodrigues, C. (2015). EFFE-On–Corpus Online de Escrita e Fala. *Saber & Educar*, 20, 24-33.
- Au, T. K., Dapretto, M., e Song, Y. K. (1994). Input vs. constraints: Early word acquisition in Korean and English. *Journal of Memory and Language*, 33, 567 – 582.
- Au, T., Knightly, L., Jun, S., e Oh, J. (2002). Overhearing a language during childhood. *Psychological Science*, 13(3), 238 – 243.
- Baganha, M. I. B. (1994). As correntes emigratórias portuguesas no século XX e o seu impacto na economia nacional. *Análise social*, 128(29), 959-980.
- Baker, C. (2006). *Foundations of bilingual education and bilingualism*. Clevedon: Multilingual Matters.
- Baker, C., e Jones, S. (1998). *Encyclopedia of Bilingualism and Bilingual Education*. Philadelphia: Clevedon Multilingual Matters.
- Barnes, J., e García, I. (2013). Vocabulary growth and composition in monolingual and bilingual Basque infants and toddlers. *International Journal of Bilingualism*, 17, 357-374.
- Barrett, M. (1982). The holophrastic hypothesis: Conceptual and empirical issues. *Cognition*, 11, 47-76.
- Barrett, M. (1997). Desenvolvimento lexical inicial. In P. Fletcher e B. MacWhinney (Eds.), *Compêndio da linguagem da criança* (pp. 299-322). Porto Alegre: Artes Médicas.
- Barron-Hauwaert, S. (2011). *Bilingual siblings*. Clevedon: Multilingual Matters.
- Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, 27, 521 – 559.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J., *et al.* (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21(1), 85-124.
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., *et al.* (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism: Language and Cognition*, 15, 616-629.

- Benmamoun, E., Montrul, S., e Polinsky, M. (2013). Heritage languages and their speakers: opportunities and challenges for linguistics. *Theoretical Linguistics*, 39(3-4), 129-181.
- Berman, R. A. (2009). Developing linguistic knowledge and language use across adolescence. In E. Hoff e M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 347-367). Malden, MA: Wiley-Blackwell.
- Best, C. T. (1995). Learning to perceive the sound pattern of English. In C. Rovee-Collier e L. Lipsitt (Eds.), *Advances in Infancy Research* (Vol. 9, pp. 217–304). Norwood, NJ: Ablex.
- Bialystok, E., Craik, F. I., Green, D. W., e Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10(3), 89-129.
- Bialystok, E., Luk, G., Peets, K. F., e Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13, 525-531.
- Blake, R. (1983). Mood Selection among Spanish Speaking Children, Ages 4 to 12. *The Bilingual Review*, 10, 21-32.
- Blom, E., Polišenská, D., e Weerman, F. (2008). Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research*, 24, 297-332.
- Bohman, T. M., Bedore, L. M., Pena, E. D., Mendez-Perez, A., e Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13, 325-344.
- Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., *et al.* (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75, 1115-1139.
- Bosch, L., e Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46, 217–243.
- Bridges, K., e Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics*, 35, 225-241.
- Bruner, J. (2007). *Como as crianças aprendem a falar*. Lisboa: Instituto Piaget.

- Burns, T. C., Yoshida, K. A., Hill, K., e Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28(3), 455-474.
- Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., e Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10, 159-199.
- Castro, S. L., e Gomes, I. (2000). *Dificuldades de Aprendizagem da Língua Materna*. Lisboa: Universidade Aberta.
- Cha, K., e Goldenberg, C. (2015). The complex relationship between bilingual home language input and kindergarten children's Spanish and English oral proficiencies. *Journal of Educational Psychology*, 107(4), 935–953.
- Chang-Song, Y. K., e Pae, S. (2009). Do Korean children acquire verbs earlier than nouns?. In C. Lee, G. B. Simpson, e Y. Kim (Eds.), *The Handbook of East Asian Psycholinguistics* (Vol. 3, pp. 63-71). Cambridge: Cambridge University Press.
- Choi, S. (2000). Caregiver input in English and Korean: Use of nouns and verbs in book reading and toy-play contexts. *Journal of Child Language*, 27, 69 – 96.
- Choi, S., e Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, 22, 497 – 529.
- Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language*, 35(1), 26-58.
- Chomsky, N. (1968). *Language and Mind*. New York: Harcourt Brace Jovanovich.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Chomsky, N. (1988). *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.
- Cobo-Lewis, A., Pearson, B., Eilers, R., e Umbel, V. (2002a). Effects of bilingualism and bilingual education on oral and written English skills: A multifactor study of standardized test outcomes. In D. K. Oller e R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 64–97). Clevedon, UK: Multilingual Matters.
- Cobo-Lewis, A., Pearson, B., Eilers, R., e Umbel, V. (2002b). Effects of bilingualism and bilingual education on oral and written Spanish skills: A Multifactor Study of Standardized Test Outcomes. In D. K. Oller e R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 98-117). Clevedon, UK: Multilingual Matters.

- Cohen, C. (2016). Relating input factors and dual language proficiency in French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, 19(3), 296-313.
- Conboy, B. T., e Montanari, S. (2016). Early Lexical Development in Bilingual Infants and Toddlers. In E. Nicoladis e S. Montanari (Eds.), *Bilingualism Across the Lifespan* (pp. 63-80). Berlin, Boston: De Gruyter.
- Conboy, B. T., e Thal, D. J. (2006). Ties Between the Lexicon and Grammar: Cross-Sectional and Longitudinal Studies of Bilingual Toddlers. *Child development*, 77(3), 712-735.
- Cornips, L., e Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Language Research*, 24, 267-296.
- Cote, L. R., e Bornstein, M. H. (2014). Productive vocabulary among three groups of bilingual American children: Comparison and prediction. *First Language*, 34, 467-485.
- Crystal, D. (1997). *English as a global language*. Cambridge, England: Cambridge University Press.
- Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *The Modern Language Journal*, 89(4), 585–592.
- D’Odorico, L., e Fasolo, M. (2007). Nouns and verbs in the vocabulary acquisition of Italian children. *Journal of Child Language*, 34(4), 891–907.
- David, A. e Wei, L. (2005). The composition of the bilingual lexicon. In J. Cohen, K.T. McAlister, K. Rolstad, e J. MacSwan (Eds.), *Proceedings of the 4th international symposium on bilingualism* (pp. 594-607). Somerville, MA: Cascadilla Press.
- David, A., e Wei, L. (2008). Individual differences in the lexical development of French–English bilingual children. *International Journal of Bilingual Education and Bilingualism*, 11(5), 598-618.
- De Houwer, A. (1990). *The acquisition of two languages from birth: A case study*. Cambridge: Cambridge University Press.
- De Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher e B. MacWhinney (Eds.), *Handbook of child language* (pp. 219-250). London: Blackwell.
- De Houwer, A. (2005). Early bilingual acquisition: Focus on morphosyntax and the separate development hypothesis. In J. Kroll e A. De Groot (Eds.), *The handbook of bilingualism* (pp. 30-48). Oxford, UK: Oxford University Press.

- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied psycholinguistics*, 28, 411-424.
- De Houwer, A. (2009). *Bilingual First Language Acquisition*. Clevedon, UK: Multilingual Matters.
- De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, 2, 221-240.
- De Houwer, A., Bornstein, M. H., e Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, 35(6), 1189-1211.
- De Houwer, A., e Gillis, S. (1998). Dutch child language: An overview. In S. Gillis e A. De Houwer (Eds.), *The acquisition of Dutch* (pp. 1–100). Amsterdam: John Benjamins.
- Dockrell, J. E., e Messer, D. A. (2004). Lexical acquisition in the early school years. In R. A. Berman (Ed.), *Language Development across Childhood and Adolescence* (pp. 35–52). Amsterdam: John Benjamins Publishing.
- Dorian, N. (1981). *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press.
- Doyle, A. B., Champagne, M., e Segalowitz, N. (1978). Some issues in the assessment of linguistic consequences of early bilingualism. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 13-20). Columbia, SC: Hornbeam Press.
- Driessen, G., Van der Slik, F., e De Bot, K. (2002). Home language and language proficiency: A large-scale longitudinal study in Dutch primary schools. *Journal of Multilingual and Multicultural Development*, 23, 175-194.
- Dromi, E. (1987). *Early lexical development*. Cambridge: Cambridge University Press.
- Duarte, I., Colaço, M., Freitas, M. J., e Gonçalves, A. (2011). *O Conhecimento da língua: desenvolver o conhecimento lexical*. Lisboa: ME/DGIDC.
- Ellis, N. C. (1998). Emergentism, Connectionism and Language Learning. *Language Learning*, 48, 631-664
- Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

- Elman, J., Bates, E. A., Johnson, M., Karmiloff-Smith, A., Parisi, D., e Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: The MIT Press.
- Fabiano-Smith, L., e Goldstein, B. A. (2010). Phonological acquisition in bilingual Spanish–English speaking children. *Journal of Speech, Language, and Hearing Research*, 53(1), 160-178.
- Field, A. P. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. Los Angeles: SAGE Publications.
- Fishman, J. (2001). 300-plus years of heritage language education in the United States. In J. K. Peyton, D. A. Ranard, e S. McGinnis (Eds.), *Heritage languages in America: Preserving a national resource* (pp. 81-89). Washington, DC e McHenry, IL: Center for Applied Linguistics & Delta Systems.
- Fishman, J. (2006). Acquisition, maintenance and recovery of heritage languages. In G. Valdés, J. Fishman, R. Chávez, e W. Pérez (Eds.), *Developing minority language resources: The case of Spanish in California* (pp. 12–22). Clevedon, UK: Multilingual Matters.
- Flores, C. (2013). Português Língua Não Materna. Discutindo conceitos de uma perspectiva linguística. In R. Bizarro, M. A. Moreira, e C. Flores (Orgs.), *Português Língua Não Materna: Investigação e Ensino* (pp. 35-46). Lisboa: LIDEL.
- Flores, C. (2015). Understanding heritage language acquisition. Some contributions from the research on heritage speakers of European Portuguese. *Lingua*, 164, 251-265.
- Flores, C., e Barbosa, P. (2014). When reduced input leads to delayed acquisition: a study on the acquisition of clitic placement by Portuguese heritage speakers. *International Journal of Bilingualism*, 18(3), 304-325.
- Flores, C., e Correia, L. (2016). O papel do grau de exposição linguística no desenvolvimento bilingue. In R. Sánches, X. e M.A. Marques (Eds.), *As ciências da linguagem no espaço galego-português. Diversidade e convergência* (pp. 243-261). Braga: ILCH/Húmus.
- Flores, C., e Melo-Pfeifer, S. (2014). O conceito “Língua de Herança” na perspectiva da Linguística e da Didática de Línguas: considerações pluridisciplinares em torno do perfil linguístico das crianças lusodescendentes na Alemanha. *Domínios de Lingu@gem*, 8(3), 16-45.

- Flores, C., Santos, A. L., Jesus, A., e Marques, R. (2016). Age and input effects in the acquisition of mood in Heritage Portuguese, *Journal of Child Language*, 1–34.
- Gathercole, V. C. M. (2002a). Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In D. K. Oller e R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 207– 219). Clevedon, UK: Multilingual Matters.
- Gathercole, V. C. M. (2002b). Command of the mass/count distinction in bilingual and monolingual children: An English morphosyntactic distinction. In D. K. Oller e R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 175–206). Clevedon, UK: Multilingual Matters.
- Gathercole, V. C. M. (2002c). Monolingual and bilingual acquisition: Learning different treatments of *that*-trace phenomena in English and Spanish. In D. K. Oller e R. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 220–254). Clevedon, UK: Multilingual Matters.
- Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: Constructivist account of morpho-syntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism*, 10(3), 224–247.
- Gathercole, V. C. M. (2014). Bilingualism matters: One size does not fit all. *International Journal of Behavioral Development*, 38, 359-366.
- Gathercole, V. C. M., e Thomas, E. M. (2009). Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Bilingualism: Language and Cognition*, 12, 213–237.
- Gathercole, V. C. M., Thomas, E. M., e Laporte, N. I. (2001). The acquisition of grammatical gender in Welsh. *Journal of Celtic Language Learning*, 6, 53–87.
- Gelman, S. A., e Tardif, T. (1998). Acquisition of nouns and verbs in Mandarin and English. In E. Clark (Ed.), *Proceedings of the 29<sup>th</sup> Annual Stanford Child Language Research Forum*. Stanford: CSLI.
- Genesee, F. (1989). Early bilingual development: one language or two?. *Journal of Child Language*, 16(1), 161–179.
- Genesee, F. (2001). Bilingual first language acquisition: exploring the limits of the language faculty. *Annual Review of Applied Linguistics*, 21, 153–168.
- Genesee, F. (2010). Dual language development in preschool children. In E. E. Garcia e E. C. Frede (Eds.), *Young English Language Learners: current research and*

- emerging directions for practice and policy* (pp. 59-79). New York, NY: Teachers College Press.
- Genesee, F., e Nicoladis, E. (2006). Bilingual acquisition. In E. Hoff e M. Shatz (Eds.), *Handbook of Language Development* (pp. 324-342). Oxford, England: Blackwell.
- Genesee, F., e Nicoladis, E. (2009). Bilingual first language acquisition. In E. Hoff e M. Shatz (Eds.), *Blackwell Handbook of Language Development* (pp. 324-342). Malden, MA: Wiley-Blackwell.
- Genesee, F., Nicoladis, E., e Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, 22, 611-631.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: language, thought and culture* (Vol. 2, pp. 301-334). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gentner, D., e Boroditsky, L. (2009). Early acquisition of nouns and verbs: Evidence from Navajo. In V. C. M. Gathercole (Ed.), *Routes to Language: Studies in honor of Melissa Bowerman* (pp. 5-36). New York: Psychology Press.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldfield, B. A. (2000). Nouns before verbs in comprehension vs. production: The view from pragmatics. *Journal of Child Language*, 27, 501 – 520.
- Gopnik, A., Choi, S., e Baumberger, T. (1996). Cross-linguistic differences in early semantic and cognitive development. *Cognitive Development*, 11, 197 – 227.
- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, Massachusetts: Harvard University Press.
- Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, Massachusetts: Harvard University Press.
- Guerreiro, H. (2007). *Processos fonológicos na fala da criança de cinco anos*. Tese de Mestrado, Instituto de Ciências da Saúde da Universidade Católica Portuguesa-Escola Superior de Saúde do Alcoitão da Santa Casa da Misericórdia de Lisboa, Lisboa.
- Gutiérrez-Clellen, V., e Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, 24(2), 267-288.

