Original Article

Artigo Original

Fernanda Marafiga Wiethan[1]
Helena Bolli Mota[1]
Anaelena Bragança de Moraes[1]

Correlations between vocabulary and phonological acquisition: number of words produced versus acquired consonants

Correlações entre aquisição do vocabulário e da fonologia: número de palavras produzidas versus consoantes adquiridas

ABSTRACT

Purpose: To verify the probable correlations between the number of word types and the number of consonants in the general phonological system in children with typical language development. **Methods**: Study participants were 186 children aged one year and six months to five years, 11 months and 29 days who were monolingual Brazilian Portuguese speakers with typical language development. Data collection involved speech, language and hearing assessments and spontaneous speech recordings. Phonology was assessed with regard to the number of acquired consonants in the general phonological system, in each syllable structure and in Implicational Model of Feature Complexity (IMFC) levels. Vocabulary was assessed with regard to number of word types produced. These data were compared across age groups. After that, correlations between the word types produced and the variables established for the phonological system were analyzed. The significance level adopted was 5%. **Results:** All phonological aspects evaluated presented gradual growth. Word types produced showed a similar behavior, though with a small regression at the age of five years. Different positive correlations occurred between the spoken word types and the variables analyzed in the phonological system. Only one negative correlation occurred with respect to the production of complex onset in the last age group analyzed. **Conclusion:** The phonology and vocabulary of the study participants present similar behaviors. There are many positive correlations between the word types produced and the different aspects of phonology, except regarding complex onset.

RESUMO

Objetivo: Verificar as possíveis correlações entre o número de tipos lexicais e o número de consoantes no sistema fonológico geral em crianças com desenvolvimento típico de linguagem. **Métodos:** 186 crianças entre um ano e seis meses e cinco anos, 11 meses e 29 dias, falantes monolíngues do Português Brasileiro, com desenvolvimento típico de linguagem, participaram do estudo. A coleta de dados envolveu avaliações fonoaudiológicas e gravação da fala espontânea. Analisou-se a fonologia quanto ao número de consoantes adquiridas no sistema fonológico geral, em cada estrutura silábica e nos níveis do Modelo Implicacional de Complexidade de Traços (MICT). O vocabulário foi analisado quanto ao número de tipos de itens lexicais produzidos. Estes dados foram comparados entre as faixas etárias. A seguir, foram analisadas as correlações entre os tipos lexicais produzidos e as variáveis estabelecidas para o sistema fonológico. O nível de significância adotado foi de 5%. **Resultados:** Todos os aspectos da fonologia avaliados apresentaram crescimento gradual. Os tipos lexicais apresentaram comportamento semelhante, porém com uma pequena regressão na idade de cinco anos. Houve diferentes correlações positivas entre os tipos lexicais produzidos e as variáveis analisadas no sistema fonológico. Houve apenas uma correlação negativa, referente à produção dos tipos lexicais com o *onset* complexo na última faixa etária analisada. **Conclusão:** A fonologia e o vocabulário apresentam comportamentos semelhantes. Há diversas correlações positivas entre os tipos lexicais produzidos e diferentes aspectos da fonologia, exceto em relação ao *onset* complexo.

Correspondence address:
Fernanda Marafiga Wiethan
Universidade Federal de Santa Maria – UFSM
Rua Júlio Nogueira, 130, Bairro Uglione, Santa Maria (RS), Brazil, CEP: 97070-510.
E-mail: fernanda_wiethan@yahoo.com.br

INTRODUCTION

In the complex and intriguing study of oral language, there are several parameters that can be considered in order to evaluate and analyze development. Therefore, researchers need to be insightful and creative to use different theories in an attempt to better explain acquisition, evolution and use, as well as differences and "flaws" when these are found. Consequently, in an attempt to circumscribe the object of study, the theme language is divided into systems: phonology, vocabulary/lexicon/semantics, syntax, morphology and pragmatics. This article will investigate typical acquisition of phonology and vocabulary of Brazilian Portuguese and relationships that occur between these systems in the course of acquisition.

Much is already known about the acquisition of Brazilian Portuguese consonants, especially in the dialectal variant of Rio Grande do Sul State, which can be exemplified by the studies conducted by researcher Regina Lamprecht and compiled in book form[1]. As for the ages of acquisition of the phonemes in the simple onset position, it has been found that at one year and six months of age, the child produces the phonemes /p, b, t, d, m, n/; at one year and nine months most of the back nasals and stops and the fricatives /f/ and /v/ are stabilized; at two years and two months the affricates and /s/ and /z/ appear; at two years and eight months the phonemes /ʒ/ and /l/ appear; at two years and ten months the child has mastered /ʃ/; at three years and four months /R/ has been mastered; at three years and ten months /ʎ/ appears; and at four years /r/ appears. With regard to syllable structures, simple onset is the earliest structure, followed by coda with /S/, acquired at two years and six months and coda with /R/ acquired at three years and ten months. Complex onset only appears at the age of five, being considered, therefore, as late-acquisition[1].

Another author[2], also from Rio Grande do Sul, aiming to describe the phonological acquisition of Brazilian Portuguese in children with phonological deviation, created the Implicational Model of Feature Complexity (IMFC), which explains the acquisition of segmental consonants through implicational relationships between features, as illustrated in Figure 1.

