Development of the Whole-Word Structure in Russian-Speaking Children: Longitudinal Study

Ingrida Balčiūnienė[1,2,*] and Aleksandr N. Kornev[2]

[1]*Vytautas Magnus University, V. Putvinskio g. 23-206, 44243 Kaunas, Lithuania*

[2]*Saint-Petersburg State Pediatric Medical University, Ul. Litovskaya d. 2, Saint-Petersburg, Russia*

Abstract: There are many studies devoted to the segmental level of a phonological development. From this perspective, the phonological development is discussed as a sequence of separate phonemes acquisition. In contrary, following the whole-word analysis [1], the main point is how a child copes with the whole-word acquisition. For the latter approach, a longitudinal corpus of spontaneous speech is the most suitable data.

The aims of the current study were: a) to estimate a distribution of word-structure patterns with different phonological and articulation complexity, produced by Russian child (2;04-6;05) during his daily conversations with adults; b) to analyze dynamic changes of this distribution. All the words were analysed according the pattern of their syllable structure by means of specially designed software [2] and classified into 22 basic groups; their percentage distribution was estimated. Statistical analysis revealed the most frequent syllabic structures: CVCVC and CVCCV (9-16%) and CVCVCV (8-10%). To verify an influence of the word-structure complexity on a frequency of the given words, their pMLU index [3] was estimated. The regression statistical analysis evidence a high significant determination of the word-structure complexity on their frequency in daily conversations; however, a force of this determination decreased from the age 2;04 to 6;05.

Keywords: Language acquisition, phonology, longitudinal study, corpus study.

INTRODUCTION

There are many studies devoted to a segmental level of a phonological development. From this perspective, the phonological development is discussed as a sequence of separate phonemes acquisition. In contrary, following the whole-word analysis [1], the main point is how a child copes with the whole-word structures acquisition.

Phonetic and syllabic structure of words considered one of the main characteristics of language and speech development [3]. Previous studies suggest that phonotactic constraints may act as a filter for a lexical acquisition [4]. In this regard, the new words acquisition and production should be influenced in children by the word structure complexity and the phonemes sequence length. Phonological complexity of words depends on the length of the given word (a number of sounds), the number and complexity of the syllables, and the presence of consonant clusters [1, 5]. Since the syllable is the main unit of motor programming and speech production, an acquisition of syllabic structures of speech is one of indicators for a formation of speech articulatory base. It is known that open syllables are the easiest to master [5, 6, 7]. The earliest syllabic structures acquired by a child are V and CV. Most of the early word shapes belong to an iterative type CVCV [5, 6]. Already in the second year of life, children start using closed CVC syllables [5, 7]. However, a data on a further dynamics (after 2 years) of the formation of syllabic structure, both on the lexical and on sub-lexical level, is extremely scarce.

Following previous studies, an acquisition of syllabic structures and mastering a pronunciation of target speech sounds are in a competitive relationship due to the struggle for cognitive resources. Difficulties in acquiring new sounds often lead to temporary compensatory simplifications of syllable and/ or sound structure in words with the given sounds [9]. This might be also recognized as so-called trade-off effect, i.e. a simplification of some characteristics due to the high complexity of others, e.g., an omission of one consonant in a cluster in multisyllabic words, but a correct pronunciation of the same cluster in mono- and disyllabic words.

According to W. Levelt and his followers, humans acquire not only a vocabulary but also so-called mental syllabary, i.e. a set of (high-frequency) syllables of a given language ready-made whole gestural scores [10]. Children learn a new word easier if its phonological structure is similar to an acquired already words [11]. The most of these findings have been based on English-speaking children's data; however, the patterns of word structure are different across languages. In Russian, the mean length of word is 3-4 syllables, and a word length varies from 2 to 18 phonemes. There are more than ten structural types of syllables that

*Address correspondence to this author at the Vytautas Magnus University, V. Putvinskio g. 23-206, 44243 Kaunas, Lithuania; Tel: +370 37 327 833; E-mail ingrida.balciuniene@vdu.lt

E-ISSN: 2311-1917/19

compose words, e.g., V, C, VC, CV, CVC, CCV, CCVC, CVCC.