- Hammer, C. S., Komaroff, E., Rodriguez, B. L., Lopez, L. M., Scarpino, S. E., e Goldstein, B. (2012). Predicting Spanish-English bilingual children's abilities. *Journal of Speech, Language and Hearing Research*, 55, 1251-1264.
- Harris, M., Barrett, M., Jones, D., e Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language*, 15, 77-94.
- Hart, B. (1991). Input frequency and children's first words. *First Language*, 11, 289-300.
- Hart, B., e Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55-88.
- Hoff, E., e Core, C. (2013). Input and language development in bilingually developing children. *Seminars in Speech and Language*, 34, 215-26.
- Hoff, E., Core, C., Place, S., Rumiche, R., Senior, M., e Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39, 1-27.
- Hoff, E., e Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, 73, 418-433.
- Hoff, E., Rumiche, R., Burrige, A., Ribot, K. M., e Welsh, S. N. (2014). Expressive vocabulary development in children from bilingual and monolingual homes: A longitudinal study from two to four years. *Early Childhood Research Quarterly*, 29, 433-444.
- Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, 19, 603-629.
- Hoff-Ginsberg, E., e Krueger, W. (1991). Older siblings as conversational partners. *Merrill Palmer Quarterly*, 37, 465-481.
- Hollowka, S., Brosseau-Lapr e, F., e Petitto, L. A. (2002). Semantic and conceptual knowledge underlying bilingual babies' first signs and words. *Language learning*, 52(2), 205-262.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., e Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27, 236-248.
- Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., e Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, 61, 343-365.

- Hyltenstam, K., e Abrahamsson, N. (2003). Maturation constraints in second language acquisition. In C. Doughty e M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 539–588). Oxford: Blackwell.
- Ingram, D. (1989). *First Language Acquisition: Method, Description and Explanation*. Cambridge, UK: Cambridge University Press.
- Ishizawa, H. (2004). Minority language use among grandchildren in multigenerational households. *Sociological Perspectives*, 47(4), 465-483.
- Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., e Gutiérrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language*, 20, 523-549.
- Jesus, A. (2014). *Aquisição do modo em orações completivas do Português Europeu: o papel dos traços de epistemicidade e veridicidade*. Tese de Mestrado, Universidade de Lisboa, Lisboa.
- Jia, G., e Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, 24, 131-161.
- Jia, G., e Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research*, 50, 1280–1299.
- Johnson, C., e Lancaster, P. (1998). The development of more than one phonology: A case study of a Norwegian-English bilingual child. *International Journal of Bilingualism*, 2(3), 265–300.
- Jones, C., e Adamson, L. (1987). Language Use in Mother-Child and Mother-Child-Sibling Interactions. *Child Development*, 58(2), 356-366.
- Junker, D. A., e Stockman, I. J. (2002). Expressive vocabulary of German-English bilingual toddlers. *American Journal of Speech-Language Pathology*, 11, 381-394.
- Kan, P. F., e Kohnert, K. (2005). Preschoolers learning Hmong and English: Lexical semantic skills in L1 and L2. *Journal of Speech, Language, and Hearing Research*, 48, 372–383.
- Kim, M., McGregor, K. K., e Thompson, C. K. (2000). Early lexical development in English- and Korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language*, 27, 224 – 254.
- Klein, W. (1986). *Second Language Acquisition*. Cambridge, UK: Cambridge University Press.

- Kopeliovich, S. (2013). Happylingual: A family project for enhancing and balancing multilingual development. In M. Schwartz e A. Verschik (Eds.). *Successful family language policy: Parents, children and educators in interaction* (pp. 249-275). Dordrecht: Springer.
- Köppe, R. (1996). Language differentiation in bilingual children: The development of grammatical and pragmatic competence. *Linguistics*, 34, 927–954.
- Krashen, S. (1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Kuczaj, S. A. (1999). The world of words: Thoughts on the development of a lexicon. In M. Barrett (Ed.), *The development of language* (pp. 133-159). Hove: Psychology Press.
- Leiria, I. (2004). Português língua segunda e língua estrangeira: investigação e ensino. *Idiomático. Revista Digital de Didáctica de PLN*, (3).
- Leiria, I., Queiroga, M. J., e Soares, N. V. (2006). *Português Língua Não Materna no Currículo Nacional. Orientações Nacionais. Perfis Linguísticos da População Escolar que Frequenta as Escolas Portuguesas*. Lisboa: Ministério da Educação, DGIDC.
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: John Wiley.
- Leopold, W. (1939-49). *Speech development of a bilingual child: a linguist's record* (Vols. 1-4). Evanston, IL: Northwestern University Press.
- Leseman, P. P. M., Mayo, A. Y., e Scheele, A. F. (2009). Old and new media in the lives of young disadvantaged bilingual children. In A. G. Bus e S. B. Neuman (Eds.), *Multimedia and literacy development* (pp. 135–155). New York, NY: Routledge.
- Li, P., Zhao, X., e MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science*, 31, 581–612.
- Libben, G., e Goral, M. (2015). How bilingualism shapes the mental lexicon. In J. W. Schwieter (Ed.), *The Cambridge Handbook of Bilingual Processing* (pp. 614-630). Cambridge: Cambridge University Press.
- Lima, R. (2000). *Linguagem infantil: da normalidade à patologia*. Braga: APPACDM.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.

- Lourenço-Gomes, M. C., Rodrigues, C., e Alves, I. (2016). EFFE-Escreves como falas–falas como escreves?. *Revue Romane. Langue et littérature. International Journal of Romance Languages and Literatures*, 51(1), 36-69.
- Lucas, R. I. G., e Bernardo, A. B. (2008). Exploring Noun Bias in Filipino-English Bilingual Children. *The Journal of genetic psychology*, 169(2), 149-164.
- Maas, E. L. H. (2014). *The Role of Bilingual Language Input in Children's Receptive and Productive Vocabulary Development*. Tese de Mestrado, Universidade de Utrecht, Utrecht.
- Macnamara, J. (1966). *Bilingualism and primary education*. Edinburgh: Edinburgh University Press.
- MacWhinney, B. (1998). Models of the emergence of language. *Annual review of psychology*, 49(1), 199-227.
- Madeira, A. (2008). Aquisição de L2. In P. Osório e R. M. Meyer (Eds.), *Português Língua Segunda e Língua Estrangeira: Da(s) Teoria(s) à(s) Prática(s)* (pp. 189-201). Lisboa: LIDEL.
- Maital, S. L., Dromi, E., Sagi, A., e Bornstein, M. H. (2000). The Hebrew Communicative Development Inventory: Language specific properties and cross-linguistic generalizations. *Journal of Child Language*, 27, 43 – 67.
- Marchman, V. A., Fernald, A., e Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language*, 37, 817–840.
- Martins, C. (2011). *Manual de análise de dados quantitativos com recurso ao IBM SPSS: Saber decidir, fazer, interpretar e redigir*. Braga: Psiquilíbrios Edições.
- McLaughlin, B. (1978). *Second language acquisition in childhood*. Hillsdale, NJ: Lawrence Erlbaum.
- Meisel, J. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam e L. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss* (pp. 13-40). Cambridge: Cambridge University Press.
- Meisel, J. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In J. Cenoz e F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 11-41). Amsterdam: John Benjamins.