The features that appear in the model are marked features, i.e. those with greater complexity, and the lines represent relationships that exist between these features, which can be strong or weak implicational relationship. The root corresponds to features that have zero complexity (/p/, /t/, /m/ and /n/ phonemes), which make up the basic representational structure given in universal grammar and present only unmarked features. From the root, or zero state, emerges a tree structure, in which the branches represent conditions of markedness and where the farther the branch is from the root, the more complex these conditions are. If in a same path, there are two or more features or combinations of features, this indicates that there is an implicational relationship between them. If one feature or combination of features is the target of two or more converging paths, this means that in order for one of these features to be specified, it is necessary for the set of features corresponding to these paths to have been specified[2].
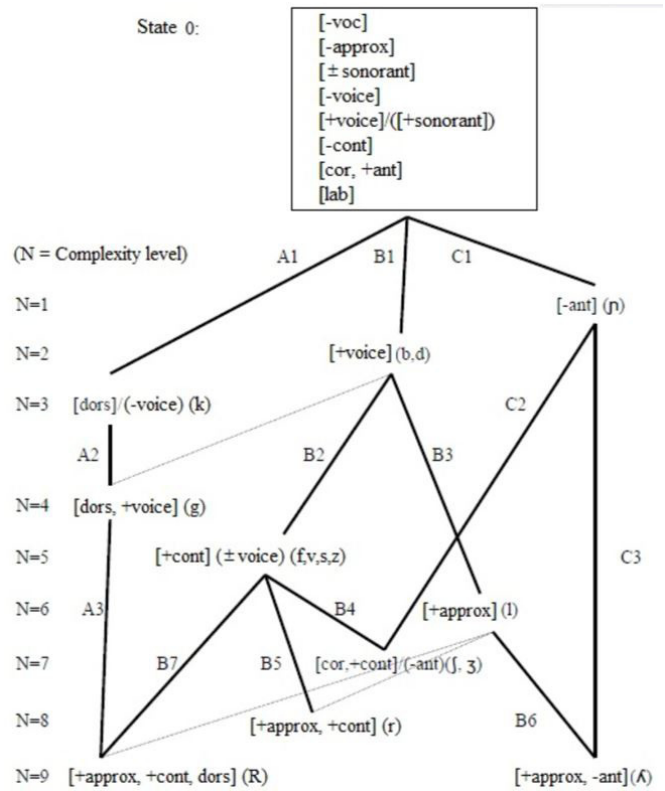

Figure 1. Implicacional Model of Feature Complexity (IMFC) proposed by Mota[2] (p. 154)

The study of vocabulary has not been as widely investigated as phonology has in Brazil. In Brazilian Portuguese, it was found that between one year and four months and one year and six months, the mean growth of vocabulary is four words per month, while between one year and ten months and two years, this growth is 25 words per month[3].

Studies that correlate different fields of language, such as vocabulary and phonology, are even more scarce. One study suggested that changes in vocabulary observed in children with specific alterations in language development can be explained by difficulties observed in skills and/or characteristics directly related to mechanisms involved in information processing, which hinder the quality and recovery of phonological and semantic representations of a new lexical item[4].

Another study conducted in 2012[5], with 15 children with typical development at two years of age, used a spontaneous naming method for words the children already knew in order to analyze the influence of frequency of lexical types, similarity of phonological types, age of acquisition and phonotactic probability on the variability and precision of production. The authors found that phonological complexity plays an important role in that words with late-acquisition phonemes and syllable structures are produced with greater variability and both the frequency of lexical types and the phonological similarity influence the observed variability in speech—the greater the effect of frequency of lexical type and phonological similarity, the less variability in productions.

In addition, children with large vocabularies have phonological systems more complex than those with small vocabularies, since the larger amount of words produced can generate the requirement of a more advanced phonological system[6].

A recent study[7] revealed that children in the early stages of phonological acquisition select the words they will say based on how "pronounceable" they are. Likewise, other authors[8] found that children who produce fricative phonemes early have better vocabulary and syntax skills than those presenting difficulties with fricatives.

This article aimed to verify possible correlations between the number of lexical types and the number of consonants in the general phonological system of children with typical language development.

METHODS

This quantitative and descriptive study with a prospective data collection is part of a project approved by the UFSM Research Ethics Committee under the number 0219.0.243.000-11. Authorization from participants' guardians was requested upon clarification, reading and signing of the informed consent, an essential condition for participation in the study.

The 186 children who participated were aged between one year and six months and five years, 11 months and 29 days, all members of monolingual families who speak Brazilian Portuguese and presented typical language development. Exclusion criteria included: children who presented hearing loss or neurological, emotional, and/or cognitive impairment detectable by means of observation; the presence of motor or organic oral alterations; or those who were participating in speech therapy.

Nine children were between 1 year, 6 months and 1 year, 11 months, 29 days; 13 children were between 2 years and 2 years, 3 months, 29 days; 13 children were between 2 years, 4 months and 2 years, 7 months, 29 days; 16 were between 2 years, 8 months and 2 years, 11 months, 29 days; and there were 15 children in each of the following nine age groups (of three months each) starting with the group with ages between 3 years and 3 years, 3 months, 29 days, and going up to the last group between 5 years, 8 months and 5 years, 11 months, 29 days.

The data were collected at eight municipal elementary schools in the city of Santa Maria-RS, located in different regions of the city. The speech assessment consisted of a questionnaire for parents or guardians, an orofacial praxis assessment, an assessment of oral language and phonetic and phonological aspects of speech and a hearing assessment.

The questionnaire, sent home to parents and guardians, sought information about pregnancy, birth, language and motor development of the child, medical history, current behavior, history of bilingualism, and general aspects of family history and dynamics.

In the oral motor assessment, an orofacial myofunctional protocol with scores (AMIOFE)[9] was adapted to the needs of this research. This protocol was used as an oral motor assessment of appearance, normal position, muscle tension and mobility. Respiratory function was also analyzed.

