

Article

# The Role of Input Factors in the Lexical Development of European Portuguese as a Heritage Language in Portuguese–German Bilingual Speakers

Liliana Correia <sup>1</sup>  and Cristina Flores <sup>2,\*</sup> 

<sup>1</sup> Centro de Estudos Humanísticos, Instituto de Letras e Ciências Humanas, Universidade do Minho, Campus de Gualtar, Braga 4710-057, Portugal; lilianamocorreia@gmail.com

<sup>2</sup> Departamento de Estudos Germanísticos e Eslavos, Instituto de Letras e Ciências Humanas, Universidade do Minho, Campus de Gualtar, Braga 4710-057, Portugal

\* Correspondence: cflores@ilch.uminho.pt; Tel.: +351-253-601658

Academic Editors: Antje Muntendam, Usha Lakshmanan and Osmer Balam

Received: 28 July 2017; Accepted: 7 December 2017; Published: 21 December 2017

**Abstract:** In light of previous research on early bilingualism, this study investigates whether 6–11-year-old child heritage speakers (HSs) of European Portuguese (EP), living in Germany, show patterns of lexical development similar to those of monolingual EP children, both in terms of vocabulary size and of lexical composition. Moreover, it assesses the role of factors related to the quantity and quality of the input in the HSs' lexical development in EP. Twenty-three bilingual 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 (nouns, verbs, adjectives), which served as a basis for between- and within-group comparisons. Information regarding the HSs' language experience was collected by means of a parental questionnaire. Results revealed significant between-group differences concerning the total corpus and the subcorpora of nouns and verbs. Within-group comparisons showed that both groups produced significantly more nouns than verbs and more verbs than adjectives. Correlation analyses revealed that the HSs' lexical knowledge is significantly correlated with the input and output quantity at home as well as with the number of EP-speaking parents. Parents emerge as the key players in the acquisition of EP as a heritage language.

**Keywords:** heritage speakers; European Portuguese; input effects; vocabulary size; lexical composition

## 1. Introduction

Over the last decades, a large body of empirical research has emerged within the field of early bilingual lexical acquisition, with a considerable amount of literature published on the bilingual development of heritage speakers (HSs) from different parts of the world (see, for instance, [1–3]). Heritage speakers are usually second- or third-generation migrants who grow up exposed both to the family language, the heritage language (HL), and to the majority societal language, either simultaneously or sequentially, during early childhood. Being a sociopolitically minority language, the HL is typically acquired under reduced input conditions in restricted language environments (i.e., at home and in immigrant communities), whereas the language of the host/dominant society—the majority language—is acquired in multifaceted settings: it is the language of schooling, of the media and of the social interactions outside (and also inside) the home environment. In general, HSs are more intensively exposed to the HL in the first years of life. However, this pattern of language exposure quickly changes when these children start attending preschool institutions and, consequently, start receiving more systematic input in the majority language. Due to the great diversity of sociolinguistic contexts in which bilingual language acquisition takes place and to the individual variation observed

in the levels of language proficiency attained by bilingual speakers in each of their languages, many of the studies on early bilingualism have attempted to verify the extent to which children exposed to two (or more) languages from an early age follow patterns of lexical development similar to those of their monolingual counterparts [4–14]. Additionally, a large number of studies have also sought to determine the role of language experience in bilingual lexical development [15–26].

Within this scope, it has been empirically proven that, when both of their languages are taken into account, bilingual children go through the same stages of language acquisition and reach basic milestones of language development within the same age span as monolingual children. Therefore, no language delay is shown due to the fact that they are acquiring two languages instead of one [7,9,10,12,27].

Empirical research has also found that, similar to monolingual lexical acquisition, there is also an asymmetry between the receptive and the productive lexical competence in bilingual children's lexical development. In fact, several studies conducted with children with different language pairs have provided evidence that their receptive lexical repertoires are consistently larger than their productive ones [6,9,27–30].

In addition, research on bilingual children's early lexical composition has revealed that their early vocabularies present patterns of lexical distribution resembling those of their monolingual counterparts. Thus, like the early lexicons of monolingual children, those of bilingual children are also initially composed of a large proportion of social words which are quickly outnumbered by lexical items belonging to the grammatical category nouns, presenting at the same time a gradual increase in the proportion of verbs, adjectives and function words [31].

Nevertheless, it has been widely reported in the literature that bilingual children do not usually develop both of their languages at the same rate [27,32,33]. In fact, a large number of studies have shown that, when only one of their languages is assessed, these children tend to have smaller lexical repertoires than their monolingual peers [9,15,17,18,21,34].

According to Oller et al. [35], the differences observed in the vocabulary sizes of bilingual children in each of their languages are due to the fact that they are usually exposed to each language in different settings, resulting in the distribution of their lexical knowledge across the two languages—the so-called “distributed characteristic” of bilingual knowledge [36,37]. Therefore, it is not surprising that the vocabulary these children know in a given language is usually restricted to the context in which its acquisition typically occurs. Thus, just as it is normal for bilingual children to have, for example, a larger vocabulary of academic terms in the language of schooling (usually the majority language) than in the language spoken at home, it is also normal for them to show a better knowledge of vocabulary related to the home environment in the language that is spoken in this context [38,39].

In fact, the outcomes of several studies on the effect of input factors on the lexical development of young bilingual children have shown that variation in the rate of lexical acquisition in each language is significantly associated with factors related to their language experience, namely the quantity [15,16,20,22,23,40] and quality [24,41,42] of input.

With respect to the minority language, it is commonly stated that a child heritage speaker starts receiving less input in the HL with the onset of schooling (i.e., immersion in the school system of the dominant society) [43,44]. Nevertheless, research on the HL of children in the first years of schooling is not as common as research on earlier stages of HL development. Moreover, many studies have focused on the different degrees of ultimate attainment of adult heritage speakers in their minority language, with some authors having found, for example, that the HSs' lexical knowledge is correlated with their levels of proficiency in other linguistic domains such as that of morphosyntax (see, for instance, [45]). Therefore, it is of utmost importance to study the specific period of HL development between early childhood and adolescence/adulthood. Thus, in light of the extant research on bilingual lexical acquisition, and given the scarcity of studies on the lexical development of HSs of European Portuguese (EP), the present study aims at investigating whether 6- to 11-year-old Portuguese–German bilingual children present patterns of lexical development similar to those of their monolingual EP counterparts, both in terms of vocabulary size and of lexical composition. Additionally, and more

importantly, it aims at evaluating the role of extra-linguistic factors related to the quantity and quality of the input in the bilingual children's lexical development with respect to their HL (i.e., EP).

The present paper is organised as follows. Sections 1.1 and 1.2 review some of the most significant findings that have emerged from empirical research on the composition of the early lexicons of bilingual children and on the effects of input factors on bilingual lexical development, respectively. Section 2 presents the research questions and predictions. In Section 3, we describe the materials and procedure adopted for the data collection (Sections 3.1, 3.3 and 3.4), and the sociolinguistic profile of the participants (Section 3.2). In Section 4, we display the results of the statistical analyses performed on the collected data and finally, in Section 5, we discuss them in light of the research questions, presenting at the same time future research directions.

### 1.1. Research on Bilingual Children's Early Lexical Composition

Despite the fact that child-directed speech comprises items of every grammatical category [46] and that verbs seem to be the most frequent open-class words used in spoken discourse by adult speakers of several languages such as EP, Spanish and French [47],<sup>1</sup> research on children's early lexical development has revealed that in initial stages of language acquisition the early lexicons of monolingual children from different language communities are mainly composed of nouns. Although some studies conducted with monolingual child speakers of Korean [49–52] and Mandarin [53–55] show contradictory results, the predominance of nouns over other grammatical categories such as verbs, adjectives and closed-class words in early vocabularies (the so-called “noun bias”) has been observed in monolingual child speakers of English [50,56,57], Italian [50,58,59], Spanish [50,60], Hebrew [50,61], French [50,62], Dutch [50,63], Navajo [64], and Brazilian Portuguese (BP) [65], among other languages. Having verified the presence of larger proportions of nouns than other word classes in the early lexicon of children from different language communities, some authors, such as Gentner [66], argue that the noun bias is universal and that semantic and perceptual factors underlie the predisposition of children to acquire nouns faster than, for instance, verbs. According to Gentner [66], the fact that verbs and other predicate terms, in comparison to concrete nouns, have semantically more complex structures and a less transparent relation regarding the perceptual world makes it easier for children to identify concepts coded by nouns than those coded by verbs.

Following this line of investigation, research on bilingual children's lexical development has attempted to determine whether their early vocabularies present the same patterns of lexical composition as those of their monolingual counterparts, with several studies showing a large proportion of nouns in early bilingual lexicons of different language pairs [4,5,14,29,67–70]. David and Li [5], for instance, found that the productive lexicons of 13 French–English bilingual children (12–36-month-olds) comprise, in each language, a higher proportion of nouns in comparison to lexical items belonging to the other grammatical categories. The only exception was the first stage of language production (i.e., productive lexicon below 50 words) in which social words were predominantly produced by the children. Regarding the other grammatical categories, the children produced more predicates than closed-class items. Taken together, the results showed that the investigated bilinguals follow patterns of lexical development similar to those of their monolingual counterparts. Dos Santos and Kern [14], in turn, found that the lexical repertoires of 29 Portuguese–French bilingual children (24–36-month-olds) showed patterns of lexical distribution identical to those of monolingual child speakers of French (i.e., the majority language), that is, their productive lexicons comprised a large number of nouns, which are followed, in descending proportions, by predicates, closed-class words, and social words. However, their lexicon in the minority language (i.e., EP) diverged from the lexical distribution observed in the majority language. Although it is mainly composed of nouns and predicates, with nouns outnumbering predicates, the children's productive lexicon in

---

<sup>1</sup> See, however, [[48]] for different results concerning the English language.

EP presented a greater proportion of social words in comparison to that of items belonging to the closed-class category. The authors suggested that the difference observed between the languages of the Portuguese–French bilingual children is due to the fact that the minority language is not at the same level of linguistic development as the majority language.

### *1.2. Research on the Role of Input Factors in Bilingual Children's Lexical Development*

Even though the amount and type of language exposure may vary during the process of monolingual language acquisition (see, for instance, the study by Hart and Risley [71]), all normally developing monolingual children end up acquiring the language they are exposed to since birth. However, the same is not true for all normally developing bilingual children simultaneously exposed to more than one language from an early age. While some children growing up in dual language environments become fully bilingual, others do not attain full language proficiency in both of their languages, becoming linguistically far more competent in one language than the other [3,72]. Since bilingual children have to divide their language exposure between their two (or more) languages and its amount tends to be unequally distributed across the languages, a large body of research has attempted to understand the role that variation in the degree and type of language exposure plays in bilingual language acquisition.

A major contribution to the understanding of how variation in the amount of language exposure affects the lexical development of bilingual children was made by Thordardottir [26]. The author analysed the receptive and productive vocabulary development of 49 5-year-old French–English simultaneous bilingual children, in each of their languages, by comparing the children to each other and to monolingual child speakers of French and English. Results revealed that their productive lexical competence is more vulnerable to effects of reduced input than the receptive one: whereas the children with balanced input in both languages (i.e., 40–60%) present lexical outcomes comparable to those of their monolingual peers on measures of receptive vocabulary, only the children with levels of exposure higher than 60% to a given language score similarly to monolingual children on measures of productive vocabulary.

In a different study, Sheng et al. analysed the lexical knowledge of two groups of Mandarin–English bilingual children (3–5-year-olds; 6–8-year-olds) in picture identification and picture naming tasks in each of their languages [30]. Unlike the bilingual children in Thordardottir's study [26], who were resident in Montreal (Canada) where both languages share the same socio-political status, these children were growing up in the U.S., where English is the majority language and Mandarin is a minority language. Results showed that the younger children's total conceptual vocabulary size is larger than their single-language vocabulary sizes, whereas the dimension of the older children's composite vocabulary is very similar to that of their English vocabulary. Moreover, both groups performed more accurately on the identification task than on the naming task and the difference between the children's accuracy level in both tasks was larger in Mandarin than in English. Taken together, these results indicate that, with age, the rate of lexical acquisition steadily increases in the majority language, while it seems to precociously stagnate in the minority language. According to Sheng et al. [30], although these children are predominantly exposed to Mandarin in the home environment, the fact that they do not have contact with other input sources in Mandarin in the dominant society leads to this lack of lexical equilibrium between the minority and the majority language.