- Meisel, J. (2006). The bilingual child. In T. K. Bhatia e W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 91-113). Malden, MA: Blackwell Publishing
- Meisel, J. (2008). Child second language acquisition or successive first language acquisition?. In B. Haznedar e E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 55–80). Amsterdam: John Benjamins Publishing.
- Meisel, J. (2011). *First and Second Language Acquisition. Parallels and differences*. Cambridge, UK: Cambridge University Press.
- Menyuk, P. (1988). *Language Development: Knowledge and Use*. Glenview: Scott, Foresman and Company.
- Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition*, 5(1), 39–68.
- Montrul, S. (2004) Subject and object expression in Spanish heritage speakers: A case of morpho-syntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125–142.
- Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins Publishing.
- Montrul, S. (2010). Current Issues in Heritage Language Acquisition, *Annual Review of Applied Linguistics*, 30, 3–23.
- Montrul, S. (2012). Is the heritage language like a second language?, *EUROSLA Yearbook*, 12(1), 1-29.
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Montrul S., e Sánchez-Walker N. (2013). Differential Object Marking in child and adult Spanish heritage speakers. *Language Acquisition*, 20(2), 109–132.
- Nation, I. S. P. (2007). Fundamental issues in modelling and assessing vocabulary knowledge. In H. Daller, J. Milton e J. Treffers-Daller (Eds.) *Modelling and assessing vocabulary knowledge* (pp. 35-43). Cambridge: Cambridge University Press.
- Nelson, K. (1973). Structure and Strategy in Learning to Talk. *Monographs of the Society for Research in Child Development*, 38, 1-135.

- Nicoladis, E. (2001). Finding first words in the input. In J. Cenoz e F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 131-147). Amsterdam, Netherlands: John Benjamins.
- Nicoladis, E., e Genesee, F. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology*, 21, 258-270.
- O'Grady, W. (2008). The emergentist program. *Lingua*, 118, 447-464.
- O'Grady, W., Lee, M., e Choo, M. (2001). The acquisition of relative clauses by heritage and nonheritage learners of Korean as a second language. A comparative study. *Journal of Korean Language Education*, 12, 283-294.
- O'Toole, C., Gatt, D., Hickey, T. M., Miękisz, A., Haman, E., Armon-Lotem, S., et al. (2017). Parent report of early lexical production in bilingual children: a cross-linguistic CDI comparison. *International Journal of Bilingual Education and Bilingualism*, 20(2), 124-145.
- Oller, D. K., Pearson, B. Z., e Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28(2), 191-230.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- Owens, R. E. J. (2012). *Language Development. An Introduction* (8th ed.). New Jersey: Pearson.
- Özcan, F. H., Altinkamiş, F. N., e Gillis, S. (2016). Early lexical composition of Turkish-Dutch bilinguals: Nouns before verbs or verbs before nouns. *Poznan Studies in Contemporary Linguistics*, 52(4), 583-604
- Padilla, B., e Ortiz, A. (2012). Fluxos migratórios em Portugal: do boom migratório à desaceleração no contexto de crise. Balanços e desafios. *Revista Interdisciplinar da Mobilidade Humana*, 20(39), 159-184
- Pan, B. A., Rowe, L. M., Singer, J. D., e Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, 76, 763-782.
- Pancsofar N., e Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology*, 27, 571-587.
- Papalia, D. E., Olds, S. W., e Feldman, R. D. (2009). *O Mundo da Criança: Da Infância à Adolescência*. São Paulo: McGraw-Hill.

- Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language Learning*, 60, 651–680.
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic approaches to bilingualism*, 1(3), 213-237.
- Paradis, J., e Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent?. *Studies in Second Language Acquisition*, 18, 1-25.
- Pascual y Cabo, D., e DeLaRosa-Prada, J. (2015). Understanding the Spanish heritage language speaker/learner. *E-JournALL, EuroAmerican Journal of Applied Linguistics and Languages*, 2(2), 1-10.
- Pascual y Cabo, D., e Rothman, J. (2012). The (Il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics*, 33(4), 450-455.
- Patterson, J. (1998). Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *American Journal of Speech-Language Pathology*, 7, 46-56.
- Patterson, J. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Applied Psycholinguistics*, 23, 493–508.
- Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics*, 28(3), 399-410.
- Pearson, B. Z. (2009). Children with two languages. In E. Bavin (Ed.), *The Cambridge Handbook of Child Language* (pp. 379-397). Cambridge: Cambridge University Press.
- Pearson, B. Z., e Fernández, S. C. (1994). Patterns of Interaction in the Lexical Growth in Two Languages of Bilingual Infants and Toddlers. *Language Learning*, 44, 617–653.
- Pearson, B. Z., Fernández, S. C., e Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, 43, 93-120.
- Pearson, B. Z., Fernández, S. C., Lewedeg, V., e Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(1), 41-58.

- Petitto, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tétreault, K., e Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of child language*, 28(2), 453-496.
- Piaget, J. (1926). *The Language and Thought of the Child*. New York: Harcourt Brace & Company.
- Piaget, J. (1967). *La Psychologie de l'Intelligence*. Paris: Armand Colin.
- Piaget, J., e Inhelder, B. (1993). *A Psicologia da Criança*. Porto: Asa.
- Pires, A., e Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211-238.
- Pires, R., Pereira, C., Azevedo, J., Espírito Santo, I., e Vidigal, I. (2016). *Emigração Portuguesa. Relatório Estatístico 2016*. Lisboa: Observatório da Emigração e Rede Migra, CIES-IUL, ISCTE-IUL.
- Place, S., e Hoff, E. (2011). Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Development*, 82, 1834–1849.
- Pokorn, N. K. (2005). *Challenging the traditional axioms: Translation into a non-mother tongue*. Amsterdam: John Benjamins Publishing.
- Polinsky, M. (2005). Word class distinctions in an incomplete grammar. In D. Ravid e H. Shyldkrot (Eds.), *Perspectives on language and language development* (pp. 419-436). Dordrecht: Kluwer.
- Polinsky, M. (2006). Incomplete Acquisition: American Russian. *Journal of Slavic Linguistics*, 14, 161-219.
- Polinsky, M. (2015). Heritage languages and their speakers: State of the field, challenges, perspectives for future work, and methodologies. *Zeitschrift für Fremdsprachwissenschaft*. 26, 7-27.
- Polinsky, M., e Kagan, O. (2007). Heritage languages: In the “wild” and in the classroom. *Language and Linguistics Compass*, 1(5), 368-395.
- Poll, G. H. (2011). Increasing the odds: Applying emergentist theory in language intervention. *Language, Speech, and Hearing Services in Schools*, 42(4), 580-591.
- Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., e Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, 17(1), 57-70.