Oral praxis in children who were three years and six months or older was assessed through a protocol for dyspraxia assessment[10]. Children under the age of three years and six months did not have this aspect evaluated because there are no reference values up to this age group.

Language assessment was performed using a behavioral observation protocol[11], which allowed us to observe cognitive and language development. This protocol was designed for children between one and four years of age, is easy to apply and provides defined reference values. For children over four years, spontaneous conversation was observed through answers to questions and an analysis of small spontaneous oral narratives.

The assessment of phonetic aspects of speech for children three years and six months or older was based on repetition of phonetically balanced words, allowing detection of articulation and phonological alterations that may occur in speech. Children under three years and six months were informally asked to repeat a few words, using playful activities.

Visual reinforcement audiometry, which is used in children between six and 24 months[12], was used to assess hearing in children up to two years, six months and 29 days of age, using a portable pediatric audiometer, with pure modulated tones (*warble)* at frequencies of 500, 1000, 2000 and 4000 Hz, in intensities from 20db to 80 dB presented in a free field. Frequencies were alternated (500, 1000, 2000, 4000 Hz) until reaching the minimum response level for each frequency. Responses between 20 and 40 dB were considered normal[13]. Although this assessment is suitable for children up to two years, it was necessary to extend the age range used, because it was found that children at that age were not able to respond to the playful audiometry.

Children between two years, seven months and five years, 11 months, 29 days underwent a hearing assessment, with conditioned play audiometry or pure tone audiometry[14], using an Interacoustics Screening Audiometer AS208, properly calibrated. Hearing thresholds in air conduction were verified from 500 to 4000 Hz at 20 dB. If there was a failure in the responses, at one or more frequencies, and in two consecutive screens, the child was sent for a complete hearing and otorhinolaryngologic assessment.

Children with appropriate responses underwent the phonological and vocabulary assessments. This assessment was performed using spontaneous speech and naming of small objects and toys, from a pre-elaborated list, selected based on the Child Phonological Assessment (Avaliação Fonológica da Criança, AFC)[15]. This instrument allows an assessment of the possibilities of occurrence for each consonant of Brazilian Portuguese, in all possible positions.

Video recordings were made using a Samsung SMX-C200 camcorder and stored in an external HD for broad phonetic transcription of the childrens' speech and alphabetical transcription of the examiners' speech. For the phonetic transcription of speech of children up to 3 yrs, 3 mths, 29 days, the method of consensus was used[16], i.e. two judges worked independently on the transcription. Then, the transcripts were compared and

discrepancies were heard by a third evaluator until they reached an agreement for all utterances/words/sounds produced by the child. If there was no agreement between at least two evaluators, the passage was deleted, thus, assuring the reliability of the transcriptions, preventing the deletion of a large number of words, since young children, even with typical development, present greater variability in productions and articulatory immaturity.

For the children from the other age groups, which present more stable productions, the following method for reliability between transcriptions was used: all samples were transcribed by an experienced evaluator in child language and a second evaluator with the same experience transcribed, independently, 20% of the same sample for reliability[5,17]. There was a mean agreement of 79.6% at the 3-yr age groups; 81.9% for 4-yr groups; and 80.1% for 5-yr groups.

Filming lasted, on average, 20 minutes, in order to obtain both a representative sample of speech and to ensure the viability of the transcriptions, considering the large number of individuals. The reference for diagnosing phonological alterations were the studies conducted by Regina Lamprecht[1], considering a margin of error, given the uniqueness of phonological acquisition. This reference was chosen since it is from the same state where the current research was conducted.

Contrastive Analysis was used initially for the speech assessments, which consists of four stages: Phonetic Description 1: a record of consonant sound execution; Phonetic description 2: a record of phonetic inventory based on categories of articulation and executions of consonant clusters; Contrastive analysis 1: a record of percentages of occurrences and possibilities of substitutions and omissions made by the child; and Contrastive Analysis 2: which presents the phonological system used by the child, including contrasts, substitutions and omissions[15].

From this analysis, the following criteria were used to establish the phonological inventory[18]: occurrence from 0% to 39% indicates that the phoneme is not acquired; occurrence between 40% and 79% indicates that the phoneme is partially acquired; occurrence equal to or greater than 80% indicates that the phoneme is acquired. To determine the features of the phonological systems of the individuals and the probabilities of producing each sound, the general phonological system of each individual was considered, analyzing consonants /p, b, t, d, k, g, f, v, s, z, ʃ, ʒ, m, n, ɲ, l, ʎ, r, R/ at positions of simple onset, /l, r/ at the second position of complex onset and /s, r/ at the coda position. It should be noted that lateral coda was produced in all cases as the semivowel [w], resulting in a diphthong[19]. Postvocalic nasal consonants were considered to behave like the nasalization of the preceding vowel, being considered, therefore, as a floating autosegment[20]. The nasal consonant may also become an element of the diphthong. Thus, nasal and postvocalic lateral consonants were not considered to be in the coda position.

After completion of the contrastive analyses, the number of phonemes acquired in the general phonological system of each child was counted for each IMFC group[2] - State 0 and Levels 1 and 2; Levels 3, 4 and 5; Levels 6, 7, 8 and 9 and in all possible syllable structures.

For the classification of vocabulary data, the number of lexical types produced was counted in accordance with a previous study[21], using a simple sum of the number of words spoken by each child from the transcriptions. It is important to note that the words spoken in repetition (said immediately after the examiner said them) were not considered.