One of the factors which has been related to the quantity of language exposure is the presence of older siblings in the household, namely in families in which both parents are speakers of a minority language. According to several authors, the presence of older siblings usually alters the patterns of language use in the family, mainly because older siblings tend to introduce the majority language in the home environment [16,73–76], which may compromise the intergenerational transmission of the minority language. In two separate studies with Spanish–English bilingual children growing up in the U.S., Bridges and Hoff [16] found that not only do bilingual toddlers receive more input in English from older siblings than from other family members, but also that children with older

siblings are more advanced in the majority language than the other participants without older siblings. Furthermore, results showed that, in comparison to children with siblings attending school, children without school-aged siblings are more proficient in the minority than in the majority language.

Another factor that has been claimed to predict bilingual children's language development is their own output [40,77–81]. According to Pearson [82], children's own language production in a target language (TL) invites more input in that same language, generating a cycle in which "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). Thus, if children do not orally express themselves, for example, in the HL, then this process may reverse itself, resulting in less language exposure and, consequently, in less linguistic proficiency in the minority language. Bohman et al. [78], for example, found that the amount of output of Spanish–English bilingual children is positively correlated with their performance on tests of semantic and morphosyntactic knowledge in Spanish and in English.

Although most of the research on the effects of the quantity of input on bilingual language development has focused on the children's current amount of language exposure, several researchers have attempted to assess the role of language exposure over time [79,81,83–86]. According to Unsworth [85], due to the fact that "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" (p. 91), when calculating these children's cumulative length of exposure (CLoE) to each language, one should take into account every and any variation in the amount of input that may have occurred in the years prior to data collection (for a detailed description on how to calculate CLoE, see [85,87]). Thus, taking Unsworth's proposal into consideration, some studies have found CLoE to be predictive of bilingual lexical development, that is, the greater the CLoE to a TL, the better the children's language performance in it [79,81,84].

Besides varying in amount, bilingual language exposure also varies in its quality. Among the multiple factors that seem to influence the richness of language exposure are the diversity of interlocutors and the regular interaction with native and non-native speakers [24,88,89], as well as the variety of different input sources, such as the media and literature, and the plurality of activities in which the TLs are used [41,42,90–93].

In fact, in a study conducted with 29 Spanish–English bilingual children, Place and Hoff found not only that the amount of input in each language is a predictor of the children's language development both in Spanish and in English, but also that specific properties of their English input, such as the number of exclusively English-speaking interlocutors, the variety of sources of English input, and the amount of English input from native speakers, are positively correlated with their lexical knowledge in this language [24]. In a different study with immigrant families living in the U.S., Ishizawa [94] found that bilingual children living with both their parents and non-English-speaking grandparents in three-generation households are more likely to use a minority language than children living with only their parents or grandparents. Regarding the role of different input sources such as the media and literature in the language development of bilingual children, Scheele et al. observed that, for example, the amount of input through the activity of storytelling (but not through that of shared book reading) in the first language (L1) of bilingual children from Moroccan-Dutch and Turkish-Dutch families living in The Netherlands is related to their receptive lexical knowledge in their L1 [42]. Moreover, although L1 input through music-related activities and educational TV watching does not correlate with their L1 vocabulary, singing in the L2 (i.e., in Dutch) is correlated with these children's L2 receptive vocabulary.

## 2. Research Questions

The present study assessed the lexical knowledge of 44 EP speakers, between the ages of 6 and 11, in a semi-spontaneous oral production task. The 44 participants were grouped into two different sets: an experimental group of 23 HSs of EP attending classes of Portuguese as a Heritage Language (PHL) in Germany, and a control group of 21 monolingual speakers of EP attending primary school in Portugal.

All the participants in the control group had taken part in a previous project, “EFFE—Escreves como falas—falas como escreves?” (see [95,96]), from which their data were extracted and analysed.

The aim of this study is two-fold: first, we aim to verify, empirically, whether HSs of EP present the same patterns of productive lexical development as their monolingual EP counterparts, both in terms of vocabulary size and in terms of lexical composition; and second, we intend to identify the input factors which may predict the lexical development of the bilingual children.

The first research question (RQ 1) and corresponding subquestions (RQs 1.1 and 1.2) to be addressed are the following:

1. Do HSs of EP, between the ages of six and 11, show patterns of lexical development identical to their monolingual Portuguese peers attending primary school in the country of origin?
  - 1.1. That is, do HSs of EP use the same quantity of different lexical items (i.e., lemmas) as their monolingual Portuguese peers in the oral description of the same thematic picture?
  - 1.2. Is it possible to identify between- and within-group similarities/differences regarding the quantitative distribution of nouns, verbs and adjectives?

Based on previous studies showing that bilingual children’s lexical repertoires tend to be smaller than those of the corresponding monolingual children when only one of their languages is assessed [9,13,15,17,18,21], we predict that, in the experimental task, the HS of EP will produce a smaller quantity of different lexical items (i.e., lemmas) when compared to that of their monolingual EP-speaking peers, regarding both the total lexical production (i.e., total corpus) and the lexical production per grammatical category (i.e., subcorpora of nouns, verbs and adjectives).

On the other hand, given previous findings on the composition of bilingual children’s early productive vocabularies demonstrating that bilinguals present the same patterns of lexical distribution as their corresponding monolingual counterparts [5,14], we predict that the HSs’ lexical repertoires will not significantly differ from those of monolingual EP-speaking children concerning the proportion of lexical items belonging to the grammatical categories of nouns, verbs and adjectives.

The second research question (RQ 2) and corresponding subquestions (RQs 2.1 and 2.2) are as follows:

2. What is the role of different input factors related to the quantity and quality of language exposure in the HSs’ lexical development?
  - 2.1. Do HSs whose language exposure to the HL is greater and more diversified show a better lexical performance than those whose language exposure is more reduced and less diversified?
  - 2.2. If there are individual differences within the bilingual group due to the amount and type of language exposure, which variables influence these children’s lexical development in the HL the most regarding their total lexical production as well as their lexical production per grammatical category?

We assume that the HSs’ lexical knowledge, assessed in the semi-spontaneous oral production task, will be characterised by individual variation—i.e., on the one hand, there will be children whose lexical knowledge will resemble that of monolingual EP-speaking children and, on the other hand, there will be children whose lexical knowledge will differ from that of their monolingual counterparts. Based on this assumption, we predict that extra-linguistic factors related to the quantity and quality of language exposure will play a major role in the development of the HSs’ lexical competence.

Thus, bearing in mind the studies on the effect of input factors on the lexical development of bilingual children presented in Section 1.2, we predict that the lexical knowledge of the experimental group will be significantly influenced by the following factors:

- The amount of exposure to the HL within the home environment (INPUT\_HOME), that is, we expect that the children who are linguistically more exposed to EP in the home environment (by means of parental input) will show a better lexical performance than the children who are less exposed to the minority language in the same context.
- The amount of output in the HL within the home environment (OUTPUT\_HOME), that is, we expect that the children who use EP more frequently in the home environment (when addressing their parents) will show a better lexical performance than the children who do not usually use the TL in this context.
- The quantity of EP-speaking parents (EP\_PARENTS), that is, we expect that the children whose parents are both native EP speakers will show a better lexical performance than the children who only have one native EP-speaking parent.
- The quantity of native EP-speaking conversational partners in the (same) migration context (CONV\_PARTNERS), that is, we expect that the children who regularly interact with a larger number of native EP speakers (i.e., grandparents and uncles/aunts) in the migration context will show a better lexical performance than the children whose contact with these input sources is more limited or even non-existent.
- The quantity of leisure activities in which EP is used (ACTIV\_EP), that is, we expect that the children who regularly participate in a larger number of ludic activities (i.e., watching television, listening to music, playing computer games, etc.) using the HL will show a better lexical performance than the children whose contact with these input sources is reduced or non-existent.
- The presence of older siblings in the household (SIBLINGS), that is, we expect that the children who do not have older siblings will show a better lexical performance than the children who do have them.
- The heritage speakers' age (AGE), that is, we expect that the older children will show a better lexical performance than the younger ones.

### 3. Materials and Methods

#### 3.1. Sociolinguistic Questionnaire

A detailed sociolinguistic questionnaire, following Gutiérrez-Clellen and Kreiter [83] and Unsworth [85], was created and filled in by one of the parents of each bilingual child. The questionnaire consisted of three sections: the first two sections were centred on biographic and sociolinguistic information about the child's parents and the last section on the child him- or herself.

Sections 1 and 2 comprised questions about the parents' age, place of birth, current occupation, and level of education. In addition, parents were asked to inform about the number of years living in the host country (i.e., Germany) and, in the event of having been born there, they were asked to mention their parents' country of origin.

Regarding the last section, the questionnaire included questions about the child's age, place of birth, number of years living in Germany, number of siblings as well as their ages. Parents were also asked to provide information not only about the child's age of first exposure to both languages and the context in which exposure took place (i.e., at home or in the school environment), but also about the people with whom the child interacts regularly and the frequency with which these people speak EP to the child and vice versa. Moreover, parents reported on the regularity with which the child spends holidays in Portugal and on the frequency with which he/she uses the HL in this context. Furthermore, parents were asked to indicate the language(s) spoken by the child when among EP-speaking friends in the migration context, the number of years the child has been attending PHL classes, as well as the number of leisure activities, such as watching television, listening to music, reading, etc., in which the child regularly takes part using the HL.

The detailed parental questionnaire described above allowed us not only to obtain crucial information to outline the sociolinguistic profile of the HSs, but also to collect data for the variables to be used in the statistical analyses centred on the impact of input factors on the HSs' lexical development.

Thus, based on the information provided by the parents, we calculated the following: (1) the child's amount of exposure to the HL within the home environment, by adding the quantity of input the child receives from each parent, whose value can vary between the minimum of zero and the maximum of 100, resulting in a scale from zero to 200 (i.e., the amount of input from both parents taken together); (2) the child's amount of output in the HL within the home environment, by adding the quantity of output the child uses to address each parent, resulting in a scale of measurement similar to the previous one; (3) the quantity of EP-speaking parents, by counting the number of parents who are native speakers of EP, which resulted in a binary scale composed of the values "both" and "mother or father"; (4) the quantity of native EP-speaking conversational partners in the same migration context (i.e., grandparents and uncles/aunts), by counting the number of native interlocutors who interact in EP with the child on a regular basis; (5) the quantity of leisure activities (i.e., watching television, listening to music, reading, etc.) in which the child uses the HL, by counting the number of ludic activities in which the child participates on a regular basis using the HL. The information provided by the parents also allowed us to verify whether factors such as (6) the presence of older siblings in the household and (7) the child's age at the time of testing play a role in the bilingual children's lexical development of the HL.

The linguistic background of the bilingual participants based on the variables listed above will be described in detail in the next subsection.

### 3.2. Participants

#### 3.2.1. Bilingual Speakers

The sample of HSs of EP was selected from a group of students attending 90-min classes of Portuguese Language and Culture, once a week and as an extracurricular activity, organised and sponsored by Camões Institute.

As selection criteria, the following parameters were adopted: (1) having Portuguese/luso-descendant parents (one or both of them); (2) having been born in Germany or emigrated to this country before the age of five; (3) being between six and 12 years old; and (4) having the caregivers' consent to participate in the study.<sup>2</sup>

Bearing in mind these criteria, 23 luso-descendant children, aged between six and 11 years (mean ( $M$ ) = 8.48; standard deviation ( $SD$ ) = 1.44), were selected to constitute the experimental group. The 23 bilingual participants, i.e., 12 male (52.2%) and 11 female (47.8%) children, were all resident in Stuttgart, Germany at the time of testing and were being raised in a bilingual environment, in which EP is the HL, mainly spoken within the home environment by the nuclear family, and German is the majority language, spoken in multifaceted settings. Twenty-one out of the 23 bilingual children were born in Germany and two migrated to the host country before the age of one. All the children were exposed to EP since birth and 16 out of the 23 children were also exposed to German before the age of three. The remaining children ( $n = 7$ ) began receiving systematic input in the majority language between the ages of three and five, when they started attending preschool institutions.

With regard to their household composition, 19 out of the 23 children (82.6%) were born to families in which both parents are native speakers of EP and only a small number of children ( $n = 4$ ; 17.4%) belong to households in which just one of the parents speaks EP as his/her native language. Moreover, most of the children ( $n = 21$ ; 91.3%) have siblings, but only 10 HSs have older siblings.

---

<sup>2</sup> The children's caregivers signed a consent form allowing their child to participate in the study.

Considering these family constellations, we observed that, within the home environment, the children's amount of language exposure to the HL varies according to the sociolinguistic background of the family, that is, children whose parents are both native EP speakers are more exposed to EP ( $M = 164.47$ ;  $SD = 37.56$ ) than those children who have only one native EP-speaking parent ( $M = 68.75$ ;  $SD = 31.46$ ).