Most studies on the phonological development that have dealt with characteristics of the syllabic structure have been based on a selective, mainly qualitative analysis. This limitation might be caused by extremely time-consuming linguistic structural analysis. It must be noted here that a developmental dynamics of word structure acquisition in Russian-speaking children has never been investigated. Therefore, in this study we had several **aims**: a) to estimate a distribution of word structure types in a discourse of Russian-speaking child during his daily conversations with his caregivers; and b) to analyze the dynamic changes of word syllabic structures from 2 to 6 years of life.

METHODOLOGY

A longitudinal evaluation method was employed for the collection and grouping of the data [12]. According to this method, 15-30-minute-long conversations between the child (2;04-6;05[1]) and his caretakers were recorded by a portable digital recorder in a home environment. The recordings were done at different times of the day, which mostly depended on the child's willingness to communicate. Most of the recordings are conversations between the boy and his mother; some dialogues are between the child and his grandmother. The collected recordings were grouped by months trying to maintain similar size and length within each month's data.

All the recordings were transcribed orthographically by a mother (professional linguist) [13] of the child and checked independently by two experts (professional linguists); the size of the corpus is 20,955 word tokens. In highly inflecting languages with rich morphology and derivation system, such as Russian, usually *word tokens*, *word types* or/ and *word lemmas* are analyzed. A number of tokens is the total number of words in a given text; types are different morphological and/or derivational forms of words; and lemmas are root forms in each word class. In this study, the *word types* were analyzed, i.e. words such as *ptica* 'a bird', *pticy* 'birds', and *ptichka* 'a bird:diminutive' were considered as different units.

The corpus was analysed by means of specially designed software for a text analysis – the PASTA (*Programme for Automated Structural Text Analysis*) [2].

All the words from different age periods (2;04-2;06, 2;09-2;11, 4;02-4;06, 5;07-5;10, and 6;00-6;06) were structurally analyzed and classified into 22 groups according to the basic types of Russian word syllable structures. It should be noted here that a huge amount of different word structures exists in Russian. Structural typing depends on a) a type and number of syllables, b) a number of clusters and their position in the word, and c) a total number of phonemes. Hence, to simplify this multitude, some structures were combined together and thus they compose 22 types from the simplest (C, V, CV, VC) to the most complex (13 phonemes with 6 syllables) ones. See [14] for more on the word structure in Russian.

To carry out a quantitative analysis of the word structure complexity, the *Phonological Complexity Index* (PCI) was estimated. The PCI represents a modified version of the *Phonological Mean Length of Utterance* (pMLU) index suggested by D. Ingram [1] for measuring a phonological complexity of utterance and words. PCI = a number of souds + a number of conconants + a number of consonant clusters (a tripple cluster was equaled to two clusters). The PCI was estimated for each type of word structure – $PCI_{i\text{-}type}$.

A distribution of the whole-word structures was estimated and submitted to regression analysis.

RESULTS

First, statistical analysis evidenced quite uneven distribution of words with a different syllabic structure. The least prevalent word structures in the corpus were those composed of 4 and more syllables. They constituted from 0.5% (at 2;04) to 4% (at 6;06) of all the word types (see Figure **1**).

The most frequent syllabic structures (starting with 2;9) were: 1) CVCVC (e.g., *domik* 'a house:diminutive') and CVCCV (e.g., *mesto* 'a place') (9–16%) and 2) trisyllabic constructions such as a) CVCVCV (e.g., *moloko* 'milk') (8–10%), b) CVCCVCVC (e.g., *razdavil* 'squashed'), CVCVCCV (e.g., *l'agushka* 'a frog'), CVCCVCV (e.g., *medvedi* 'bears') or CVCVCVC (e.g., *chelovek* 'a person') (8–9%). Word structures such as CV, VC, and CVC were relatively rare in all age samples (4–6%). Single-syllable words like CCV were extremely rare (<1% of all the word types).

In general, a dynamics of a percentage of different words structures during the analyzed period was not

[1]Hereinafter this format refers to the age of the child: 2;04 means two years and four months.