All the words spoken by each child were entered into a spreadsheet (one item per cell). Then, the repeated items were deleted, leaving only one item of each word. In this way, all the different words produced by the child were considered for the sum of the number of words spoken.

In cases where a word kept the same radical, but changed the suffix, e.g. for number, gender and verb inflection, it was considered a single lexical item (one spoken word). For example, boy/boys; like/liked; *menino/menina* (gender inflection in Portuguese).

In the case of verbal phrases, when two words were used together with a single meaning, as in "going to eat", this was also considered a single lexical item (one spoken word), since the first verb only contributes to express the time frame, but does not contribute to the meaning. Contractions, such as "*pra*" (*para a*), were recorded as a single lexical item (one spoken word).

The mean number of lexical types produced (mean number of words spoken) were then calculated in each age group. As mentioned previously, the different grammatical classes produced were not considered in this study. Since the objective was to analyze the quantity of lexical items produced (how many words the children produce) at a certain age, only the number of words spoken by each individual in each filming was counted in order to establish correlations between these data and that of consonant production.

The numerical data of lexical types (mean number of words spoken at the time determined by age group) and of the phonemes in the general phonological system in each syllable structure and in the different IMFC levels[2] produced by each individual were compared across age groups, using the Statistica 9.1 program and applying the Kruskal-Wallis non-parametric test, followed by multiple comparison tests.

The same program was used to calculate correlations between the number of lexical types produced and the variables established for the phonological system of the individuals in each age group surveyed, using the Spearman rank correlation coefficient, followed by the Student's t-test to check the significance of the correlation. Some correlation coefficients could not be calculated because the values remained constant. For both tests, a significance level of 5% was used ($p \leq 0.05$).

RESULTS

Table 1 presents data for consonants acquired in the general phonological system and in each possible syllable structure by age, as well as the comparison between these means.

To understand this table, it is important to keep a few things in mind: the maximum number of possible phonemes in the

general phonological system and the simple onset position is the same, which is, 19; the maximum number of possible phonemes for complex onset and for coda is two.

Table 2 presents the data for consonant production in each IMFC level[2] by age group and the comparison between these means.

To understand Table 2, it is important to remember that there are seven phonemes in 0 State and Levels 1 and 2: /p, t, m, n, ɲ, b, d/; six phonemes for Levels 3, 4 and 5: /k, g, f, v, s, z/; and six phonemes for Levels 6, 7, 8 and 9: /l, ʃ, ʒ, r, R, ʎ/.

Table 3 presents the mean number of lexical types produced by age group and the comparison between these means.

Table 4 presents the correlations between the number of lexical types produced and the number of phonemes acquired in the phonological system, at the different IMFC levels[2] and in each possible syllable structure in the age groups surveyed. Few correlations showed statistical significance and most of the correlations were positive, that is, with the increase in one of the variables, there was an increase in another.

The only significant negative correlation found was in the age group 5 yrs, 8 months- 5 yrs, 11 months, 29 days between the lexical types produced and the number of phonemes acquired in the complex onset position. This means that the more lexical types children at this age produced, the more unstable the complex onset production was or, conversely, the fewer lexical types

Table 1. Number of consonants acquired in the general phonological system (GPS) and in each syllable structure by age group

Age group	GPS/Simple onset				Complex onset				Coda			
	Mean	Min	Max	SD	Mean	Min	Max	SD	Mean	Min	Max	SD
(A)1:6-1: 11;29	4.8 (F – M)	3	8	1.8	0 (K – M)	0	0	0	0 (H. J – M)	0	0	0
(B) 2:0-2:3;29	9.9 (G – M)	5	15	3.0	0 (K – M)	0	0	0	0 (H – M)	0	0	0
(C) 2:4-2:7;29	12.8 (G – M)	8	18	3.1	0 (K – M)	0	0	0	0.1(H. J – M)	0	1	0.4
(D) 2:8-2: 11;29	15.2 (G – M)	13	17	1.3	0.1 (K – M)	0	1	0.2	0.4(H. J – M)	0	1	0.2
(E) 3:0-3:3;29	16.3 (J – M)	13	18	1.8	0 (K – M)	0	0	0	0.4(H. J – M)	0	1	0.5
(F) 3:4-3:7;29	17.4 (A)	15	19	1.2	0.2 (K – M)	0	1	0.4	0.7(H. J – M)	0	1	0.5
(G) 3:8-3: 11;29	17.7 (A – B)	15	19	1.2	0.3 (K – M)	0	2	0.6	0.8(H. J – M)	0	1	0.4
(H) 4:0-4:3;29	18.4 (A – D)	17	19	0.8	0.3 (L.M)	0	2	0.6	1.9 (A – G)	1	2	0.3
(I) 4:4-4:7;29	18.3 (A – D)	17	19	0.7	0.8	0	2	0.9	1.3 (B)	1	2	0.5
(J) 4:8-4: 11;29	18.7 (A - E)	18	19	0.5	0.9	0	2	0.8	1.9 (A – G)	1	2	0.3
(K) 5:0-5:3;29	18.9 (A - E)	18	19	0.3	1.5 (A – G)	0	2	0.7	1.9(A – G)	1	2	0.3
(L) 5:4-5:7;29	18.7 (A - E)	18	19	0.5	1.5 (A – H)	1	2	0.5	1.9(A – G)	1	2	0.3
(M) 5:8-5: 11;29	18.9 (A - E)	18	19	0.3	1.5 (A – H)	0	2	0.6	1.9(A – G)	1	2	0.3