As for the parents' educational level, it ranged from low (i.e., Primary Education) to high (i.e., Tertiary Education). However, most of the bilingual children's parents only completed Primary Education (i.e., six years of schooling) or Secondary General Education (i.e., 12 years of schooling).

A common feature of these bilingual children is the attendance of HL classes. Nevertheless, due to its extracurricular status, the number of years of schooling in the minority language differs from child to child, even within the same age range. Thus, 22 children began schooling in the HL between the ages of six and seven, and one child started attending PHL classes at the age of nine.

Besides being exposed to EP within the home environment and in the extracurricular EP classes, 15 out of the 23 children interact regularly with other native EP speakers, usually members of the extended family, such as grandparents and uncles/aunts, who live in the same migration context. With respect to the languages spoken when among EP-speaking friends living in the host country, the majority of the bilingual children ( $n = 14$ ; 60.9%) use more German than EP. As for the remaining HSs, one child only uses EP; two children use more EP than German; two children use equal amounts of EP and German; and four children only use German.

As far as other sources of language exposure are concerned, most of the bilingual children ( $n = 21$ ; 91.3%) were exposed to EP from the media and literature on a regular basis.

Furthermore, 11 HSs always spend holidays in Portugal, nine HSs almost always go on holidays to Portugal and three HSs only sometimes visit the country of origin during school holidays. When they visit Portugal, most of the HSs ( $n = 13$ ; 56.5%) always use EP in their social interactions, four HSs usually use it, three HSs only sometimes communicate in EP, and one HS seldom speaks in EP in this context. Two parents did not give information about their children's language use when visiting the country of origin.

As for the language background data of the seven input factors selected for the inferential statistics regarding the role of these variables in the lexical development of the bilingual children (i.e., INPUT\_HOME, OUTPUT\_HOME, CONV\_PARTNERS, ACTIV\_EP, EP\_PARENTS, SIBLINGS, and AGE), the data we will present refer to only 22 HSs, since an outlier was detected.

Concerning the amount of exposure to the HL within the home environment, it ranges from a minimum value of 25 to a maximum value of 200 ( $M = 151.14$ ;  $SD = 50.28$ ) and its median value is 150.

With respect to the amount of output in the HL within the home environment, it also varies from a minimum value of 25 to a maximum value of 200 ( $M = 112.50$ ;  $SD = 63.03$ ). However, the data distribution is different from that of the previous variable (i.e., INPUT\_HOME): the median value is 100.

As far as the quantity of native EP-speaking conversational partners in the same migration context is concerned, the number of native EP speakers (i.e., grandparents and uncles/aunts) with whom the bilingual children regularly interact ranges from a minimum of zero to a maximum of four ( $M = 1.64$ ;  $SD = 1.33$ ). Its distribution is as follows: seven out of the 22 participants (31.8%) do not regularly have contact with native EP speakers; one participant (4.5%) regularly interacts with only one native EP speaker; nine participants (40.9%) frequently have contact with two native EP speakers; three participants (13.6%) interact with three native EP speakers on a regular basis; and two participants (9.1%) usually have contact with four native EP speakers.

Regarding the quantity of leisure activities in which EP is used, the number of ludic activities in which the bilingual children frequently participate using the HL varies between a minimum of zero and a maximum of five ( $M = 2.45$ ;  $SD = 1.22$ ); that is, two out of the 22 participants (9.1%) do not usually take part in leisure activities in which EP is used; one participant (4.5%) uses EP in only one

leisure activity; eight participants (36.4%) use it in two activities; nine participants (40.9%) use it in three activities; and two participants (9.1%) use it in five activities.

As for the quantity of EP-speaking parents, 19 out of the 22 participants (86.4%) were born to families in which both parents are native speakers of EP and only three participants (13.6%) belong to households in which just one of the parents speaks EP as his/her native language.

With regard to the presence of older siblings in the household, 10 out of the 22 participants (45.5%) have older siblings.

Concerning the heritage speakers' age, the participants' ages range from 6–11 years ( $M = 8.36$ ;  $SD = 1.36$ ). The distribution is as follows: two participants (9.1%) are six years old; four participants (18.2%) are seven; five participants (22.7%) are eight; eight participants (36.4%) are nine; one participant (4.5%) is 10; and two participants (9.1%) are 11 years old.

The language background data of each participant (i.e., age, gender, number of EP-speaking parents, presence of older siblings in the household, amount of language input and output in the home environment, number of native EP-speaking conversational partners, and other sources of language input) can be consulted in Table A1 (Appendix A).

### 3.2.2. Monolingual Speakers

The control group was selected from a group of 92 monolingual children attending the second grade of primary school in Portugal. All the children were participants in project EFFE (see [95,96]), and participated in the project's oral narrative-descriptive task regarding the thematic picture *Sala* 'living room' (see Figure A1 in Appendix B). Twenty-one monolingual EP children, aged between seven and 10 years ( $M = 7.20$ ;  $SD = 0.70$ ), were selected to constitute the control group of the present study (11 male (52.38%) and 10 female (47.62%)). The children were resident either in Lisbon ( $n = 11$ ) or in Oporto ( $n = 10$ ) and, just like in the bilingual group, their parents' levels of education were heterogeneous, ranging from very low (i.e., Illiterate) to high (i.e., Tertiary Education).

### 3.3. Semi-Spontaneous Oral Production Task

So that a corpus of semi-spontaneous oral production could be built for the comparative analyses between the lexical knowledge of both groups under investigation, the HS group took part in a semi-spontaneous oral production task identical to the one in which the control group participated under the project EFFE.

Therefore, for the experimental task, we used the picture *Sala* 'living room' (see Figure A1 in Appendix B), chosen from a set of five thematic pictures—*Casa de Banho* 'bathroom', *Cidade* 'city', *Cozinha* 'kitchen', *Floresta* 'forest' and *Sala* 'living room'. These pictures, originally developed by Yavas et al. [97] to assess the phonological competence of BP-speaking children over the age of three, were adapted to EP by Guerreiro [98] and used by project EFFE's research team to evaluate oral and written language skills of EP-speaking children attending primary school. The project EFFE's research group selected these pictures because they contained key elements that trigger the use of crucial target words for the building of a corpus made of oral and written data from EP-speaking children in the first years of schooling [96]. Our choice for the picture *Sala* 'living room' was mainly due to the fact that it contains key elements for the production of vocabulary related to the context in which the acquisition of HLs usually takes place—the home environment.

The oral production data were collected in individual sessions, at a single time point and without a time limit, in the schools where the PHL classes were taking place. At the beginning of each session, children were given an A4-sized printed version of the selected picture and were presented a fictitious situation in which they had to imagine they had just arrived home and found the scene represented in the picture. Besides being asked to narrate the events prior to that moment and to indicate possible reasons for the scene presented, the children were encouraged to name and describe elements displayed in the picture by means of questions such as "What do you see here?", "What is the father doing?", "How is the living room?", among others, so that they would produce as many lexical items as possible.

By including both a narrative and a descriptive component in the experimental task, we aimed at eliciting lexical items of different grammatical categories, namely nouns, verbs and adjectives. The oral data were collected by one of the researchers who is also a native speaker of EP. All the participants were informed that the researcher did not speak German so that they would not use this language during the experimental task.

### 3.4. Corpus Building

To build the corpus of semi-spontaneous oral production, the digital audio recordings of both the HS group and the control group were orthographically transcribed and the lexical items belonging to the grammatical categories nouns, verbs and adjectives were selected and entered into three databases. The inflected forms of the produced lexical items were lemmatised so that they would constitute a single entry in the databases. If a participant produced, for example, three different inflected forms of a verb, such as *canto* 'I sing', *cantei* 'I sang' and *cantarei* 'I will sing', these were counted as a single lexical item (that is, *cantar* 'to sing').

Moreover, words that were metonymically used, such as *Playstation* for *consola de jogos* 'video game console' or *Monopoly* for *jogo de tabuleiro* 'board game', as well as foreign words, such as *bordeaux* or *puzzle*, constitute entries in the databases. However, lexical items which were produced either in English or in German and which are not used by the EP-speaking population were excluded from the corpus. Additionally, multilexical units, such as *dar conta de* 'to realise', appear as a single entry in the corpus due to the fact that they have a unitary meaning.

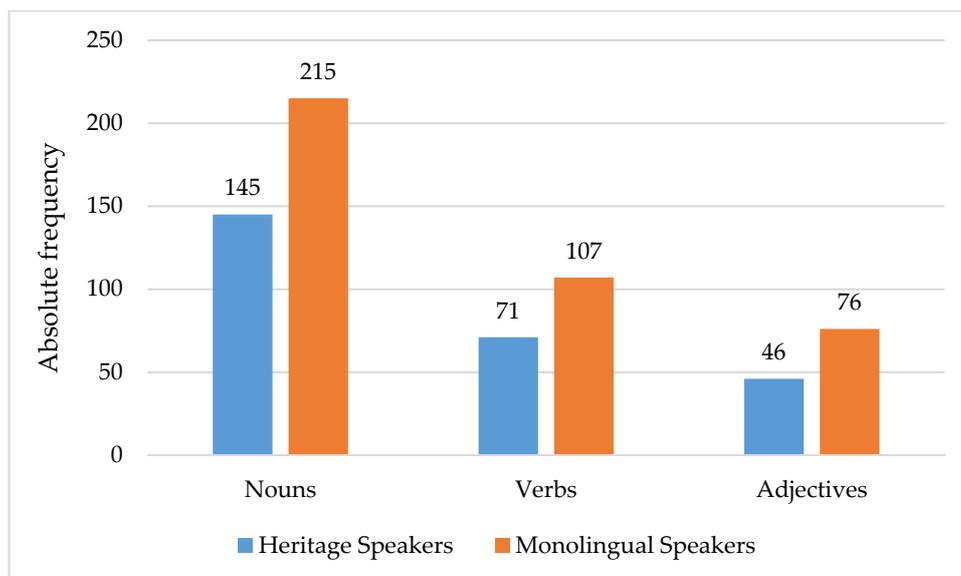
Bearing in mind these criteria, we initially built two corpora: one with the lexical items produced by the experimental group and the other with the oral data of the control group. These corpora, each one composed of three subcorpora (nouns, verbs and adjectives), were used to measure the vocabulary sizes of both groups under investigation. Subsequently, these corpora were merged into a single corpus for the between- and within-group comparisons regarding the lexical knowledge assessed in the experimental task.

## 4. Results

In this section, we will present the results of the descriptive and inferential statistical analyses of the data collected from the semi-spontaneous oral production task, which allowed us not only to assess the HSS' lexical knowledge in the HL, but also to evaluate the effect that the different input factors have on their lexical development. All the statistical analyses were performed using the statistical software IBM SPSS (IBM, Armonk, NY, USA, version 23.0).

### 4.1. Semi-Spontaneous Oral Production Task

Figure 1 presents the total number (i.e., absolute frequency) of lexical items produced by each group in each grammatical category. By analysing it in detail, one can see that the HSS' corpus is composed of 145 nouns (55.3%), 71 verbs (27.1%), and 46 adjectives (17.6%), resulting in a total number of 262 lexical items. The monolingual speakers' corpus is, in turn, composed of 215 nouns (54.0%), 107 verbs (26.9%), and 76 adjectives (19.1%), resulting in a total number of 398 lexical items. The combination of the two corpora results in a total corpus of 487 lexical items, since words occurring in both corpora were counted only once. These are distributed as follows: 269 nouns (55.2%), 127 verbs (26.1%), and 91 adjectives (18.7%). In this total corpus, 91 out of the 269 nouns (33.8%), 51 out of the 127 verbs (40.2%), and 31 out of the 91 adjectives (34.1%) were commonly produced by both groups. The lexical distribution per group of EP speakers, illustrated in Figure 1, shows that the monolingual speakers produced more lexical items in every grammatical category than the HSS and that both groups produced more nouns than verbs and more verbs than adjectives.



**Figure 1.** Quantity of lexical items produced in each grammatical category per group of European Portuguese (EP) speakers.