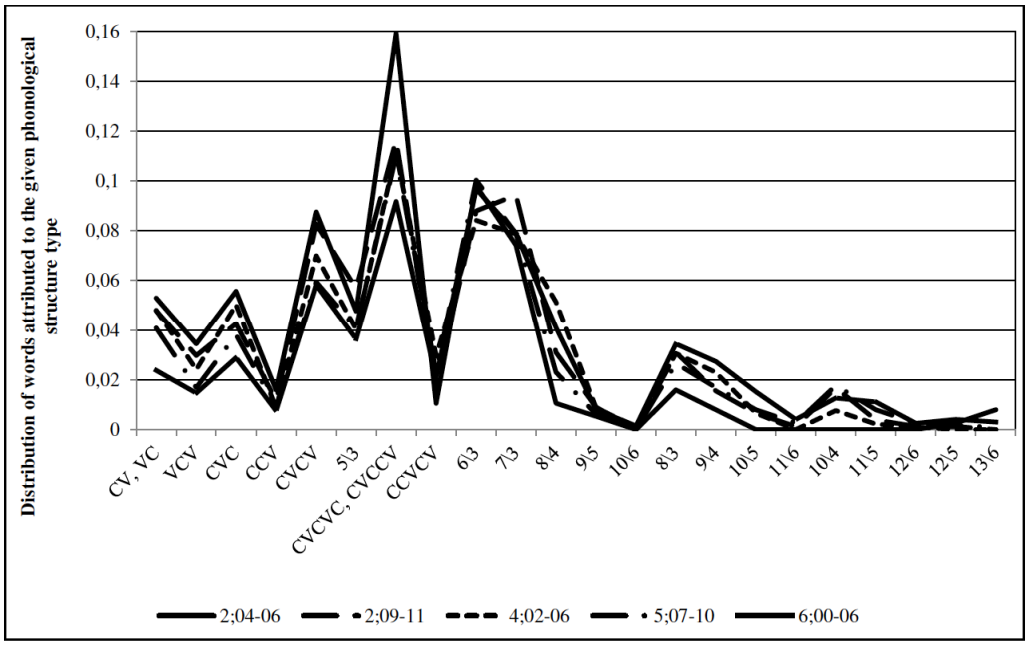

Figure 1: Distribution of word types with different syllabic structures. The numerical labeling of the word structure types: the first figure refers to a number of phonemes in the given word structure type; the second figure refers to a number of syllables in the given word structure type.

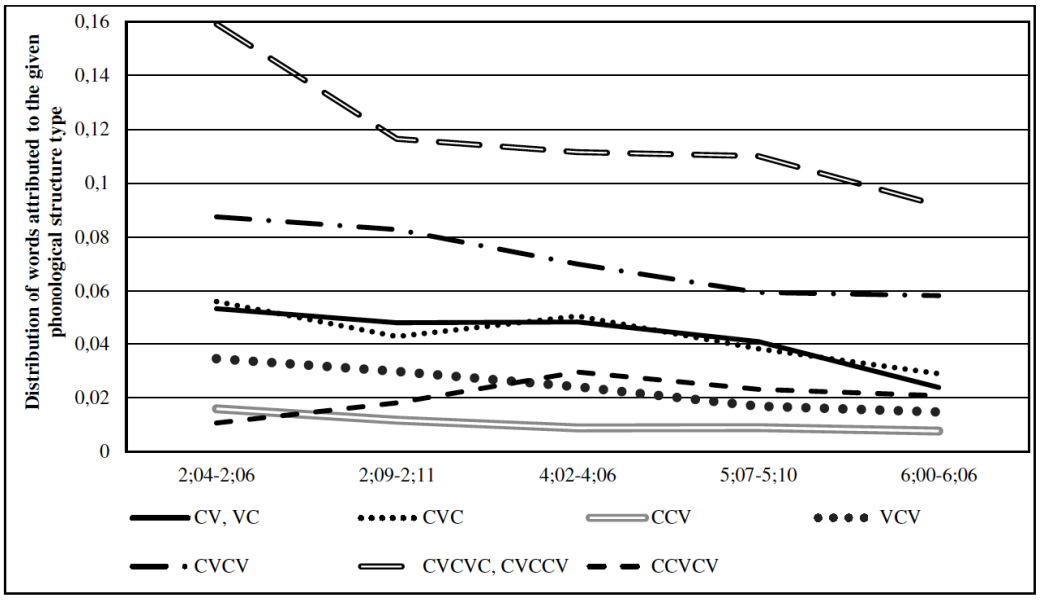

Figure 2: Dynamics of distribution of word types with simple syllabic structures.