Caption: Statistical test: Kruskal-Wallis. Level of significance: p ≤ 0.05. Min = minimum; Max = maximum; SD = standard deviation. The superscript letters in parentheses represent age groups with statistically significant difference. The dash (-) indicates significant difference from one age group to another and the comma (,) indicates significant difference between one age group and another. Example: Coda Production in the age group 1:6 to 1:11;29 (A) differs from that of the groups 4:0-4:3;29 (H) and 4:8-4:11;29 (J), 5:0-5:3;29 (K), 5:4-5:7;29 (L), 5:8-5:11;29 (M)

Table 2. Descriptive measures of the number of consonants in each IMFC level[2]

Age group	S0, L1, L2				L3, L4, L5				L6, L7, L8, L9			
	Mean	Min	Max	SD	Mean	Min	Max	SD	Mean	Min	Max	SD
(A)1:6-1: 11;29	2.9 (D – M)	1	5	1.3	1.7 (E – M)	0	3	0.9	0.2 (G – M)	0	1	0.4
(B) 2:0-2:3;29	5.7	2	7	1.5	3.3 (F – M)	1	6	1.5	0.9 (G – M)	0	3	0.9
(C) 2:4-2:7;29	6.5	4	7	1.0	3.9 (F. H– M)	1	6	1.5	2.5 (H – M)	0	5	1.7
(D) 2:8-2: 11;29	7.0 (A)	7	7	0	5.2	4	6	0.7	3.0 (H – M)	1	5	1.0
(E) 3:0-3:3;29	7.0 (A)	7	7	0	5.3 (A)	3	6	1.1	3.9 (J – M)	2	5	1.1
(F) 3:4-3:7;29	7.0 (A)	7	7	0	5.9 (A – C)	4	6	0.5	4.5	2	6	1.0
(G) 3:8-3: 11;29	7.0 (A)	7	7	0	5.9 (A. B)	5	6	0.3	4.9 (A. B)	2	6	1.1
(H) 4:0-4:3;29	7.0 (A)	7	7	0	6.0 (A – C)	6	6	0	5.4 (A – D)	4	6	0.8
(I) 4:4-4:7;29	7.0 (A)	7	7	0	6.0 (A – C)	6	6	0	5.3 (A – D)	4	6	0.7
(J) 4:8-4: 11;29	7.0 (A)	7	7	0	6.0 (A – C)	6	6	0	5.7 (A – E)	5	6	0.5
(K) 5:0-5:3;29	7.0 (A)	7	7	0	6.0 (A – C)	6	6	0	5.9 (A – E)	5	6	0.3
(L) 5:4-5:7;29	7.0 (A)	7	7	0	6.0 (A – C)	6	6	0	5.7 (A – E)	5	6	0.5
(M) 5:8-5: 11;29	7.0 (A)	7	7	0	6.0 (A – C)	6	6	0	5.9 (A – E)	5	6	0.3

Caption: Statistical Test: Kruskal-Wallis; level of significance: p ≤ 0.05. Min = minimum; Max = maximum; SD = standard deviation. S0 = 0 State; L = Level. Superscript letters in parentheses represent the age groups with significant difference. The dash (-) indicates significant difference from one age group to another and the comma (,) indicates significant difference between one age group and another. Example: Phoneme production in L3, 4 and 5 in age group 2:4-2:7;29 (C) differs from groups 3:43:7;29 (F) and 4:0-4:3;29 (H); 4:4-4:7;29 (I); 4:8-4:11;29 (J); 5:0-5:3;29 (K); 5:4-5:7;29 (L); 5:8-5:11;29 (M)

Table 3. Descriptive measures of the number of lexical types produced by age group

Age group	Mean of lexical types	Minimum	Maximum	Standard deviation
(A)1:6-1: 11;29	20.7 [F - M]	8	37	10.1
(B) 2:0-2:3;29	57.9 [F - M]	30	129	28.0
(C) 2:4-2:7;29	96.0 [I,J]	60	158	26.5
(D) 2:8-2: 11;29	107.9 [J]	73	176	28.6
(E) 3:0-3:3;29	103.0 [I,J]	57	143	23.4
(F) 3:4-3:7;29	126.2 [A,B]	73	169	28.9
(G) 3:8-3: 11;29	139.9 [A,B]	78	191	31.3
(H) 4:0-4:3;29	139.8 [A,B]	92	178	22.3
(I) 4:4-4:7;29	172.9 [A - C,E]	89	310	60.8
(J) 4:8-4: 11;29	164.8 [A - E]	98	232	37.9
(K) 5:0-5:3;29	127.2 [A,B]	60	173	34.9
(L) 5:4-5:7;29	132.5 [A,B]	66	187	43.6
(M) 5:8-5: 11;29	135.9 [A,B]	77	205	40.0

Caption: Statistical Test: Kruskal-Wallis; level of significance: p ≤ 0.05. Min = minimum; Max = maximum; SD = standard deviation. Superscript letters in parentheses represent the age groups with significant difference. The dash (-) indicates significant difference from one age group to another and the comma (,) indicates significant difference between one age group and another. Example: Lexical type production in the age group 2:0 to 2:3;29 (B) differs from age groups 3:4-3:7;29 (F), 3:8-3:11;29 (G), 4:0-4:3;29 (H), 4:4:7;29 (I), 4:8-4:11;29 (J), 5:0-5:3;29 (K), 5:4-5:7;29 (L), 5:8-5:11;29 (M); the age group 3:4-3:7;29 (F) differs from groups 1:6-1:11;29 (A) and 2:0-2:3;29 (B)

these children produced, the higher the frequency of correct complex onset production.