- Quay, S. (1995). The bilingual lexicon: implications for studies of language choice. *Journal of Child Language*, 22(2), 369–387.
- Ramírez, N. F., Ramírez, R. R., Clarke, M., Taulu, S., e Kuhl, P. K. (2016). Speech discrimination in 11-month-old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science*, 1-16.
- Redlinger, W. E., e Park, T. Z. (1980). Language mixing in young bilinguals. *Journal of Child Language*, 3, 449–455.
- Ribot, K. M., Hoff, E., e Burrige, A. (2017). Language Use Contributes to Expressive Language Growth: Evidence From Bilingual Children. *Child Development*, 1-12.
- Rigolet, S. A. (2000). *Os três P. Comunicação e linguagem para uma plena expressão*. Porto: Porto Editora.
- Rinke, E., e Flores, C. (2014) Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals, *Bilingualism: Language and Cognition*, 17(4), 681–699.
- Romaine, S. (1999). Bilingual language development. In M. Barrett (Ed.), *The development of language* (pp. 251-275). Hove: Psychology Press.
- Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue*. Paris: Champion.
- Rothman, J. (2007). Heritage speaker competence differences, language change and input type: Inflected infinitives in heritage Brazilian Portuguese. *International Journal of Bilingualism*, 11(4), 359–389.
- Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism*, 13(2), 155-163.
- Rothman, J., e Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: heritage speaker bilinguals are natives too!. *Applied linguistics*, 35(1), 93-98.
- dos Santos, C., e Kern, S. (2013, maio). *Early lexical development of French-Portuguese bilingual children: a CDI-adaptation study*. Poster apresentado na Conferência “*Child Language Impairment in Multilingual Contexts*”, Cracóvia, Polónia.
- Scheele, A. F., Leseman, P. P. M., e Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, 31, 117-140.

- Scherer, S., e Souza, A. P. R. (2011). *Types e tokens* na aquisição típica de linguagem por sujeitos de 18 a 32 meses falantes do português brasileiro. *Revista CEFAC*, 13, 838-845.
- Schlichting, L. (1996). *Discovering syntax: An empirical study in Dutch language acquisition*. Nijmegen: Nijmegen University Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Macmillan.
- Schwartz, M. (2010). Family language policy: Core issues of an emerging field. *Applied Linguistics Review*, 1, 171-192.
- Sheng, L., Lu, Y., e Kan, P. F. (2011). Lexical development in Mandarin–English bilingual children. *Bilingualism: Language and Cognition*, 14, 579–587.
- Silva-Corvalán, C. (2014). *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge, UK: Cambridge University Press.
- Sim-Sim, I. (1998). *Desenvolvimento da Linguagem*. Lisboa: Universidade Aberta.
- Singleton, D. (2016). *Language and the Lexicon. An introduction*. New York: Routledge.
- Sinka, I., e Schelleter, C. (1998). Morphosyntactic development in bilingual children. *International Journal of Bilingualism*, 2, 301-326.
- Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Slobin, D. (2006). Cross-Linguistic Comparative Approaches to Language Acquisition. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (Vol. 3, pp. 299-301). Oxford: Elsevier.
- Smithson, L., Paradis, J., e Nicoladis, E. (2014). Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference?. *Bilingualism: Language and Cognition*, 17, 810-821.
- Spinassé, K. (2006). Os conceitos de Língua Materna, Segunda Língua e Língua Estrangeira e os falantes de línguas alóctones minoritárias no sul do Brasil. *Revista Contingentia*, 1, 1-8.
- Spolsky, B. (2009). *Language management*. Cambridge, UK: Cambridge University Press.
- Stern, H. H. (1991). *Fundamental Concepts of Language Teaching*. Oxford: Oxford University Press.

- Stoel-Gammon, C., e Sosa, A. V. (2009). Phonological Development. In E. Hoff e M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 347-367). Malden, MA: Wiley-Blackwell.
- Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, 32, 492 – 504.
- Tardif, T., Shatz, M., e Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, 24, 535 – 565.
- Thomas, E. M., e Gathercole, V. M. (2007). Children's productive command of grammatical gender and mutation in Welsh: An alternative to rule-based learning. *First Language*, 27(3), 251-278.
- Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, 15, 426-445.
- Thordardottir, E. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech and Language Pathology*, 17, 97–114.
- Tomasello, M. (2005). Beyond formalities: The case of language acquisition. *The Linguistic Review*, 22, 183–197.
- Tomasello, M. (2006). Acquiring linguistic constructions. In W. Damon, R. M. Lerner, D. Kuhn, e R. Siegler (Eds.), *Handbook of child psychology: Cognitive perception and language* (Vol. 2, pp. 255–298). Hoboken, NJ: Wiley.
- Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., et al. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6–12 months: language-specific developmental changes in speech perception abilities. *The Emergence of Human Cognition and Language*, 3, 57-61.
- Umbel, V. C., Pearson, B. Z., Fernández, M., e Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, 63, 1012–1020.
- UNESCO (2008). *Mother Tongue Matters: Local Language as a Key to Effective Learning*. Paris: UNESCO.
- Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., e Tsimpli, I. (2014). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics*, 35(4), 765-805.

- Unsworth, S. (2011). *Utrecht bilingual language exposure calculator*. Manuscrito não publicado, Universidade de Utrecht, Utrecht.
- Unsworth, S. (2012). Quantity-oriented and quality-oriented exposure variables in simultaneous bilingual acquisition. In N. de Jong, K. Juffermans, M. Keijzer e L. Rasier (Eds.), *Papers of the Anéla 2012 Applied Linguistics Conference* (pp. 13-22). Delft, The Netherlands: Eburon Uitgeverij B.V.
- Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition*, 16(1), 86-110.
- Unsworth, S. (2016). Quantity and Quality of Language Input in Bilingual Language Development. In E. Nicoladis e S. Montanari (Eds.), *Bilingualism Across the Lifespan* (pp. 103-122). Berlin, Boston: De Gruyter.
- Valdés, G. (2000). Introduction. In L. A. Sandstedt (Ed.), *Spanish for native speakers: AATSP Professional Development Series handbook for teachers K-16* (pp. 1-20). New York: Harcourt College.
- Valdés, G. (2001). Heritage language students: profiles and possibilities. In J. Peyton, J. Ranard, e S. McGinnis (Eds.), *Heritage Languages in America: Preserving a national resource* (pp. 37-80). McHenry, IL: The Center for Applied Linguistics and Delta Systems.
- Volterra, V., e Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, 5, 311-326.
- Weizman, Z. O., e Snow, C. E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, 37, 265-279.
- Werker, J. F., e Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 103, 23-79.
- Xavier, M. F., e Mateus, M. H. (Orgs.) (1990). *Dicionário de Termos Linguísticos* (Vol. 1). Lisboa: Edições Cosmos.
- Xuan, L., e Dollaghan, C. (2013). Language-specific noun bias: evidence from bilingual children. *Journal of Child Language*, 40(5), 1057-1075.