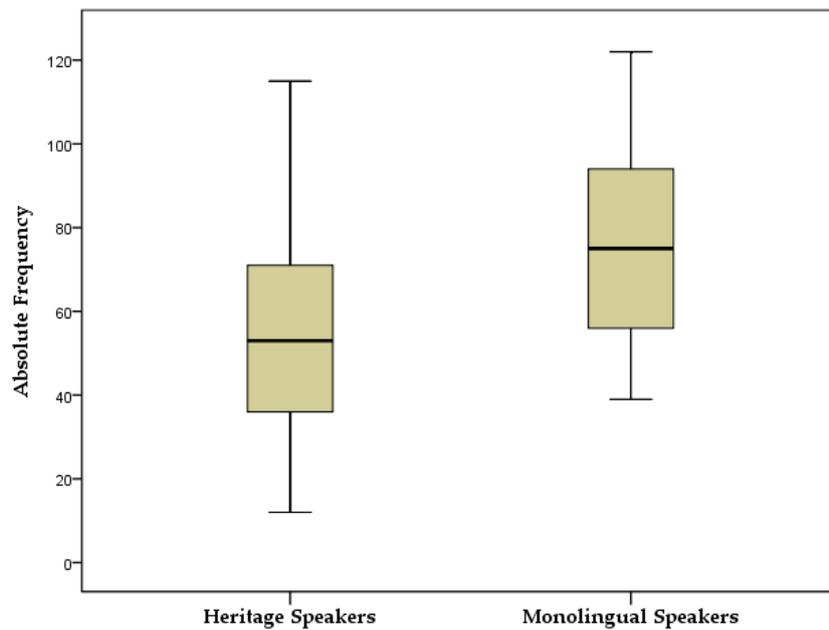
An analysis of the absolute frequencies shows that, on average, the HSs produced 53.74 ( $SD = 25.21$ ) lexical items, whereas the monolingual speakers produced 77.52 ( $SD = 23.41$ ). Regarding the mean number of lexical items per grammatical category, the HSs produced on average 28.74 ( $SD = 13.36$ ) nouns, 14.22 ( $SD = 7.21$ ) verbs and 10.78 ( $SD = 5.68$ ) adjectives, while the monolingual speakers produced 44.57 ( $SD = 14.24$ ) nouns, 19.76 ( $SD = 6.49$ ) verbs and 13.19 ( $SD = 6.29$ ) adjectives.

The data distribution displayed in Figure 2 shows the lexical variation across groups and within each group in relation to the total number of lexical items (i.e., nouns, verbs and adjectives) produced per group. As can be seen from the boxplot regarding of the HSs' data, the middle value of the sample (i.e., the median) stood at 53 lexical items and its interquartile range (IQR) was between 32 and 72 lexical items. The boxplot also shows that the bilingual sample presented a minimum value of 12 lexical items and a maximum of 115. On the other hand, the boxplot concerning the monolingual children's data shows that the median of the control group stood at 75 lexical items and that its IQR was between 55.5 and 96 lexical items. Moreover, the boxplot also shows that the minimum and maximum values of the control group were 39 and 122 lexical items, respectively. From the data described above, we can observe that the values of the control group were higher than those of the experimental group in every measure (i.e., median, IQR, minimum and maximum).

Additionally, the data distribution displayed in Figure 3 and in Table 1 allows us to discriminate the variation previously described per grammatical category. As far as nouns and verbs are concerned, the HSs presented lower values than the control group in every observable measure (i.e., median, IQR, maximum and minimum), although the differences seemed to be more pronounced in the grammatical category of nouns. However, the two groups did not seem to differ from each other regarding the use of adjectives.

First, exploratory analyses of the data were performed in order to verify whether they met the assumptions required to run parametric tests<sup>3</sup>. Kolmogorov–Smirnov and Shapiro–Wilk tests showed that the data were approximately normally distributed. Levene's tests, in turn, revealed that variances were equal for both groups.

<sup>3</sup> According to Field [99], the assumptions underlying parametric tests are (1) normally distributed data; (2) homogeneity of variance; (3) interval data; and (4) independence.



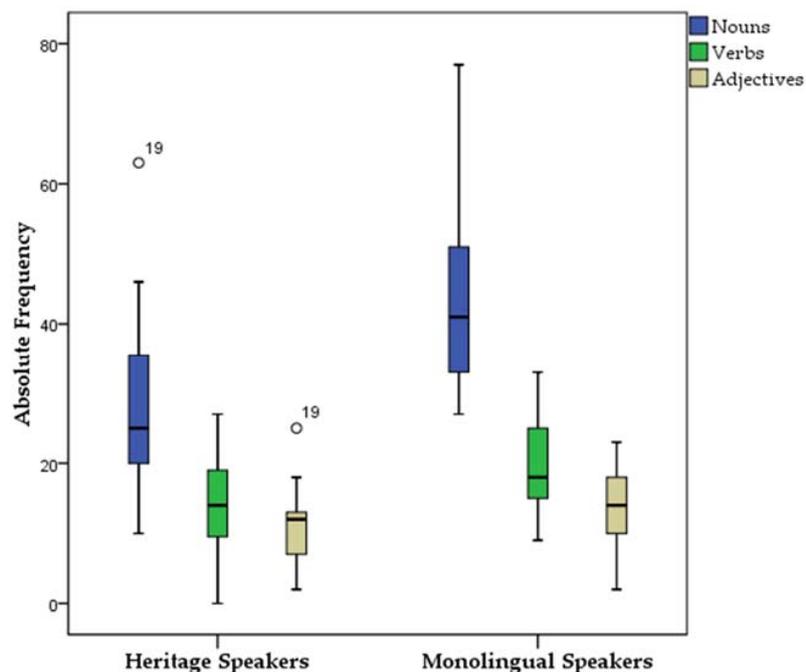
**Figure 2.** Boxplot: Lexical distribution and variation per group of EP speakers regarding the total corpus.

Thus, so as to examine whether the experimental group differed significantly from the control group in relation to the quantity of lexical items produced in the experimental task, parametrical *t*-tests for independent samples were applied not only to the total corpus, but also to the three subcorpora (i.e., nouns, verbs and adjectives). The results revealed that there were significant differences between the HSs and the monolingual children regarding the total corpus ( $t(42) = -3.23, p = 0.002$ ), that is, the experimental group produced significantly fewer lexical items than the control group. As for the subcorpora, the results of the *t*-tests showed that there were significant between-group differences regarding the grammatical categories nouns ( $t(42) = -3.81, p < 0.001$ ) and verbs ( $t(42) = -2.67, p = 0.011$ ), that is, the HSs of EP produced fewer nouns and fewer verbs than their monolingual counterparts. However, no statistical differences were found between both groups of participants regarding the use of adjectives ( $t(42) = -1.33, p = 0.189$ ). Additionally, since an outlier was identified in the subcorpora nouns and adjectives, further *t*-tests were run without this atypical value in order to evaluate whether its presence influenced the results of the statistical analyses. The results without the outlier showed that there was still a significant difference between the bilingual and monolingual children regarding the grammatical category nouns ( $t(41) = -4.44, p < 0.001$ ) and that both groups did not significantly differ from each other with respect to the subcorpus adjectives ( $t(41) = -1.79, p = 0.082$ ).

**Table 1.** Lexical distribution and variation per group of European Portuguese (EP) speakers regarding the subcorpora.

	Nouns		Verbs		Adjectives	
	HS	MS	HS	MS	HS	MS
Median	25	41	14	18	12	14
Minimum	10	27	0	9	2	2
Maximum *	46/63	77	27	33	18/25	23
IQR	20–36	32.5–55	9–20	15–25.5	6–13	8–18.5

Note: HS = Heritage Speakers; MS = Monolingual Speakers; IQR = Interquartile Range; \* Due to the existence of an outlier in the HS group regarding the grammatical categories of nouns and adjectives, the first (the outlier's) and second highest values are presented.



**Figure 3.** Boxplot: Lexical distribution and variation per group of EP speakers regarding the subcorpora.

In a next step, the analysis focused on the identification of within-group differences regarding the quantitative distribution of the lexical items produced per grammatical category. A repeated-measures analysis of variance (ANOVA) was performed for each group of EP speakers. Since the data did not meet the assumption of sphericity ( $p < 0.05$ ), the Greenhouse–Geisser correction values were used to test for significance (see [99]). The results of the ANOVAs indicated that there were statistically significant differences both within the experimental group ( $F(1.204, 26.489) = 82.18, p < 0.001$ ) and within the control group ( $F(1.389, 27.788) = 113.24, p < 0.001$ ) regarding the quantity of lexical items produced per grammatical category. Further Bonferroni pairwise comparisons (see Table A2 in Appendix C) revealed that, in the HS group, noun production was significantly superior to verb production (mean difference ( $MD$ ) = 14.522,  $p < 0.001$ ); and to adjective production ( $MD = 17.952, p < 0.001$ ); verb production was also significantly superior to adjective production ( $MD = 3.435, p < 0.001$ ). As for the monolingual group, the difference between the mean number of items produced in the category nouns was also significantly superior to verbs ( $MD = 24.81, p < 0.001$ ) and to adjectives ( $MD = 31.381, p < 0.001$ ). Verbs were also used significantly more than adjectives ( $MD = 6.561, p < 0.001$ ).

In a final step, we centred our analysis on the frequency distribution of the lexical items per grammatical category in each group. Figures 4 and 5, which depict the distribution of word frequency per grammatical category and per group of speakers, allow us to have a wider perspective of the EP-speaking children's lexical performance. A closer look at Figure 4 shows that, regarding the HSs' corpus, only a small percentage of nouns (7.6%), of verbs (4.2%) and of adjectives (6.5%) had frequencies equal to or higher than 75% (i.e., were produced by 75% or more of the participants), whereas a large percentage of nouns (76.6%), of verbs (73.2%) and of adjectives (65.2%) presented frequencies lower than 25% (i.e., were produced by less than 25% of the participants). With respect to the monolingual speakers' corpus, the data displayed in Figure 5 also show that only a small percentage of nouns (9.8%), of verbs (4.7%) and of adjectives (3.9%) had frequencies equal to or higher than 75%, while a large percentage of nouns (77.7%), of verbs (76.6%) and of adjectives (80.3%) presented frequencies lower than 25%. From the data described above, one can see that the distribution of word frequency per grammatical category was very similar across both groups of EP speakers. Besides, the fact that the majority of the lexical items produced in both groups in the experimental task were not consistently used by all speakers shows that the participants' lexical knowledge is very diverse.

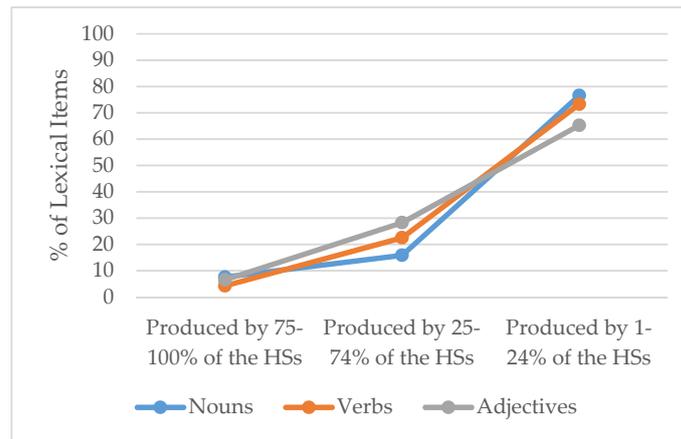


Figure 4. Distribution of word frequency per grammatical category in the heritage speakers (HS) group.

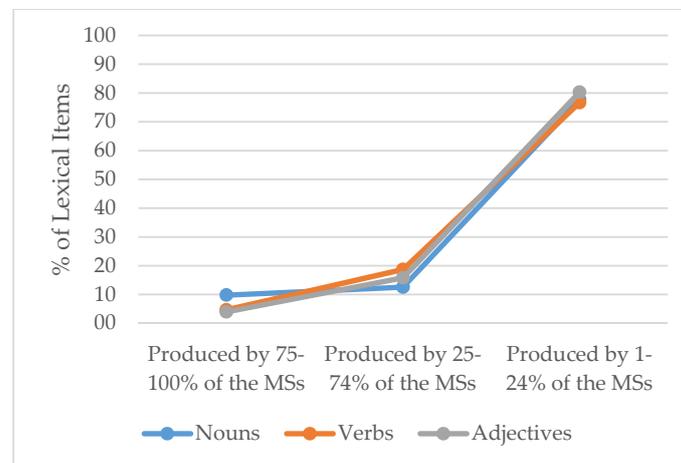


Figure 5. Distribution of word frequency per grammatical category in the monolingual group.