DISCUSSION

There was a gradual increase in the mean number of consonants acquired in all possible positions (Table 1) and the phenomenon of regression, which has often been mentioned[1,22] was not observed. Regression is usually analyzed as a single phenomenon and is more easily observed in longitudinal studies[22]. We believe this phenomenon was not observed here due to having considered the means of acquisition of a large number of children and having counted only the number of consonants, without analyzing them separately.

Studies carried out in Rio Grande do Sul, the same state in which this survey was conducted, have suggested complete acquisition of the Brazilian Portuguese phonological system at between five years and five years and two months, with complex onset being the last component to be acquired[1]. However, in the present study, not all children presented a complete phonological system, even at the age of five. This may be due to either individual variations[1] or environmental interferences, since language is the result of the interrelation between the initial state (language acquisition system) and the course of experience[23].

In relation to syllable structures, acquisition followed the sequence: simple onset → *coda* → complex onset, demonstrated by the statistically significant difference across age groups for each structure, as occurred in previous research[1]. This can be explained by the degree of articulatory difficulty in the production

of syllable components, where consonant-vowel (CV) is the simplest structure, followed by C-V-C and then CCV.

With regard to IMFC levels[2], quick stabilization and linear growth were observed in the consonants of 0 State and Levels 1 and 2, since they correspond to stops and nasal phonemes, which are acquired early in Brazilian Portuguese[1,19].

The phonemes of Levels 3, 4 and 5 also showed linear growth, however later stabilization, in the range from four years to four years, three months and 29 days. These levels correspond to stops /k, g/ and fricatives /f, v, s, z/. According to the literature, the first of these can be acquired from one year and seven or eight months[1] up to two years or two years and one month[19]. The fricatives /f, v/ are acquired between one year and eight or nine months[1] up to two years or two years and one month[19]. Acquired a little later, the fricatives /s, z/ were found at between two years and two years and six months[1] up to two years and four months and two years and 11 months[19]. Generally speaking, the children in this study presented later acquisition, because only in the first four-year age group did 100% of the children present all the phonemes for Levels 3, 4 and 5.

For Levels 6, 7, 8 and 9, corresponding to the posterior fricatives and liquids, there was not 100% acquisition at any age group. However, at the age of four years to four years, three months and 29 days, most children had all these consonants acquired, which is consistent with the literature[1,19].

In relation to the lexical types produced, in general there was a gradual increase, as was observed in other studies[24]. The significant increase in lexical types produced between the first age group and that of three years, four months and three years, seven months, 29 days was similar to that found in a German study with a similar methodology[24].

There was a more pronounced increase in the means of lexical types between the age groups in the early stages of acquisition, similar to a study with French children, which found a marked progression in measures of lexical diversity between 24 and 36 months[25]. This may be related to the vocabulary explosion phase, a period of rapid growth in the number of lexical items produced by children[21].

The constant and gradual increase of vocabulary occurs because it is the only language element with an infinite possibility of growth, that is, we acquire new words throughout our entire lives, even though the items used in our daily lives are predominantly acquired in the initial period of acquisition[21].

Also in terms of number of lexical types produced, the standard deviation analysis showed great variability among children within the same age group. This occurs due to individual variations and external factors that may influence the acquisition of language, as already mentioned[23].

As for the interrelations found, the only negative correlation found can be explained by studies that have found that children with large vocabularies have more complex phonological systems than those with small vocabularies. A large vocabulary may generate a demand for a more advanced phonological system[6]. Likewise, changes in vocabulary in children with language changes are due to difficulties in information processing,

Table 4. Correlations between the number of lexical types produced and different phonological aspects of individuals

Age group		Phonological Aspects	Coefficient (r)	p-value
1:6-1: 11;29	Types X	Phonemes (GPS)/simple onset	0.514	0.156
		Phonemes E0, L1, L2	0.699	0.036*
		Phonemes L3, L4, L5	- 0.071	0.855
		Phonemes L6, L7, L8, L9	0.215	0.578
2:0-2:3;29	Types X	Phonemes (GPS)/simple onset	0.585	0.036*
		Phonemes E0, L1, L2	0.240	0.429
		Phonemes L3, L4, L5	0.431	0.142
		Phonemes L6, L7, L8, L9	0.738	0.004*
2:4-2:7;29	Types X	Phonemes (GPS)/simple onset	0.658	0.014*
		Phonemes E0, L1, L2	0.550	0.052
		Phonemes L3, L4, L5	0.748	0.003*
		Phonemes L6, L7, L8, L9	0.214	0.483
		Coda	0.679	0.011*
2:8-2: 11;29	Types X	Phonemes (GPS)/simple onset	- 0.237	0.375
		Phonemes L3, L4, L5	-0.083	0.761
		Phonemes L6, L7, L8, L9	-0.246	0.359
		Complex onset	- 0.204	0.449
		Coda	0.539	0.031*
3:0-3:3;29	Types X	Phonemes (GPS)/simple onset	- 0. 139	0.622
		Phonemes L3, L4, L5	- 0.335	0.222
		Phonemes L6, L7, L8, L9	0.108	0.701
		Coda	- 0.163	0.562
3:4-3:7;29	Types X	Phonemes (GPS)/simple onset	0.160	0.568
		Phonemes L3, L4, L5	- 0.840	0.766
		Phonemes L6, L7, L8, L9	0.235	0.398
		Complex onset	0.444	0.098
		Coda	0.312	0.258
3:8-3: 11;29	Types X	Phonemes (GPS)/simple onset	- 0.248	0.374
		Phonemes L3, L4, L5	0.011	0.968
		Phonemes L6, L7, L8, L9	- 0.289	0.296
		Complex onset	0.029	0.918
		Coda	- 0.377	0.166
4:0-4:3;29	Types X	Phonemes (GPS)/simple onset	0.237	0.396
		Phonemes L6, L7, L8, L9	0.237	0.396
		Complex onset	0.052	0.854
		Coda	0.196	0.484
4:4-4:7;29	Types X	Phonemes (GPS)/simple onset	0.066	0.815
		Phonemes L6, L7, L8, L9	0.066	0.815
		Complex onset	0.094	0.738
		Coda	- 0.342	0.212
4:8-4: 11;29	Types X	Phonemes (GPS)/simple onset	- 0.431	0.109
		Phonemes L6, L7, L8, L9	- 0.431	0.109
		Complex onset	- 0.024	0.934
		Coda	0.013	0.963
5:0-5:3;29	Types X	Phonemes (GPS)/simple onset	0.258	0.353
		Phonemes L6, L7, L8, L9	0.258	0.353
		Complex onset	0.345	0.207
		Coda	0.532	0.041*
5:4-5:7;29	Types X	Phonemes (GPS)/simple onset	0.125	0.657
		Phonemes L6, L7, L8, L9	0.125	0.657
		Complex onset	0.115	0.683
		Coda	0.041	0.884
5:8-5: 11;29	Types X	Phonemes (GPS)/simple onset	0.283	0.307
		Phonemes L6, L7, L8, L9	0.283	0.307
		Complex onset	- 0.663	0.007*
		Coda	- 0.478	0.072