Yang, C. (2002). *Knowledge and learning in natural language*. New York: Oxford University Press.

Yavas, M., Hernandorena, C. L. M., e Lamprecht, R. R. (1991). *Avaliação fonológica da criança*. Porto Alegre: Artes Médicas.



# **Anexos**



## Anexo A

### Autorização

Exmo/a. Senhor/a Encarregado de educação:

Eu, Lílíana Manuela Osório Correia, aluna do Mestrado em Português Língua Não Materna/Língua Estrangeira e Língua Segunda pela Universidade do Minho, venho por este meio solicitar a sua colaboração para o estudo que me proponho realizar.

Este projeto, que se enquadra na área do ensino de Português Língua de Herança e que se intitula *O efeito da exposição linguística sobre a competência lexical de crianças lusodescendentes residentes na Alemanha*, visa estudar crianças que crescem com duas línguas e tem como objetivo melhorar o processo de ensino-aprendizagem da Língua Portuguesa no estrangeiro.

Contudo, este objetivo só será possível com a valiosa colaboração de V. Exa. ao permitir que o seu educando participe no estudo, respondendo a breves questionários e realizando pequenas tarefas.

Os resultados serão confidenciais e apenas utilizados para efeitos da investigação. Se desejar, no final do estudo, poderei deslocar-me à escola para apresentar as suas conclusões.

Aguardando pelo seu consentimento, subscrevo-me com a mais elevada consideração,

Atenciosamente,

\_\_\_\_\_  
(Lílíana Manuela Osório Correia)

Braga, 19 de outubro de 2015

-----  
Eu, \_\_\_\_\_, encarregado/a de educação do/a aluno/a \_\_\_\_\_, aceito participar e autorizo a participação do meu educando no estudo *O efeito da exposição linguística sobre a competência lexical de crianças lusodescendentes residentes na Alemanha*,

Data: \_\_\_\_\_, \_\_\_\_ de \_\_\_\_\_ de 2015

Assinatura: \_\_\_\_\_







## Anexo C

### Questionário sociolinguístico

Caro(a) Encarregado(a) de Educação,

Em primeiro lugar, gostaria de agradecer-lhe por ter aceitado participar no estudo que estou a realizar sobre crianças que crescem com duas línguas, o Português e o Alemão. De facto, não seria possível realizá-lo sem a sua participação e a do seu educando.

Agora, nesta fase do estudo, eu precisava de recolher algumas informações sobre o uso do Português pelo seu educando e pelas pessoas que estão em contacto com ele/ela. Assim, gostaria de pedir-lhe que preenchesse estas informações no seguinte questionário que será essencial para a conclusão deste estudo. Gostaria ainda de salientar que os dados que preencherá serão confidenciais e apenas utilizados para efeitos de investigação. Nenhum dado pessoal será publicado.

Agradecendo antecipadamente a sua colaboração, subscrevo-me com a mais elevada consideração.

Atenciosamente,

Liliana Manuela Osório Correia

### Questionário

*Preencha por escrito a informação em falta e/ou assinale com um X as opções mais adequadas/corretas.*

#### Sobre o Encarregado de Educação

1. Nome: \_\_\_\_\_ . (opcional)
2. Idade: \_\_\_\_\_ anos.
3. Grau/Relação de parentesco com o aluno:  
 Mãe       Pai       Outro. Qual? \_\_\_\_\_.
4. País de nascimento:  
 Portugal       Alemanha       Outro. Qual? \_\_\_\_\_.
5. Se nasceu na Alemanha:
  - a. qual é o país de nascimento dos seus pais?
    - i. Mãe: \_\_\_\_\_
    - ii. Pai: \_\_\_\_\_
6. Se não nasceu na Alemanha, há quantos anos vive nesse/neste país? Há \_\_\_\_\_ anos.
7. De que região de Portugal é originária a sua família? \_\_\_\_\_.
8. Qual é a sua profissão atual? \_\_\_\_\_.
9. Qual é o seu nível de escolaridade?  
 1.º ciclo (1.º - 4.º ano)       2.º ciclo (5.º - 6.º ano)       3.º ciclo ((7.º - 9.º ano)  
 Secundário (10.º - 12.º ano)       Universitário
10. Se não nasceu na Alemanha, frequentou a escola neste/nesse país?  Sim       Não

#### Sobre o seu cônjuge (marido/esposa, companheiro(a))

1. Idade: \_\_\_\_\_ anos.
2. Grau/Relação de parentesco com o seu educando:  
 Mãe       Pai       Outro. Qual? \_\_\_\_\_.

3. País de nascimento:  
 Portugal  Alemanha  Outro. Qual? \_\_\_\_\_.
4. Se não nasceu na Alemanha, há quantos anos vive o seu cônjuge nesse/neste país? Há \_\_\_\_\_ anos.
5. Qual o país de nascimento dos pais do seu cônjuge?  
 i. Mãe: \_\_\_\_\_  
 ii. Pai: \_\_\_\_\_
6. Se algum dos pais do seu cônjuge é português, de que região de Portugal é originário? \_\_\_\_\_.
7. Qual é a profissão atual do seu cônjuge? \_\_\_\_\_.
8. Qual é o nível de escolaridade do seu cônjuge?  
 1.º ciclo (1.º - 4.º ano)  2.º ciclo (5.º - 6.º ano)  3.º ciclo ((7.º - 9.º ano)  
 Secundário (10.º - 12.º ano)  Universitário
9. Se o seu cônjuge não nasceu na Alemanha, ele/ela frequentou a escola neste/nesse país?  
 Sim  Não
10. Se o Português **não** é a língua materna do seu cônjuge, qual é a sua capacidade de compreender e de falar Português? Assinale com um X a resposta mais adequada/correta.

Falar	
<input type="checkbox"/>	Quase nenhuma fluência/ Não é capaz;
<input type="checkbox"/>	Fluência limitada/ Com muitas dificuldades;
<input type="checkbox"/>	Fluência razoável/ Com algumas dificuldades;
<input type="checkbox"/>	Bastante fluência/ Com poucas dificuldades;
<input type="checkbox"/>	Muita fluência/ Quase nenhuma dificuldade;
<input type="checkbox"/>	Fluência nativa/ Sem nenhuma dificuldade.