However, an analysis of the nouns, verbs and adjectives with the 10 highest frequencies within each group of speakers (see Table A3 in Appendix D) showed that, despite presenting different frequencies, most of the lexical items that occupy the highest ranks were commonly produced by bilingual and monolingual children (in the total corpus). In fact, in the category nouns, all the items were common to both groups; in the category verbs, only one item (i.e., *ficar* ‘to be/stay/become’) used by eight monolingual children was not produced by any of the HSs; and, in the category adjectives, only two items (i.e., *lilás* ‘lilac’ and *português* ‘Portuguese’) used by the bilingual group were not produced by any monolingual speaker and only two items (i.e., *escuro* ‘dark’ and *estranho* ‘strange’) used by the monolingual children were not produced by any HS.

As for the total corpora, 57.7% of all the nouns, 52.3% of all the verbs, and 59.2% of all the adjectives used by the monolingual children were not produced by any bilingual child; whereas only 37.2% of all nouns, 28.2% of all verbs and 32.6% of all adjectives used by the HSs were not produced by any monolingual speaker. That is, there were lexical items of the three grammatical categories that were only produced by one of the groups. The monolingual group showed the highest percentage/amount of items only produced by one group.

#### 4.2. Relation between the Input Factors and the HSs’ Lexical Knowledge

So as to avoid biased results, the outlier was excluded from the correlation analyses between the input factors and the HSs’ lexical knowledge. The sample is, henceforth, composed of 22 participants.

Before statistically analysing the influence of the seven input factors selected for this study on the HSs’ lexical development, we performed exploratory analyses of the interval variables (i.e., INPUT\_HOME; OUTPUT\_HOME; CONV\_PARTNERS; ACTIV\_EP; and AGE) in order to verify whether the data met the assumptions required to perform parametric tests for correlation. The results of the Kolmogorov–Smirnov and Shapiro–Wilk tests showed that only the variable AGE had normally distributed data (i.e.,  $p > 0.05$ ). Thus, the Pearson correlation coefficient ( $r$ ) was applied to the interval variable AGE and the Spearman’s correlation coefficient ( $r_s$ ) was used with the remaining interval variables whose data were not normally distributed (i.e., INPUT\_HOME; OUTPUT\_HOME; CONV\_PARTNERS; and ACTIV\_EP) (following Field [99]). As for the dichotomous nominal variables EP\_PARENTS and SIBLINGS, the point–biserial correlation coefficient ( $r_{pb}$ ) was applied to the correlation analysis with these variables.

Table 2 displays the results of the correlation tests between each of the input factors and the HSs’ lexical performance assessed in the experimental task regarding their total lexical production (i.e., total corpus) and their lexical production per grammatical category (i.e., subcorpora nouns, verbs and adjectives).

**Table 2.** Correlations between the input factors and the lexical performance of the HSs.

Correlation Coefficient	Input Factors	Total Corpus	Subcorpora		
			Nouns	Verbs	Adjectives
$r$	AGE	0.40	0.34	0.43	0.41
	INPUT_HOME	0.63 **	0.56 **	0.65 **	0.67 **
$r_s$	OUTPUT_HOME	0.60 **	0.55 **	0.67 **	0.54**
	CONV_PARTNERS	−0.38	−0.44	−0.36	−0.26
	ACTIV_EP	0.39	0.40	0.42	0.28
$r_{pb}$	EP_PARENTS	0.46 *	0.38	0.48 *	0.51 *
	SIBLINGS	−0.33	−0.36	−0.24	−0.31

Note:  $r$  = Pearson;  $r_s$  = Spearman;  $r_{pb}$  = Point-biserial; Significance level: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

An analysis of the correlation coefficients and their significance levels revealed that the variables INPUT\_HOME and OUTPUT\_HOME were positively correlated with the HSs’ vocabulary size regarding both the total corpus and the three subcorpora, that is, the HSs who were more exposed to EP and who interacted more frequently in EP within the home environment produced more lexical items in every context under analysis (i.e., total corpus and subcorpora) than those who were less exposed to EP and who communicated less frequently in this language within the same context. Moreover, the results of the point–biserial correlation tests showed that the variable EP\_PARENTS was significantly and positively correlated with the HSs’ lexical scores regarding the total corpus and the subcorpora verbs and adjectives; that is, the HSs whose parents were both native EP speakers presented a better lexical performance with respect to the total corpus and the subcorpora verbs and adjectives than those who only had one native EP-speaking parent. This association suggests that the HSs whose parents were both native speakers of EP consequently received more input and produced more output in the HL, resulting, therefore, in a higher level of language proficiency. In fact, a correlation test between the amount of exposure to the HL within the home environment and the quantity of EP-speaking parents revealed that the two variables were positively correlated with each other ( $r_{pb} = 0.68, p < 0.001$ ). In addition, a correlation test between the amount of output in the HL within the home environment and the number of parents who were native speakers of EP also showed that there was a positive correlation between these variables ( $r_{pb} = 0.46, p < 0.05$ ). The remaining input variables (CONV\_PARTNERS; ACTIV\_EP; AGE; and SIBLINGS) were not significantly correlated with the HSs’ performance in the experimental task.

To sum up, the results revealed that, on the one hand, there were input factors that were significantly correlated with the HSs’ productive lexical knowledge in the HL and that, on the other hand, there were also input factors which were not found to significantly correlate with these children’s lexical knowledge.

## 5. Discussion

In the present section, we will discuss the findings of the present study in light of the research questions on which it is founded. As far as the first research question (RQ 1) and subquestion RQ 1.1 are concerned, the results revealed that the lexical knowledge of Portuguese–German bilingual children, aged between six and 11 years, differed from that of monolingual EP children attending primary school in the country of origin. Specifically, the group of HSs of EP produced significantly fewer lexical items (i.e., lemmas) than the group of monolingual EP speakers in the same oral narrative-descriptive task. Thus, one can conclude that, similar to previous findings on the lexicon size of bilingual children with different language pairs [9,13,15,17,18,21], the Portuguese–German bilingual children also have smaller lexical repertoires than their monolingual EP peers when only one of their languages is assessed.

Regarding the subquestion 1.2, the results showed that the between-group differences observed in the overall size of the lexical repertoires lay mainly in the categories of nouns and verbs, but not in that of adjectives. Thus, the bilingual children produce significantly fewer nouns and verbs than the monolingual children, but both groups of EP speakers produce similar quantities of adjectives. This contradicts our prediction that the bilingual children would produce fewer lexical items than their monolingual counterparts in every grammatical category. The absence of inter-group differences concerning the production of adjectives may be related to the fact that the studies on the composition of the bilingual children’s lexicon [5,14], on which some of this study’s predictions were founded, grouped the lexical items belonging to the grammatical categories verbs and adjectives into the category predicates. Thus, it is not possible to discriminate the exact proportion of verbs and adjectives in the children’s early vocabularies. Nevertheless, Figure 3 and Table 1 show that, in comparison to the other grammatical categories, both bilingual and monolingual children produced a small quantity of adjectives in the experimental task. Therefore, this may be one of the reasons for the statistically non-significant difference between both groups regarding the category adjectives.

However, despite the differences between the experimental and the control group concerning their total production of nouns and verbs, the within-group statistical analyses revealed that the lexical repertoires of both groups of EP speakers presented identical patterns of lexical distribution with respect to the proportion of lexical items belonging to the three grammatical categories under analysis; that is, both groups produced significantly more nouns than verbs and more verbs than adjectives in the experimental task. This finding is consistent with the results of previous research conducted with bilingual and monolingual children from different language communities [4,5,14,29,50,56,60], according to which the lexical repertoires of the investigated children are predominantly composed of nouns. Our results further suggest that, with respect to lexical composition, the “noun bias” observed in the early lexicons of monolingual and bilingual children from diverse language populations continues in subsequent stages of language development in EP-speaking children. However, further studies should be carried out so that one can verify whether the noun bias still stands with different measures of vocabulary assessment (i.e., types, tokens, etc.) and different thematic pictures, or/and whether, for example, like in spoken corpora of EP-speaking adults (see, for instance, [100]), the frequency of verbs is higher than that of nouns.

A comparison of the words used by both groups showed that there was a considerable overlap of the lexical items used by monolingual and by bilingual children. For instance, the words with the ten highest frequencies were, in every grammatical category, almost the same in both groups. However, the heritage speakers also produced a significant amount of lexical items that monolingual children did not use (and vice versa), so we can conclude that the bilingual children possess a rich, diverse lexical repertoire on their own that is not a small subpart of the monolinguals’ repertoire.

On the other hand, and in line with recent conclusions drawn by authors such as Gathercole [2], Montrul [3], and Unsworth [101], the descriptive analysis of our data showed that the lexical knowledge of the Portuguese–German bilingual children is characterised by individual variation, i.e., there were HSs of EP who presented the same patterns of lexical development as their monolingual EP counterparts, whereas others diverged. This conclusion can be confirmed by the re-examination

of Figures 2 and 3. In these side-by-side boxplots, one can observe that there were values of the experimental group which overlapped with those of the control group, while there were also values of the bilingual group which did not fall within the data range of the monolingual group. For instance, if we consider the control group's lexical range (i.e., 39–122), in Figure 2, as representative of the patterns of monolingual EP-speaking children's lexical development and take into account both the lexical range (i.e., 12–115) and the median value of the HSs' total corpus (i.e., 53), one can conclude that at least 50 percent of the bilingual children have vocabulary sizes which fall within the patterns of lexical development of their monolingual counterparts.

Since the prediction that the HSs' lexical knowledge would be characterised by individual variation has been confirmed, it became essential to determine the extent to which input factors affect the development of the HSs' lexical competence in the HL.

Within this scope, and as far as the second research question (RQ 2) and corresponding subquestions (RQ 2.1 and RQ 2.2) are concerned, the results of the correlational analyses between the lexical performance of the HSs and the seven input factors selected for the present study revealed that only the variables INPUT\_HOME, OUTPUT\_HOME and EP\_PARENTS were significantly and positively correlated with the HSs' lexical performance in the experimental task. The remaining input factors (i.e., CONV\_PARTNERS; ACTIV\_EP; AGE; and SIBLINGS) were not found to be significantly correlated with the children's lexical knowledge in the HL.

Thus, similar to previous research on the effects of the quantity of language input and output on the productive lexical competence of bilingual children with other TLs [15,22,24,40,79], the results of the present investigation showed that the lexical knowledge of the HSs of EP was significantly influenced by the quantity of language exposure and production which occurred during the verbal interactions within the home environment. The finding that the HSs who received more input in EP had, overall, better lexical outcomes than those who were less exposed to EP further supports the idea that the HL must play a major role in the daily interactions within the home environment if the intergenerational transmission of the family language is to be successful. It is important to note that, in bilingual sociolinguistic contexts—in which one of the child's languages (i.e., the majority language) is spoken by the dominant society in multifaceted communication contexts and the other child's language is, in most cases, used in more restricted contexts—the minority language is more vulnerable to the effects of reduced input than the societal language.

Besides showing positive correlations between the HSs' lexical performance and the amount of input and output within the home environment, the results also revealed that the quantity of EP-speaking parents was positively correlated with the HSs' total lexical repertoires, specifically with their vocabulary sizes regarding the grammatical categories verbs and adjectives; that is, the HSs whose parents were both native EP speakers produced more verbs and adjectives than those who only had one native EP-speaking parent. Therefore, one can conclude that the HSs need to receive more input and to produce more output in the HL to acquire verbs and adjectives than to acquire nouns. This fact may be related to Gentner's theoretical assumption [66] that it is easier for children to identify concepts coded by nouns than those coded by verbs and other predicates, since the latter have semantically more complex structures and a less transparent relation regarding the perceptual world. Thus, the positive correlations between the HSs' lexical performance and the variables INPUT\_HOME and OUTPUT\_HOME, on the one hand, and between these two variables and the variable EP\_PARENTS, on the other hand, seem to suggest that, in order to establish the complex semantic-perceptual relations that the acquisition of lexical items belonging to predicates requires, the HSs need to be more exposed to EP and to communicate more frequently in this language.

Moreover, the existence of positive correlations between the HSs' lexical performance and the number of EP-speaking parents, and between the latter variable and the amount of exposure to the HL within the home environment is in line with the findings by Place and Hoff [24]. They observed that the lexical knowledge of HSs of Spanish is significantly influenced by the amount of language exposure to the minority language, with the children with two native Spanish-speaking parents presenting larger

vocabularies in Spanish than the children with only one native Spanish-speaking parent, although the difference was not found to be statistically significant.