Caption: Student's t-test for significance of the Spearman's rank correlation coefficient; level of significance: p ≤ 0.05; for some variables it was not possible to calculate the correlation coefficient; GPS = General Phonological System; S0 = 0 State; L = Level; Asterisk (*) indicates values with significance difference

which involves the phonological and semantic representations corresponding to a new lexical item[4].

One study[26] that compared performance on an expressive vocabulary test in children with and without phonological deviation showed that the correct naming of items (nouns) on the test was significantly higher in children with typical phonological development. This indicates that the better the phonology, the better the vocabulary will be in general, which is also in corroboration with our results.

Words with phonetic properties that mirror pre-linguistic vocalizations will be acquired sooner than words with features or syllable structures that are not present in the child's pre-linguistic repertoire[6]. This explains why most of the correlations found in the initial age groups were related to phonemes in the simple onset position and to the first levels of IMFC[2].

Most of the positive correlations found in this study were related to coda. A study of four-year-old American children showed that rhyme recognition is directly related to incidental acquisition of new words[27]. Therefore, since coda structures come at the end of the syllable and may also be at the end of the word, it is assumed that children with better rhyming recognition skills will more easily recognize and produce them, which also influences vocabulary expansion.

The only statistically significant negative correlation was found between the production of lexical types and the production of phonemes in the complex onset position in the last age group. One possible explanation for this is that, in this age group, children have a more diverse vocabulary with more difficult grammatical classes and a greater semantic requirement may lead to more errors in phonological production[28]. In addition, this phonological difficulty is more pronounced in the more complex syllable structure and in later acquisition. Moreover, some consonant clusters take longer to be acquired, even if they are high frequency words in the child's input[29].

Thus, it was confirmed that vocabulary and phonology present a similar behavior during typical language development and there are interrelations between them, most of which are positive correlations. It is important to note that these data corroborate a number of international studies[6,27-29], demonstrating that there is a relationship between lexicon and phonology that is common to different languages.

CONCLUSION

In typical language development, the phonological system grows gradually with the child's advancing age, both for the number of consonants acquired in the general phonological system and the different syllable structures related to IMFC levels[2]. In terms of syllable structures, the order of acquisition was simple onset, followed by coda and complex onset. Neither coda nor complex onset presented 100% acquisition, even in the oldest age groups surveyed. In terms of IMFC levels[2], the order of acquisition was: 0 State and Level 1 and 2 phonemes → Levels 3, 4 and 5 → Levels 6, 7, 8 and 9. The latter group did not reach 100% acquisition in any age group.

The behavior of lexical types produced was similar to that of consonants, e.g. there was a gradual progression. However, in the first five-year group there was a small regression, although without statistical significance.

Correlations between vocabulary and phonology were found and most of these were positive, indicating that these systems are interdependent, that is, the more lexical types the child produces, the better his or her phonological system will be in general. Only complex onset presented a statistically significant negative correlation, which was in respect to lexical types in the oldest age group surveyed, indicating that the corresponding syllable structure may require more linguistic processing, negatively impacting vocabulary.

ACKNOWLEDGEMENTS

REFERENCES

1. Lamprecht R, editor. Aquisição fonológica do português: perfil de desenvolvimento e subsídios para terapia. Porto Alegre: Artes Médicas; 2004.

2. Mota HB. Aquisição segmental do português: um modelo implicacional de complexidade de traços [tese]. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul; 1996.

3. Bastos JC, Ramos AP, Marques J. Estudo do vocabulário infantil: limitações das metodologias tradicionais de coleta. Rev Soc Bras Fonoaudiol. 2004;9(1):1-9.

4. Gândara JP, Befi-Lopes DM. Tendências da aquisição lexical em crianças em desenvolvimento normal e crianças com alterações específicas no desenvolvimento da linguagem. Rev Soc Bras Fonoaudiol. 2010;15(2):297-304. http://dx.doi.org/10.1590/S1516-80342010000200024.