Entender/Compreender	
<input type="checkbox"/>	Não compreende;
<input type="checkbox"/>	Compreende, mas com muitas dificuldades;
<input type="checkbox"/>	Compreende, mas com algumas dificuldades;
<input type="checkbox"/>	Compreende bem, com poucas dificuldades;
<input type="checkbox"/>	Compreende muito bem, com quase nenhuma dificuldade;
<input type="checkbox"/>	Compreende como um nativo/sem nenhuma dificuldade;

### Sobre o seu educando

1. Nome: \_\_\_\_\_
2. Data de nascimento: \_\_\_\_\_.
3. País de nascimento:  
 Portugal  Alemanha  Outro. Qual? \_\_\_\_\_.
4. Se o seu filho não nasceu na Alemanha, com que idade é que ele/ela chegou à Alemanha?  
 \_\_\_\_\_
5. O seu filho tem irmãos?  Sim  Não
6. Se respondeu sim, quantos irmãos tem e de que idade? \_\_\_\_\_

7. Quando é que o seu filho entrou pela **primeira vez** em contacto com:

a. o Português:

- quando nasceu;
- quando tinha \_\_\_\_ ano(s);

b. o Alemão:

- quando nasceu;
- quando tinha \_\_\_\_ ano(s);

8. Como é que o seu filho entrou em contacto pela primeira vez com o **Português**? Através de:

(pode assinalar mais do que uma opção)

- mãe;
- pai;
- ambos os pais;
- irmãos mais velhos;
- avós;
- escola;
- outros: \_\_\_\_\_.

9. Qual é a capacidade de compreender e de falar **Português** do seu filho? Assinale com um X a resposta mais adequada/correta.

<b>Falar</b>	
<input type="checkbox"/>	Quase nenhuma fluência/ Não é capaz;
<input type="checkbox"/>	Fluência limitada/ Com muitas dificuldades;
<input type="checkbox"/>	Fluência razoável/ Com algumas dificuldades;
<input type="checkbox"/>	Bastante fluência/ Com poucas dificuldades;
<input type="checkbox"/>	Muita fluência/ Quase nenhuma dificuldade;
<input type="checkbox"/>	Fluência nativa/ Sem nenhuma dificuldade.

<b>Entender/Compreender</b>	
<input type="checkbox"/>	Não compreende;
<input type="checkbox"/>	Compreende, mas com muitas dificuldades;
<input type="checkbox"/>	Compreende, mas com algumas dificuldades;
<input type="checkbox"/>	Compreende bem, com poucas dificuldades;
<input type="checkbox"/>	Compreende muito bem, com quase nenhuma dificuldade;
<input type="checkbox"/>	Compreende como um nativo/sem nenhuma dificuldade;

10. Como é que o seu filho entrou em contacto pela primeira vez com o **Alemão**? Através de:

(pode assinalar mais do que uma opção)

- mãe;
- pai;
- ambos os pais;
- irmãos mais velhos;
- avós;
- escola;
- outros: \_\_\_\_\_.



14. Qual a capacidade de **falar Português** de cada uma das pessoas assinaladas? Assinale com um X a resposta mais adequada/correta.

	Não consegue.	Tem muitas dificuldades.	Tem algumas dificuldades.	Tem poucas dificuldades.	Não tem quase nenhuma dificuldade.	Não tem dificuldades.
Mãe						
Pai						
Irmão/Irmã 1						
Irmão/Irmã 2						
Avó materna						
Avó materno						
Avó paterna						
Avó paterno						
Outro: _____						
Outro: _____						

15. O seu filho passa as férias em Portugal?

- Nunca    Poucas vezes    Às vezes    Quase sempre    Sempre

16. Quando o seu filho está de férias, ele comunica em Português:

- Nunca    Raramente    Às vezes    Frequentemente    Sempre

17. As pessoas com quem o seu filho comunica durante as férias são falantes nativos do Português?

- Sim    Não

18. Alguns dos amigos do seu filho falam Português?    Sim    Não

19. Quando o seu filho está com amigos que falam Português, que língua é que eles usam?

- Apenas Português;  
 Mais Português do que Alemão;  
 Português e Alemão em quantidades idênticas/iguais;  
 Mais Alemão do que Português;  
 Apenas Alemão;

20. Há quanto tempo é que o seu filho frequenta a Escola Portuguesa?

- É a primeira vez;  
 Há \_\_\_\_\_ anos.

21. Assinale com um X o tempo que o seu filho passa por semana usando a língua portuguesa nas seguintes atividades:

	Jogar computador	Navegar na internet	Ver televisão	Ouvir música	Ler livros, revistas, etc.	Outra(s)
Nenhum tempo;						
Menos de 1 hora por semana						
1-5 horas por semana						
6-10 horas por semana						
Mais do que 10 horas por semana						

Obrigada pela colaboração!

Liliana Correia

## Anexo D

**Tabela D1 - Dados sociolinguísticos dos participantes bilingues**

Participante	Dados sociolinguísticos										
	Idade	Sexo	País falantes do PE	Região de origem (país)	Idade de emigração (país)		Irmãos mais velhos	Quantidade de <i>input</i> (país)	Quantidade de <i>output</i> (país)	Quantidade de parceiros conversacionais falantes do PE	Quantidade de outras fontes de <i>input</i>
					Mãe	Pai					
FH_1	11	M	ambos	Norte/RAA**	NA	Adulto	Não	100	100	0	3
FH_2	10	F	ambos	Norte	Adulto	Adulto	Sim	200	200	0	2
FH_3	9	F	ambos	Norte	Adulto	Adulto	Sim	200	200	0	1
FH_4	9	M	pai	Centro	NA	Adolescente	Não	100	50	2	3
FH_5	11	F	ambos	Norte/Centro	Adulto	Adulto	Não	200	200	2	3
FH_6	7	M	ambos	Centro	NA	Adulto	Sim	150	150	3	2
FH_7	9	M	ambos	Centro	NA	Adulto	Não	150	100	3	2
FH_8	9	M	ambos	Norte/Lisboa e Vale do Tejo	Adolescente	NA	Sim	150	150	2	3
FH_9	8	M	ambos	Norte/Centro	Adolescente	Adolescente	Não	125	50	1	5
FH_10	9	M	ambos	Norte/Centro	NA	Adulto	Não	175	75	2	3
FH_11	9	F	ambos	Norte/Centro	NA	Adulto	Não	175	75	2	3
FH_12	9	F	ambos	Centro	NA	Adolescente	Não	150	50	2	2
FH_13	7	F	ambos	Norte/Centro	NA	Adulto	Sim	200	200	3	3
FH_14	7	F	ambos	Centro	Adulto	Adulto	Não	200	200	0	5
FH_15	8	M	ambos	Centro	NA	Adolescente	Sim	150	50	2	2
FH_16	8	F	ambos	Norte	Adolescente	Adulto	Não	200	150	0	3
FH_17	8	F	ambos	Norte	Adulto	Adulto	Sim	200	150	0	2
FH_18	6	F	ambos	Norte	Adolescente	Adulto	Sim	200	150	0	3
FH_19*	11	F	mãe	Norte	Adulto	NA	Não	75	50	0	2
FH_20	8	M	mãe	Centro	Criança	NA	Não	25	25	2	0
FH_21	7	M	ambos	Norte/Centro	NA	NA	Sim	100	50	4	2
FH_22	6	M	mãe	Centro	Criança	NA	Sim	75	50	2	0
FH_23	9	M	ambos	Norte/Centro	NA	NA	Não	100	50	4	2

Nota: M = Masculino; F = Feminino; NA = não se aplica \* Por ser um *outlier*, o participante FH\_19 não entra nos testes estatísticos de associação entre os fatores de *input* e o desempenho lexical dos FH; \*\* Região Autónoma dos Açores.