significant. The simplest word structures (composed of 1–3 syllables) remained the most frequent, although slightly decreased. The greatest decrease was evident in word structures CVCVC and CVCCV: they constituted from 16% (at 2;04) to 9% (at 6;06) of all word types (see Figure **2**).

More complex word structures, in contrast, slightly increased during the analyzed period. The most evident increase was observed in the CVCVCV structure: such word types constituted from 1% (at 2;04) to 4% (at 6;06) of all word types (see Figure **3**).

To determine a potential effect of the syllabic complexity of words (PCI measure) on the probability of their use by the child, the regression analysis carried out. The coefficient of regression (ß)[2] between the

[2]The ß coefficient of regression measures the extent of the determination of the dependent variable by the independent variable.

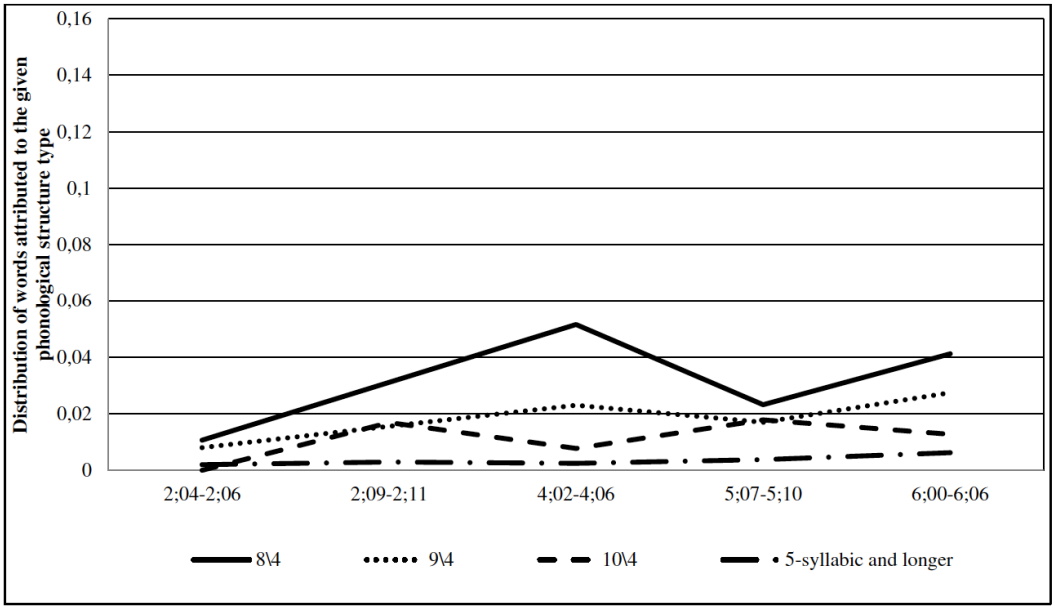

Figure 3: Dynamics of distribution of word types with complex syllabic structures. The numerical labeling of the word structure types the same as in Figure **1**.

$PCI_{i\text{-}type}$ and the frequency of word types with different syllabic structures was calculated. The coefficient ß was ranged from -0.69 to -0.46 in the age periods and for all of them the degree of significance was quite high (from P = 0.0001 to P = 0.043) (see Table **1**).