5. Sosa AV, Stoel-Gammon C. Lexical and phonological effects in early word production. J Speech Lang Hear Res. 2012;55(2):596-608. http://dx.doi.org/10.1044/1092-4388(2011/10-0113). PMid:22207699.

6. Stoel-Gammon C. Relationships between lexical and phonological development in young children. J Child Lang. 2011;38(1):1-34. http://dx.doi.org/10.1017/S0305000910000425. PMid:20950495.

7. Vihman MM, DePaolis RA, Keren-Portnoy T. The role of production in infant word learning. Lang Learn. 2014;64(2):121-40. http://dx.doi.org/10.1111/lang.12058.

8. Sotto CD, Redle E, Bandaranayake D, Neils-Strunjas J, Creaghead NA. Fricatives at 18 months as a measure for predicting vocabulary and grammar at 24 and 30 months. J Commun Disord. 2014;49:1-12. http://dx.doi.org/10.1016/j.jcomdis.2014.02.003. PMid:24674713.

9. Felício CM, Ferreira CL. Protocol of orofacial myofunctional evaluation with scores. Int J Pediatr Otorhinolaryngol. 2008;72(3):367-75. http://dx.doi.org/10.1016/j.ijporl.2007.11.012. PMid:18187209.

10. Hage SRV. Distúrbio específico do desenvolvimento da linguagem: subtipos e correlações neuroanatômicas [tese]. Campinas: Faculdade de Ciências Médicas, UNICAMP; 2000.

11. Zorzi JL, Hage SRV. PROC: protocolo de observação comportamental: avaliação de linguagem e aspectos cognitivos infantis. São José dos Campos: Pulso Editorial; 2004.

12. Lidén G, Kankkunen A. Visual reinforcement audiometry. Acta Otolaryngol. 1969;67(2):281-92. PMid:5374646.

13. Vieira EP, Azevedo MF. Audiometria de reforço visual com diferentes estímulos sonoros em crianças. *Pró-Fono R. Atual. Cient.* 2007;19(2):185-94. http://dx.doi.org/10.1590/S0104-56872007000200007.

14. Northern JL, Downs MP. Avaliação auditiva comportamental. 5. ed. Rio de Janeiro: Guanabara Koogan; 2005. Audição na infância; p. 129-167.

15. Yavas M, Hernandorena CLM, Lamprecht RR. Avaliação fonológica da criança: reeducação e terapia. Porto Alegre: Artes Médicas; 1991.

16. Shriberg LD, Kwiatkowski J, Hoffmann KA. A procedure for phonetic transcription by consensus. J Speech Hear Res. 1984;27(3):456-65. http://dx.doi.org/10.1044/jshr.2703.456. PMid:6482415.

17. McLeod S, Harrison LJ, McCormack J. The intelligibility in context scale: validity and reliability of a subjective rating measure. J Speech Lang Hear Res. 2012;55(2):648-56. http://dx.doi.org/10.1044/1092-4388(2011/10-0130). PMid:22215036.

18. Bernhardt B. Developmental implications of nonlinear phonological theory. Clin Linguist Phon. 1992;6(4):259-81. http://dx.doi.org/10.3109/02699209208985536. PMid:20670203.

19. Hernandorena CLM. Aquisição da fonologia do Português: estabelecimento de padrões com base em traços distintivos [tese]. Porto Alegre: Pontificia Universidade Católica do Rio Grande do Sul; 1990.

20. Mateus MHM, D'Aandrade E. The phonology of portuguese. Oxford: Oxford University Press; 2000.

21. Vidor DCGM. Aquisição lexical inicial por crianças falantes de Português Brasileiro: discussão do fenômeno da explosão do vocabulário e da atuação da hipótese do vies nominal [tese]. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul; 2008.

22. Wiethan FM, Mota HB. Inter-relações entre aquisição fonológica e lexical: um estudo longitudinal. Dist Comun Hum. 2014;26(3):518-27.

23. Chomsky N. Novos horizontes no estudo da linguagem e da mente. São Paulo: Editora UNESP; 2005.

24. Kauschke C, Hofmeister C. Early lexical development in German: a study on vocabulary growth and vocabulary composition during the second and third year of life. J Child Lang. 2002;29(4):735-57. http://dx.doi.org/10.1017/S0305000902005330. PMid:12471971.

25. Le Normand M, Parisse C, Cohen H. Lexical diversity and productivity in French preschoolers: developmental, gender and sociocultural factors. Clin Linguist Phon. 2008;22(1):47-58. http://dx.doi.org/10.1080/02699200701669945. PMid:18092219.

26. Brancalioni AR, Marini C, Cavalheiro LG, Keske-Soares M. Desempenho em prova de vocabulário de crianças com desvio fonológico e com desenvolvimento fonológico normal. Rev CEFAC. 2011;13(3):428-36. http://dx.doi.org/10.1590/S1516-18462010005000011.

27. Ramachandra V, Hewitt LE, Brackenbury T. The relationship between phonological memory, phonological sensitivity, and incidental word learning. J Psycholinguist Res. 2011;40(2):93-109. http://dx.doi.org/10.1007/s10936-010-9157-8. PMid:20872250.

28. Leonard LB, Schwartz R, Morris B, Chapman K. Factors influencing early lexical acquisition; lexical orientation and phonological composition. Child Dev. 1981;52(3):882-7. http://dx.doi.org/10.2307/1129090.

29. Ota M, Green SJ. Input frequency and lexical variability in phonological development: a survival analysis of word-initial cluster production. J Child Lang. 2013;40(3):539-66. http://dx.doi.org/10.1017/S0305000912000074. PMid:22452825.

Author contributions