## Anexo E

Resultados dos testes estatísticos de normalidade e de homogeneidade

**Tabela E1** – Resultados dos testes de normalidade Kolmogorov-Smirnov e Shapiro-Wilk referentes à distribuição dos dados do *corpus* total e dos *subcorpora* por grupo de participantes (com *outlier*).

<i>Corpora</i>	Grupo	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Estatística	gl	Sig.	Estatística	gl	Sig.
Nomes, Adjetivos e Verbos	Falantes de Herança	,068	23	,200*	,980	23	,913
	Falantes Monolíngues	,107	21	,200*	,970	21	,740
Nomes	Falantes de Herança	,132	23	,200*	,946	23	,242
	Falantes Monolíngues	,143	21	,200*	,922	21	,093
Verbos	Falantes de Herança	,085	23	,200*	,982	23	,938
	Falantes Monolíngues	,152	21	,200*	,951	21	,357
Adjetivos	Falantes de Herança	,131	23	,200*	,953	23	,332
	Falantes Monolíngues	,123	21	,200*	,943	21	,246

\*. Este é um limite inferior da significância verdadeira.

a. Correlação de Significância de Lilliefors

**Tabela E2** – Resultados dos testes de normalidade Kolmogorov-Smirnov e Shapiro-Wilk referentes à distribuição dos dados do *corpus* total e dos *subcorpora* por grupo de participantes (sem *outlier*).

<i>Corpora</i>	Grupo	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Estatística	gl	Sig.	Estatística	gl	Sig.
Nomes, Adjetivos e Verbos	Falantes de Herança	,081	22	,200*	,975	22	,817
	Falantes Monolíngues	,107	21	,200*	,970	21	,740
Nomes	Falantes de Herança	,122	22	,200*	,944	22	,236
	Falantes Monolíngues	,143	21	,200*	,922	21	,093
Verbos	Falantes de Herança	,098	22	,200*	,988	22	,992
	Falantes Monolíngues	,152	21	,200*	,951	21	,357
Adjetivos	Falantes de Herança	,149	22	,200*	,936	22	,162
	Falantes Monolíngues	,123	21	,200*	,943	21	,246

\*. Este é um limite inferior da significância verdadeira.

a. Correlação de Significância de Lilliefors

**Tabela E3** – Resultados dos testes de homogeneidade das variâncias dos dois grupos em comparação (com *outlier*).

		Estatística de Levene	gl1	gl2	Sig.
Nomes, Adjetivos e Verbos	Com base em média	,014	1	42	,905
	Com base em mediana	,019	1	42	,891
	Com base em mediana e com gl ajustado	,019	1	41,067	,891
	Com base em média aparada	,014	1	42	,907
Nomes	Com base em média	,183	1	42	,671
	Com base em mediana	,094	1	42	,760
	Com base em mediana e com gl ajustado	,094	1	41,981	,760
	Com base em média aparada	,148	1	42	,702
Verbos	Com base em média	,297	1	42	,589
	Com base em mediana	,387	1	42	,537
	Com base em mediana e com gl ajustado	,387	1	41,846	,537
	Com base em média aparada	,316	1	42	,577
Adjetivos	Com base em média	,466	1	42	,499
	Com base em mediana	,333	1	42	,567
	Com base em mediana e com gl ajustado	,333	1	42,000	,567
	Com base em média aparada	,435	1	42	,513

**Tabela E4** – Resultados dos testes de homogeneidade das variâncias dos dois grupos em comparação (sem *outlier*).

		Estatística de Levene	gl1	gl2	Sig.
Nomes, Adjetivos e Verbos	Com base em média	,154	1	41	,697
	Com base em mediana	,125	1	41	,726
	Com base em mediana e com gl ajustado	,125	1	40,768	,726
	Com base em média aparada	,151	1	41	,700
Nomes	Com base em média	1,227	1	41	,275
	Com base em mediana	,712	1	41	,404
	Com base em mediana e com gl ajustado	,712	1	38,326	,404
	Com base em média aparada	1,020	1	41	,318
Verbos	Com base em média	,068	1	41	,796
	Com base em mediana	,136	1	41	,715
	Com base em mediana e com gl ajustado	,136	1	40,390	,715
	Com base em média aparada	,077	1	41	,783
Adjetivos	Com base em média	1,499	1	41	,228
	Com base em mediana	1,068	1	41	,308
	Com base em mediana e com gl ajustado	1,068	1	38,285	,308
	Com base em média aparada	1,460	1	41	,234

**Tabela E5** – Resultados dos testes de normalidade Kolmogorov-Smirnov e Shapiro-Wilk referentes aos fatores de *input* (sem *outlier*).

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	gl	Sig.	Estatística	gl	Sig.
<i>idade</i>	,184	22	,051	,930	22	,120
<i>input_pais</i>	,198	22	,025	,865	22	,006
<i>output_pais</i>	,203	22	,019	,851	22	,004
<i>ativ_PE</i>	,237	22	,002	,873	22	,009
<i>par_conv_PE</i>	,244	22	,001	,858	22	,005

a. Correlação de Significância de Lilliefors



## Anexo F

Resultados das comparações intragrupois por método *pairwise* de *Bonferroni*

**Tabela F1** – Resultados das comparações por método *pairwise* de *Bonferroni* dentro do grupo de FH.

(I) Categoria gramatical	(J) Categorias gramaticais	Diferença média (I-J)	Erro	Sig. <sup>b</sup>	95% Intervalo de Confiança para Diferença <sup>b</sup>	
					Limite inferior	Limite superior
1 - Nomes	2 - Verbos	14,522*	1,589	,000	10,404	18,639
	3 - Adjetivos	17,957*	1,894	,000	13,049	22,864
2 - Verbos	1 - Nomes	-14,522*	1,589	,000	-18,639	-10,404
	3 - Adjetivos	3,435*	,722	,000	1,564	5,305
3 - Adjetivos	1 - Nomes	-17,957*	1,894	,000	-22,864	-13,049
	2 - Verbos	-3,435*	,722	,000	-5,305	-1,564

Baseado em médias marginais estimadas

\*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: *Bonferroni*.

**Tabela F2** – Resultados das comparações por método *pairwise* de *Bonferroni* dentro do grupo de FM.

(I) Categoria gramatical	(J) Categorias gramaticais	Diferença média (I-J)	Erro	Sig. <sup>b</sup>	95% Intervalo de Confiança para Diferença <sup>b</sup>	
					Limite inferior	Limite superior
1 - Nomes	2 - Verbos	24,810*	2,417	,000	18,495	31,124
	3 - Adjetivos	31,381*	2,641	,000	24,482	38,280
2 - Verbos	1 - Nomes	-24,810*	2,417	,000	-31,124	-18,495
	3 - Adjetivos	6,571*	1,303	,000	3,167	9,976
3 - Adjetivos	1 - Nomes	-31,381*	2,641	,000	-38,280	-24,482
	2 - Verbos	-6,571*	1,303	,000	-9,976	-3,167

Baseado em médias marginais estimadas

\*. A diferença média é significativa no nível ,05.

b. Ajustamento para diversas comparações: *Bonferroni*.