Based on previous research [24,41,42,91,92,94], we also predicted that the HSs' language performance would be significantly influenced by qualitative input factors such as the number of native EP speakers in the same migration context (i.e., grandparents and uncles/aunts) with whom the children interact on a regular basis as well as the quantity of leisure activities in which the children regularly participate using the HL. However, the statistical results showed no significant correlations between the HSs' lexical knowledge and the variables CONV\_PARTNERS and ACTIV\_EP. Thus, the results discussed so far seem to suggest that, with the bilingual children under investigation, the amount of input and output within the home environment play a more important role in the development of their productive lexical competence (in the HL) than the regular contact with other native speakers of EP and other input sources closer to standard EP.

Furthermore, contrary to what we predicted, the presence of older siblings in the household did not have an effect on the development of the HSs' productive lexical competence. Since, in the present study, neither the amount of input from older siblings nor the amount of output to them were taken into account, we consider it important for future research to include these variables in the statistical model so that one can have a clearer perspective of the effect that the presence of older siblings in the family has on the bilingual children's language environment and, consequently, on their heritage language competence.

Finally, regarding the factor AGE, the results revealed that chronological age was not a predictive factor of the HSs' lexical development. This may be due to the fact that the age of the bilingual children, (i.e., between six and 11 years old), may not be a representative measure of the quantity of input that they actually received over time. In fact, if we consider Unsworth's proposal [85] for the measurement of the CLoE, i.e., to take into account every and any variation in the amount of input that may have occurred in the years prior to data collection, then it is highly likely that, in our experimental group, two children who share the same chronological age differ from each other with regard to their CLoE to the HL. For instance, the sociolinguistic characterisation of the HSs revealed that 16 out of the 23 bilingual children exposed to EP since birth were also exposed to German before the age of three, whereas the remaining seven bilinguals only began receiving systematic input in the societal language between the ages of three and five. If we focus our attention on the variation observed in the amount of input to which each child had been exposed during the first years of life, then one can conclude that, a priori, the CLoE to EP of an eight-year-old child whose age of onset to German was at five is superior to that of an eight-year-old child who, since birth, had to divide his/her time between EP and German. Hence, we suggest that, in future studies, as an alternative to chronological age, one should adopt Unsworth's criteria [85] for the measurement of the bilingual children's language exposure over time.

To conclude, it is important to highlight that, although we believe that the findings of the present study contribute to a better understanding of how lexical acquisition occurs in HSs of EP, the field of bilingual language acquisition would have benefited from a comparative analysis between the HSs' productive and receptive lexical competence. This comparative analysis would have allowed us to verify whether, similar to several studies conducted with bilingual children with different language pairs [6,9,30], the receptive lexical knowledge of the HSs of EP is greater than their productive lexical knowledge. In addition, it would also have allowed us to investigate whether, the productive lexical competence of the bilingual children is more vulnerable to the effects of reduced input than the receptive one, as previous research has found [26,30]. Moreover, in a future study, an analysis on both the receptive and productive lexical knowledge of the bilingual children, in each of their languages, should be carried out. In fact, a contrastive analysis between the minority and the majority language of the Portuguese–German bilingual children would have allowed us to examine whether these children—as other bilingual speakers from other sociolinguistic contexts [8,13,21]—tend to have receptive and productive vocabulary sizes inferior to those of their Portuguese and German monolingual counterparts, when only one of their languages is assessed. Such an analysis would also have allowed us to examine

whether the total and conceptual vocabulary sizes of these bilingual children are similar to those of their monolingual counterparts.

Taking into consideration the results discussed in the current section, one of the most noteworthy findings to emerge from the present study is the fact that, despite being raised in a bilingual environment, more than half of the investigated Portuguese–German bilingual children presented patterns of lexical development in the HL identical to those of monolingual EP children—at least with respect to the lexical knowledge related to the environment in which the acquisition of HLs usually takes place, that is, the home environment. Nevertheless, the fact that the bilingual children belonged to households in which at least one of the parents was a native speaker of EP did not guarantee the acquisition of the Portuguese language within the patterns of language development presented by the monolingual EP-speaking children in the country of origin, even when the picture used in the experimental task was mainly—but not exclusively—related to the domestic context. As could be seen, this was due to the variation in the amount of EP input received and the amount of EP output produced in the verbal interactions between parents and children within the home environment, which makes us highlight the fact that, as primary sources of input in the minority language and as main conversational partners of the Luso-descendant bilingual children, the parents are, in fact, the key players in the acquisition of EP as a heritage language.

**Acknowledgments:** We would like to express our gratitude to the Portuguese emigrant community living in Stuttgart, Germany, for having accepted to participate in this study. We would also like to thank Diogo Oliveira, the bilingual children’s teacher of Portuguese as a Heritage Language, for all the support during data collection.

**Author Contributions:** L.C. carried out the whole empirical study, described the method and participants and wrote the results section, while C.F. contributed to the discussion, literature review and reviewed the text.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Sociolinguistic data of the bilingual speakers.

Participant	Age	Gender	EP-Speaking Parents	Older Siblings	Amount of Input (Parents)	Amount of Output (Parents)	Quantity of EP-Speaking Conversational Parents	Quantity of Other Input Sources
HS_1	11	M	both	No	100	100	0	3
HS_2	10	F	both	Yes	200	200	0	2
HS_3	9	F	both	Yes	200	200	0	1
HS_4	9	M	father	No	100	50	2	3
HS_5	11	F	both	No	200	200	2	3
HS_6	7	M	both	Yes	150	150	3	2
HS_7	9	M	both	No	150	100	3	2
HS_8	9	M	both	Yes	150	150	2	3
HS_9	8	M	both	No	125	50	1	5
HS_10	9	M	both	No	175	75	2	3
HS_11	9	F	both	No	175	75	2	3
HS_12	9	F	both	No	150	50	2	2
HS_13	7	F	both	Yes	200	200	3	3
HS_14	7	F	both	No	200	200	0	5
HS_15	8	M	both	Yes	150	50	2	2
HS_16	8	F	both	No	200	150	0	3
HS_17	8	F	both	Yes	200	150	0	2
HS_18	6	F	both	Yes	200	150	0	3
HS_19*	11	F	mother	No	75	50	0	2
HS_20	8	M	mother	No	25	25	2	0
HS_21	7	M	both	Yes	100	50	4	2
HS_22	6	M	mother	Yes	75	50	2	0
HS_23	9	M	both	No	100	50	4	2

Note: M = Male; F = Female; HS = Heritage Speaker; \* Due to being an outlier, participant HS\_19 was excluded from the correlation analyses between the input factors and the HSs’ lexical knowledge.

## Appendix B



Figure A1. Thematic picture *Sala* ‘Living Room’.

## Appendix C

Table A2. Pairwise comparisons between the grammatical categories (nouns vs. verbs/adjectives and verbs vs. adjectives).

Group	Grammatical Category	Grammatical Categories	Mean Difference	Significance
EG	Nouns	Verbs	14,522	***
		Adjectives	17,957	***
	Verbs	Adjectives	3435	***
CG	Nouns	Verbs	24,810	***
		Adjectives	31,381	***
	Verbs	Adjectives	6571	***

Note: EG = Experimental Group; CG = Control Group; Significance level: \*\*\*  $p < 0.001$ .

## Appendix D

Table A3. Nouns, verbs and adjectives with the 10 highest frequencies per group of speakers.

Rank	Nouns	Quantity of HSs (AF and RF)	Rank	Nouns	Quantity of MSs (AF and RF)
1	<i>bola</i> ‘ball’	23 (100%)	1	<i>almofada</i> ‘cushion’	21 (100%)
2	<i>televisão</i> ‘television’	22 (95.7%)		<i>bruxa</i> ‘witch’	21 (100%)
3	<i>dinheiro</i> ‘money’	21 (91.3%)		<i>caixa</i> ‘box’	21 (100%)
	<i>jogo</i> ‘game’	21 (91.3%)		<i>desenho</i> ‘drawing’	21 (100%)
4	<i>livro</i> ‘book’	20 (87%)		<i>dinheiro</i> ‘money’	21 (100%)
	<i>sofá</i> ‘sofa’	20 (87%)		<i>jornal</i> ‘newspaper’	21 (100%)
5	<i>bruxa</i> ‘witch’	19 (82.6%)		<i>lápiz</i> ‘pencil’	21 (100%)
6	<i>jornal</i> ‘newspaper’	18 (78.3%)		<i>rádio</i> ‘radio’	21 (100%)
	<i>mesa</i> ‘table’	18 (78.3%)		<i>televisão</i> ‘television’	21 (100%)
	<i>rádio</i> ‘radio’	18 (78.3%)		<i>tesoura</i> ‘scissors’	21 (100%)
	<i>tesoura</i> ‘scissors’	18 (78.3%)	2	<i>igreja</i> ‘church’	20 (95.2%)
7	<i>gravata</i> ‘tie’	17 (73.9%)		<i>quadro</i> ‘painting’	20 (95.2%)
	<i>lápiz</i> ‘pencil’	17 (73.9%)	<i>vassoura</i> ‘broom’	20 (95.2%)	
8	<i>chapéu</i> ‘hat’	16 (69.6%)	3	<i>jogo</i> ‘game’	18 (85.7%)
9	<i>caixa</i> ‘box’	15 (65.2%)		<i>livro</i> ‘book’	18 (85.7%)
10	<i>planta</i> ‘plant’	14 (60.9%)		<i>martelo</i> ‘hammer’	18 (85.7%)
	<i>sol</i> ‘sun’	14 (60.9%)		<i>pai</i> ‘father’	18 (85.7%)
				<i>palhaço</i> ‘clown’	18 (85.7%)
				<i>prateleira</i> ‘shelf’	18 (85.7%)

Table A3. Cont.

Rank	Nouns	Quantity of HSs (AF and RF)	Rank	Nouns	Quantity of MSs (AF and RF)
			4	<i>sofá</i> 'sofa'	17 (81%)
			5	<i>bola</i> 'ball'	16 (76.2%)
			6	<i>gato</i> 'cat'	15 (71.4%)
			7	<i>cofre</i> 'safe' <i>prego</i> 'nail'	14 (66.7%) 14 (66.7%)
			8	<i>mesa</i> 'table'	13 (61.9%)
			9	<i>planta</i> 'plant'	12 (57.1%)
			10	<i>casa</i> 'house/home' <i>coisa</i> 'thing' <i>cor</i> 'colour' <i>saco</i> 'bag'	11 (52.4%) 11 (52.4%) 11 (52.4%) 11 (52.4%)
Rank	Verbs	Quantity of HSs (AF and RF)	Rank	Verbs	Quantity of MSs (AF and RF)
1	<i>ver</i> 'to see/watch'	20 (87.0%)	1	<i>estar</i> 'to be'	20 (95.2%)
2	<i>estar</i> 'to be'	19 (82.6%)	2	<i>ser</i> 'to be'	19 (90.5%)
3	<i>voar</i> 'to fly'	18 (78.3%)	3	<i>cantar</i> 'to sing'	17 (81.0%)
4	<i>brincar</i> 'to play' <i>saber</i> 'to know' <i>ser</i> 'to be'	17 (73.9%) 17 (73.9%) 17 (73.9%)	4	<i>ter</i> 'to have' <i>ver</i> 'to see/watch'	16 (76.2%) 16 (76.2%)
5	<i>cantar</i> 'to sing'	14 (60.9%)	5	<i>poder</i> 'can' <i>voar</i> 'to fly'	14 (66.7%) 14 (66.7%)
6	<i>poder</i> 'can'	13 (56.5%)	6	<i>ir</i> 'to go'	13 (61.9%)
7	<i>arrumar</i> 'to tidy' <i>jogar</i> 'to play' <i>pintar</i> 'to colour'	11 (47.8%) 11 (47.8%) 11 (47.8%)	7	<i>arrumar</i> 'to tidy'	12 (57.1%)
8	<i>chamar</i> 'to call' <i>ter</i> 'to have'	10 (43.5%) 10 (43.5%)	8	<i>pôr</i> 'to put' <i>saber</i> 'to know'	11 (52.4%) 11 (52.4%)
9	<i>ir</i> 'to go'	8 (34.8%)	9	<i>achar</i> 'to think' <i>dever</i> 'should' <i>dizer</i> 'to say'	9 (42.9%) 9 (42.9%) 9 (42.9%)
10	<i>falar</i> 'to speak' <i>fazer</i> 'to do/make' <i>gostar</i> 'to like'	7 (30.4%) 7 (30.4%) 7 (30.4%)	10	<i>fazer</i> 'to do/make' <i>costumar</i> 'to be used to' <i>*ficar</i> 'to be/stay/become'	9 (42.9%) 8 (38.1%) 8 (38.1%)
Rank	Adjectives	Quantity of HSs (AF and RF)	Rank	Adjectives	Quantity of MSs (AF and RF)
1	<i>amarelo</i> 'yellow'	19 (82.6%)	1	<i>amarelo</i> 'yellow'	17 (81.0%)
2	<i>azul</i> 'blue' <i>vermelho</i> 'red'	18 (78.3%) 18 (78.3%)	2	<i>verde</i> 'green' <i>vermelho</i> 'red'	17 (81.0%) 16 (76.2%)
3	<i>verde</i> 'green'	17 (73.9%)	3	<i>azul</i> 'blue' <i>branco</i> 'white' <i>desarrumado</i> 'untidy'	15 (71.4%) 15 (71.4%) 15 (71.4%)
4	<i>fechado</i> 'closed'	16 (69.6%)	4	<i>preto</i> 'black'	14 (66.7%)
5	<i>preto</i> 'black'	15 (65.2%)	5	<i>castanho</i> 'brown'	13 (61.9%)
6	<i>aberto</i> 'open'	14 (60.9%)	6	<i>cinzento</i> 'grey'	12 (57.1%)
7	<i>branco</i> 'white'	13 (56.5%)	7	<i>arrumado</i> 'tidy'	11 (52.4%)
8	<i>desarrumado</i> 'untidy' <i>cinzento</i> 'grey'	10 (43.5%) 8 (34.8%)	8	<i>cor-de-rosa</i> 'pink' <i>laranja</i> 'orange'	8 (38.1%) 8 (38.1%)
9	<i>cor de laranja</i> 'orange' <i>cor-de-rosa</i> 'pink' <i>*lilás</i> 'lilac'	8 (34.8%) 8 (34.8%) 8 (34.8%)	9	<i>cor de laranja</i> 'orange' <i>roxo</i> 'purple' <i>verde-escuro</i> 'dark green'	7 (33.3%) 7 (33.3%) 7 (33.3%)
10	<i>castanho</i> 'brown' <i>*português</i> 'portuguese' <i>rosa</i> 'pink'	6 (26.1%) 6 (26.1%) 6 (26.1%)	10	<i>*escuro</i> 'dark' <i>*estranho</i> 'strange' <i>verde-claro</i> 'light green'	5 (23.8%) 5 (23.8%) 5 (23.8%)

Note: HS = Heritage Speaker; MS = Monolingual Speaker; AF = Absolute Frequency; RF = Relative Frequency; \* Lexical items which were produced by only one group of speakers.

## References

1. Flores, C. Understanding heritage language acquisition. Some contributions from the research on heritage speakers of European Portuguese. *Lingua* **2015**, *164*, 251–265. [[CrossRef](#)]
2. Gathercole, V.C.M. Bilingualism matters: One size does not fit all. *Int. J. Behav. Dev.* **2014**, *38*, 359–366. [[CrossRef](#)]
3. Montrul, S. *The Acquisition of Heritage Languages*; Cambridge University Press: Cambridge, UK, 2016.
4. Conboy, B.T.; Thal, D.J. Ties between the Lexicon and Grammar: Cross-Sectional and Longitudinal Studies of Bilingual Toddlers. *Child Dev.* **2006**, *77*, 712–735. [[CrossRef](#)] [[PubMed](#)]
5. David, A.; Li, W. The composition of the bilingual lexicon. In Proceedings of the 4th International Symposium on Bilingualism, Arizona State University, Tempe, AZ, USA, 30 April–3 May 2003; Cohen, J., McAlister, K.T., Rolstad, K., MacSwan, J., Eds.; Cascadilla Press: Somerville, MA, USA, 2005; pp. 594–607.
6. De Houwer, A.; Bornstein, M.H.; Putnick, D.L. A bilingual–monolingual comparison of young children’s vocabulary size: Evidence from comprehension and production. *Appl. Psycholinguist.* **2014**, *35*, 1189–1211. [[CrossRef](#)]
7. Holowka, S.; Brosseau-Lapr e, F.; Petitto, L.A. Semantic and conceptual knowledge underlying bilingual babies’ first signs and words. *Lang. Learn.* **2002**, *52*, 205–262. [[CrossRef](#)]
8. Junker, D.A.; Stockman, I.J. Expressive vocabulary of German-English bilingual toddlers. *Am. J. Speech Lang. Pathol.* **2002**, *11*, 381–394. [[CrossRef](#)]
9. Pearson, B.Z.; Fern andez, S.C.; Oller, D.K. Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Lang. Learn.* **1993**, *43*, 93–120. [[CrossRef](#)]
10. Pearson, B.Z.; Fern andez, S.C. Patterns of Interaction in the Lexical Growth in Two Languages of Bilingual Infants and Toddlers. *Lang. Learn.* **1994**, *44*, 617–653. [[CrossRef](#)]
11. Patterson, J. Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *Am. J. Speech Lang. Pathol.* **1998**, *7*, 46–56. [[CrossRef](#)]
12. Petitto, L.A.; Katerelos, M.; Levy, B.G.; Gauna, K.; T etreault, K.; Ferraro, V. Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *J. Child Lang.* **2001**, *28*, 453–496. [[CrossRef](#)] [[PubMed](#)]
13. Poulin-Dubois, D.; Bialystok, E.; Blaye, A.; Polonia, A.; Yott, J. Lexical access and vocabulary development in very young bilinguals. *Int. J. Biling.* **2013**, *17*, 57–70. [[CrossRef](#)] [[PubMed](#)]
14. Dos Santos, C.; Kern, S. Early Lexical Development of French-Portuguese Bilingual Children: A CDI-Adaptation Study. In Presented at the Child Language Impairment in Multilingual Contexts, Krakow, Poland, May 2013; Available online: [http://www.ddl.ish-lyon.cnrs.fr/fulltext/santos/dos%20Santos\\_2013\\_CLIMC.pdf](http://www.ddl.ish-lyon.cnrs.fr/fulltext/santos/dos%20Santos_2013_CLIMC.pdf) (accessed on 25 April 2017).
15. Cote, L.R.; Bornstein, M.H. Productive vocabulary among three groups of bilingual American children: Comparison and prediction. *First Lang.* **2014**, *34*, 467–485. [[CrossRef](#)] [[PubMed](#)]
16. Bridges, K.; Hoff, E. Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Appl. Psycholinguist.* **2014**, *35*, 225–241. [[CrossRef](#)] [[PubMed](#)]
17. Cobo-Lewis, A.; Pearson, B.; Eilers, R.; Umbel, V. Effects of bilingualism and bilingual education on oral and written English skills: A multifactor study of standardized test outcomes. In *Language and Literacy in Bilingual Children*; Oller, D.K., Eilers, R., Eds.; Multilingual Matters: Clevedon, UK, 2002; pp. 64–97.
18. Cobo-Lewis, A.; Pearson, B.; Eilers, R.; Umbel, V. Effects of bilingualism and bilingual education on oral and written Spanish skills: A multifactor study of standardized test outcomes. In *Language and Literacy in Bilingual Children*; Oller, D.K., Eilers, R., Eds.; Multilingual Matters: Clevedon, UK, 2002; pp. 98–117.
19. David, A.; Li, W. Individual differences in the lexical development of French–English bilingual children. *Int. J. Biling. Educ. Biling.* **2008**, *11*, 598–618. [[CrossRef](#)]
20. Gathercole, V.C.M.; Thomas, E.M. Bilingual first-language development: Dominant language takeover, threatened minority language take-up. *Biling. Lang. Cogn.* **2009**, *12*, 213–237. [[CrossRef](#)]
21. Hoff, E.; Core, C.; Place, S.; Rumiche, R.; Se nor, M.; Parra, M. Dual language exposure and early bilingual development. *J. Child Lang.* **2012**, *39*, 1–27. [[CrossRef](#)] [[PubMed](#)]
22. Hoff, E.; Rumiche, R.; Burrige, A.; Ribot, K.M.; Welsh, S.N. Expressive vocabulary development in children from bilingual and monolingual homes: A longitudinal study from two to four years. *Early Child. Res. Q.* **2014**, *29*, 433–444. [[CrossRef](#)] [[PubMed](#)]

23. Pearson, B.Z.; Fernández, S.C.; Lewedeg, V.; Oller, D.K. The relation of input factors to lexical learning by bilingual infants. *Appl. Psycholinguist.* **1997**, *18*, 41–58. [[CrossRef](#)]
24. Place, S.; Hoff, E. Properties of dual language exposure that influence two-year-olds' bilingual proficiency. *Child Dev.* **2011**, *82*, 1834–1849. [[CrossRef](#)] [[PubMed](#)]
25. Smithson, L.; Paradis, J.; Nicoladis, E. Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference? *Biling. Lang. Cogn.* **2014**, *17*, 810–821. [[CrossRef](#)]
26. Thordardottir, E. The relationship between bilingual exposure and vocabulary development. *Int. J. Biling.* **2011**, *15*, 426–445. [[CrossRef](#)]
27. De Houwer, A. *Bilingual First Language Acquisition*; Multilingual Matters: Clevedon, UK, 2009.
28. Kan, P.F.; Kohnert, K. Preschoolers learning Hmong and English: Lexical semantic skills in L1 and L2. *J. Speech Lang. Hear. Res.* **2005**, *48*, 372–383. [[CrossRef](#)]
29. Özcan, F.H.; Altinkamiş, F.N.; Gillis, S. Early lexical composition of Turkish-Dutch bilinguals: Nouns before verbs or verbs before nouns. *Pozn. Stud. Contemp. Linguist.* **2016**, *52*, 583–604. [[CrossRef](#)]
30. Sheng, L.; Lu, Y.; Kan, P.F. Lexical development in Mandarin–English bilingual children. *Biling. Lang. Cogn.* **2011**, *14*, 579–587. [[CrossRef](#)]
31. Conboy, B.T.; Montanari, S. Early Lexical Development in Bilingual Infants and Toddlers. In *Bilingualism across the Lifespan*; Nicoladis, E., Montanari, S., Eds.; De Gruyter: Berlin, Germany, 2016; pp. 63–80.
32. Bialystok, E.; Craik, F.I.; Green, D.W.; Gollan, T.H. Bilingual minds. *Psychol. Sci. Public Interest* **2009**, *10*, 89–129. [[CrossRef](#)] [[PubMed](#)]
33. Pearson, B.Z. Children with two languages. In *The Cambridge Handbook of Child Language*; Bavin, E., Ed.; Cambridge University Press: Cambridge, UK, 2009; pp. 379–397.
34. Bialystok, E.; Luk, G.; Peets, K.F.; Yang, S. Receptive vocabulary differences in monolingual and bilingual children. *Biling. Lang. Cogn.* **2010**, *13*, 525–531. [[CrossRef](#)] [[PubMed](#)]
35. Oller, D.K.; Pearson, B.Z.; Cobo-Lewis, A.B. Profile effects in early bilingual language and literacy. *Appl. Psycholinguist.* **2007**, *28*, 191–230. [[CrossRef](#)] [[PubMed](#)]
36. Oller, D.K. The distributed characteristic in bilingual learning. In *ISB4: Proceedings of the 4th International Symposium on Bilingualism*; Cohen, J., McAlister, K.T., Rolstad, K., MacSwan, J., Eds.; Cascadilla Press: Somerville, MA, USA, 2005; pp. 1744–1749.
37. Oller, D.K.; Pearson, B.Z. Assessing the effects of bilingualism: A background. In *Language and Literacy in Bilingual Children*; Oller, D.K., Eilers, R.E., Eds.; Multilingual Matters: Clevedon, UK, 2002; pp. 3–21.
38. Hoff, E. How social contexts support and shape language development. *Dev. Rev.* **2006**, *26*, 55–88. [[CrossRef](#)]
39. Hoff, E.; Core, C. Input and language development in bilingually developing children. *Semin. Speech Lang.* **2013**, *34*, 215–226. [[CrossRef](#)] [[PubMed](#)]
40. Ribot, K.M.; Hoff, E.; Burrige, A. Language Use Contributes to Expressive Language Growth: Evidence from Bilingual Children. *Child Dev.* **2017**, 1–12. [[CrossRef](#)] [[PubMed](#)]
41. Patterson, J. Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Appl. Psycholinguist.* **2002**, *23*, 493–508. [[CrossRef](#)]
42. Scheele, A.; Leseman, P.; Mayo, A. The home language environment of monolingual and bilingual children and their language proficiency. *Appl. Psycholinguist.* **2010**, *31*, 117–140. [[CrossRef](#)]
43. Montrul, S. *Incomplete Acquisition in Bilingualism. Re-Examining the Age Factor*; John Benjamins: Amsterdam, The Netherlands, 2008.
44. Polinsky, M.; Kagan, O. Heritage Languages: In the 'Wild' and in the Classroom. *Lang. Linguist. Compass* **2007**, *1*, 368–395. [[CrossRef](#)]
45. Polinsky, M. Incomplete acquisition: American Russian. *J. Slav. Linguist.* **2006**, *14*, 191–262.
46. Snedeker, J. Word learning. In *Encyclopedia of Neuroscience*; Squire, L.R., Ed.; Elsevier: Amsterdam, The Netherlands, 2009; pp. 503–508.
47. Cresti, E.; Moneglia, M. *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*; John Benjamins: Amsterdam, The Netherlands, 2005.
48. Biber, D.; Johansson, S.; Leech, G.; Conrad, S.; Finegan, E. *The Longman Grammar of Spoken and Written English*; Longman: London, UK, 1999.
49. Au, T.K.; Dapretto, M.; Song, Y.K. Input vs. constraints: Early word acquisition in Korean and English. *J. Mem. Lang.* **1994**, *33*, 567–582. [[CrossRef](#)]

50. Bornstein, M.H.; Cote, L.R.; Maital, S.; Painter, K.; Park, S.-Y.; Pascual, L.; Pêcheux, M.-G.; Ruel, J.; Venuti, P.; Vyt, A. Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Dev.* **2004**, *75*, 1115–1139. [[CrossRef](#)] [[PubMed](#)]
51. Choi, S.; Gopnik, A. Early acquisition of verbs in Korean: A cross-linguistic study. *J. Child Lang.* **1995**, *22*, 497–529. [[CrossRef](#)] [[PubMed](#)]
52. Kim, M.; McGregor, K.; Thompson, C. Early lexical development in English- and Korean-speaking children: Language-general and language-specific patterns. *J. Child Lang.* **2000**, *27*, 224–254. [[CrossRef](#)]
53. Gelman, S.A.; Tardif, T. Acquisition of nouns and verbs in Mandarin and English. In Proceedings of the Twenty-Ninth Annual Child Language Research Forum, Cordura Hall, Center for the Study of Language and Information (CSLI), Stanford University, Stanford, CA, USA, 25–27 April 1997; Clark, E., Ed.; CSLI: Stanford, CA, USA, 1998; pp. 27–36.
54. Tardif, T. Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Dev. Psychol.* **1996**, *32*, 492–504. [[CrossRef](#)]
55. Tardif, T.; Shatz, M.; Naigles, L. Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *J. Child Lang.* **1997**, *24*, 535–565. [[CrossRef](#)] [[PubMed](#)]
56. Bates, E.; Marchman, V.; Thal, D.; Fenson, L.; Dale, P.; Reznick, J.S.; Reilly, J.; Hartung, J. Developmental and stylistic variation in the composition of early vocabulary. *J. Child Lang.* **1994**, *21*, 85–124. [[CrossRef](#)] [[PubMed](#)]
57. Nelson, K. Structure and Strategy in Learning to Talk. *Monogr. Soc. Res. Child Dev.* **1973**, *38*, 1–135. [[CrossRef](#)]
58. Caselli, M.C.; Bates, E.; Casadio, P.; Fenson, J.; Fenson, L.; Sanderl, L.; Weir, J. A cross-linguistic study of early lexical development. *Cogn. Dev.* **1995**, *10*, 159–199. [[CrossRef](#)]
59. D'Odorico, L.; Fasolo, M. Nouns and verbs in the vocabulary acquisition of Italian children. *J. Child Lang.* **2007**, *34*, 891–907. [[CrossRef](#)] [[PubMed](#)]
60. Jackson-Maldonado, D.; Thal, D.; Marchman, V.; Bates, E.; Gutiérrez-Clellen, V. Early lexical development in Spanish-speaking infants and toddlers. *J. Child Lang.* **1993**, *20*, 523–549. [[CrossRef](#)] [[PubMed](#)]
61. Maital, S.; Dromi, E.; Sagi, A.; Bornstein, M.H. The Hebrew Communicative Development Inventory: Language specific properties and cross-linguistic generalizations. *J. Child Lang.* **2000**, *27*, 43–67. [[CrossRef](#)] [[PubMed](#)]
62. Bassano, D. Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *J. Child Lang.* **2000**, *27*, 521–559. [[CrossRef](#)] [[PubMed](#)]
63. Schlichting, L. *Discovering Syntax: An Empirical Study in Dutch Language Acquisition*; Nijmegen University Press: Nijmegen, The Netherlands, 1996.
64. Gentner, D.; Boroditsky, L. Early acquisition of nouns and verbs: Evidence from Navajo. In *Routes to Language: Studies in Honor of Melissa Bowerman*; Gathercole, V.C.M., Ed.; Psychology Press: New York, NY, USA, 2009; pp. 5–36.
65. Scherer, S.; Souza, A.P.R. Types e tokens na aquisição típica de linguagem por sujeitos de 18 a 32 meses falantes do português brasileiro. *Rev. CEFAC* **2011**, *13*, 838–845. [[CrossRef](#)]
66. Gentner, D. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In *Language Development: Language, Thought and Culture*; Kuczaj, S.A., Ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1982; Volume 2, pp. 301–334.
67. Barnes, J.; García, I. Vocabulary growth and composition in monolingual and bilingual Basque infants and toddlers. *Int. J. Biling.* **2013**, *17*, 357–374. [[CrossRef](#)]
68. Lucas, R.I.; Bernardo, A.B. Exploring Noun Bias in Filipino-English Bilingual Children. *J. Genet. Psychol.* **2008**, *169*, 149–164. [[CrossRef](#)] [[PubMed](#)]
69. Nicoladis, E. Finding first words in the input. In *Trends in Bilingual Acquisition*; Cenoz, J., Genesee, F., Eds.; John Benjamins: Amsterdam, The Netherlands, 2001; pp. 131–147.
70. Xuan, L.; Dollaghan, C. Language-specific noun bias: Evidence from bilingual children. *J. Child Lang.* **2013**, *40*, 1057–1075. [[CrossRef](#)] [[PubMed](#)]
71. Hart, B.; Risley, T.R. *Meaningful Differences in the Everyday Experience of Young American Children*; Brookes Publishing: Baltimore, MD, USA, 1995.
72. De Houwer, A. Language input environments and language development in bilingual acquisition. *Appl. Linguist. Rev.* **2011**, *2*, 221–240. [[CrossRef](#)]
73. Barron-Hauwaert, S. *Bilingual Siblings*; Multilingual Matters: Clevedon, UK, 2011.

74. Kopeliovich, S. Happylingual: A family project for enhancing and balancing multilingual development. In *Successful Family Language Policy: Parents, Children and Educators in Interaction*; Schwartz, M., Verschik, A., Eds.; Springer: Dordrecht, The Netherlands, 2013; pp. 249–275.
75. Schwartz, M. Family language policy: Core issues of an emerging field. *Appl. Linguist. Rev.* **2010**, *1*, 171–192. [[CrossRef](#)]
76. Spolsky, B. *Language Management*; Cambridge University Press: Cambridge, UK, 2009.
77. Bedore, L.M.; Peña, E.D.; Summers, C.L.; Boerger, K.M.; Resendiz, M.D.; Greene, K.; Bohman, T.M.; Gillam, R.B. The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Biling. Lang. Cogn.* **2012**, *15*, 616–629. [[CrossRef](#)] [[PubMed](#)]
78. Bohman, T.M.; Bedore, L.M.; Peña, E.D.; Mendez-Perez, A.; Gillam, R.B. What you hear and what you say: Language performance in Spanish-English bilinguals. *Int. J. Biling. Educ. Biling.* **2010**, *13*, 325–344. [[CrossRef](#)] [[PubMed](#)]
79. Cohen, C. Relating input factors and dual language proficiency in French-English bilingual children. *Int. J. Biling. Educ. Biling.* **2016**, *19*, 296–313. [[CrossRef](#)]
80. Hammer, C.S.; Komaroff, E.; Rodriguez, B.L.; Lopez, L.M.; Scarpino, S.E.; Goldstein, B. Predicting Spanish-English bilingual children's abilities. *J. Speech Lang. Hear. Res.* **2012**, *55*, 1251–1264. [[CrossRef](#)]
81. Unsworth, S. Quantity-oriented and quality-oriented exposure variables in simultaneous bilingual acquisition. In *Papers of the Anéla 2012 Applied Linguistics Conference*; Jong, N., Juffermans, K., Keijzer, M., Rasier, L., Eds.; Eburon: Delft, The Netherlands, 2012; pp. 13–22.
82. Pearson, B.Z. Social factors in childhood bilingualism in the United States. *Appl. Psycholinguist.* **2007**, *28*, 399–410. [[CrossRef](#)]
83. Gutiérrez-Clellen, V.; Kreiter, J. Understanding child bilingual acquisition using parent and teacher reports. *Appl. Psycholinguist.* **2003**, *24*, 267–288. [[CrossRef](#)]
84. Maas, E.L.H. The Role of Bilingual Language Input in Children's Receptive and Productive Vocabulary Development. Master's Thesis, Utrecht University, Utrecht, The Netherlands, 2014.
85. Unsworth, S. Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Biling. Lang. Cogn.* **2013**, *16*, 86–110. [[CrossRef](#)]
86. Unsworth, S.; Argyri, F.; Cornips, L.; Hulk, A.; Sorace, A.; Tsimpli, I. The role of age of onset and input in early child bilingualism in Greek and Dutch. *Appl. Psycholinguist.* **2014**, *35*, 765–805. [[CrossRef](#)]
87. Unsworth, S. Utrecht Bilingual Language Exposure Calculator. Unpublished Manuscript. 2011.
88. Cornips, L.; Hulk, A. Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second Lang. Res.* **2008**, *24*, 267–296. [[CrossRef](#)]
89. Driessen, G.; van der Slik, F.; De Bot, K. Home language and language proficiency: A large-scale longitudinal study in Dutch primary schools. *J. Multiling. Multicult. Dev.* **2002**, *23*, 175–194. [[CrossRef](#)]
90. Cha, K.; Goldenberg, C. The complex relationship between bilingual home language input and kindergarten children's Spanish and English oral proficiencies. *J. Educ. Psychol.* **2015**, *107*, 935–953. [[CrossRef](#)]
91. Jia, G.; Aaronson, D. A longitudinal study of Chinese children and adolescents learning English in the United States. *Appl. Psycholinguist.* **2003**, *24*, 131–161. [[CrossRef](#)]
92. Jia, G.; Fuse, A. Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *J. Speech Lang. Hear. Res.* **2007**, *50*, 1280–1299. [[CrossRef](#)]
93. Leseman, P.; Mayo, A.; Scheele, A. Old and new media in the lives of young disadvantaged bilingual children. In *Multimedia and Literacy Development*; Bus, A.G., Neuman, S.B., Eds.; Routledge: New York, NY, USA, 2009; pp. 135–155.
94. Ishizawa, H. Minority language use among grandchildren in multigenerational households. *Sociol. Perspect.* **2004**, *47*, 465–483. [[CrossRef](#)]
95. Alves, I.; Costa, P.; Lourenço-Gomes, M.C.; Rodrigues, C. EFFE-On—Corpus Online de Escrita e Fala. *Saber Educar* **2015**, *20*, 24–33. [[CrossRef](#)]
96. Lourenço-Gomes, M.C.; Rodrigues, C.; Alves, I. EFFE-Escreves como falas—Falas como escreves? *Rev. Romane* **2016**, *51*, 36–69. [[CrossRef](#)]
97. Yavas, M.; Hernandorena, C.L.M.; Lamprecht, R.R. *Avaliação Fonológica da Criança*; Artes Médicas: Porto Alegre, Brazil, 1991.
98. Guerreiro, H. Processos Fonológicos na Fala da Criança de Cinco anos. Master's Thesis, Instituto de Ciências da Saúde da Universidade Católica Portuguesa, Lisbon, Portugal, 2007.

99. Field, A.P. *Discovering Statistics Using SPSS: (And Sex and Drugs and Rock 'n' Roll)*, 3rd ed.; SAGE Publications: Los Angeles, CA, USA; London, UK, 2009.
100. Bacelar do Nascimento, M.F.; Gonçalves, J.B.; Veloso, R.; Antunes, S.; Barreto, F.; Amaro, R. The Portuguese corpus. In *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*; Cresti, E., Moneglia, M., Eds.; John Benjamins: Amsterdam, The Netherlands, 2005; pp. 163–207.
101. Unsworth, S. Quantity and Quality of Language Input in Bilingual Language Development. In *Bilingualism across the Lifespan*; Nicoladis, E., Montanari, S., Eds.; De Gruyter: Berlin, Germany, 2016; pp. 103–122.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).