Table 1: Regression between the $PCI_{i\text{-}type}$ and word Structure Types Incidence

Age period	β	Sig	R^2
2;04-2;06	-0.57	0.004	0.33
2;09-2;11	-0.69	0.000	0.48
4;02-4;06	-0.69	0.000	0.47
5;07-5;10	-0.67	0.000	0.45
6;00-6;06	-0.46	0.043	0.18

This means that the phonological complexity of the given word structure determinates the incidence of using the word in the child discourse. At the age of 2;04 to 2;11, the coefficient ß was the highest (ß = -0.57 and ß = -0.69) and the index R^2 was equal to 0.48[3]. This means that such variable as the word structural complexity controls 48% of the word structures distribution. In this regard, a phonological complexity of a word structure should be recognized as a very influential on word shapes distribution.

[3]The R^2 index measures how much the part of the dependent variable dispersion is controlled by the independent variable.

DISCUSSION AND CONCLUSIONS

Traditional linguistic approach to phonological development aims at tracing an acquisition of separate phonemes. Studies in speech development, on the other hand, usually address a progress in mastering separate sound articulation.

Both the approaches are very fruitful; however, in both of them, the natural speech act is dissected into several rather artificial units. The natural units of speech programming are syllables [1, 6, 10]. Moreover, in the early stage of speech development children master their word production holistically, as a sincret [15]. Following the usage-based approach, the whole-word structure analysis (WWSA) [1] seems to be the most naturalistic methodology for studies in phonological development. A word is a multifaceted entity characterized by particular sounds' pattern, articulation gestures' score, and a cluster of meanings. The current study addresses the words as the structural entities. WWSA and word-structure typology shed a light on some specific line of speech development, i.e. on a mastering the more and more complex articulatory patterns.

The syllabic structure seems to be some template for the phonological and the articulatory word building. The more templates a child has been mastered, the more various word forms occur in his/ her productive vocabulary. The first question is how early a child acquires rich enough devices to produce the multiple

word structures. The second question addresses the strategic patterns of this device development.

The data revealed in current study evidenced that the vocabulary child used in his daily conversations comprised a wide range of word structure types. The most prevalent were the simple disyllabic structures CV and CVC. These results are close to other findings obtained in English-speaking toddlers [1, 6, 8]. However, an unexpected fact was obtained in our corpus: the same ratio remained from the 2;04 until the 7th year of life. Moreover, generally the word shapes distribution did not change considerably in the child over the four years. Two explanations might be suggested. First, the child had mastered an articulation of almost all phonological structures at the 2nd year of life. Second, a genre of the daily conversation required quite low demand for using complex and long words. In this regard, the child's verbal behavior displayed the minimal level of his language competence. Our previous findings published elsewhere have evidenced in favor the second explanation.

The new approach to evaluation of phonological and lexical development seems rather fruitful both in the TD children and the clinical population.

ACKNOWLEDGEMENTS

The study was carried out with the financial support from the Russian Science Foundation, grant No. 18-18-00114.

REFERENCES

[1] Ingram D. The measurement of whole-word productions. Journal of Child Language 2001; 29: 713-733. https://doi.org/10.1017/S0305000902005275

[2] Kornev AN, Balčiūnienė I, Nedorya AE. Stanovlenie zvuko-slogovoy struktury y rebenka: novyy metod analiza korpusnyx dannyx [Development of word structure: A new methodological approach]. In Sed'moy mezhdisciplinarnyy seminar "Analiz razgovornoy russkoy rechi" (AR3-2017) 2017; pp 53-58.

[3] Zamuner TS, Kharlamov V. Phonotactics and syllable in infant speech perception. In: Lidz J, Snyder W, Pater J, Eds. The Oxford Handbook of the Developmental Linguistics. Oxford University Press 2016; pp. 27-42. https://doi.org/10.1093/oxfordhb/9780199601264.013.3

[4] Vihman MM. Variable paths to early word production. Journal of Phonetics 1993; 21: 61-82. https://doi.org/10.1016/S0095-4470(19)31321-X

[5] Gvozdev AN. Voprosy izucheniya detskoy rechi [Issues of child language studies]. M. 1963.

[6] Ingram D. The measurement of whole-word productions. Journal of Child Language 2002; 29: 713-733. https://doi.org/10.1017/S0305000902005275

[7] Eliseeva MB. Foneticheskoe i leksicheskoe razvitie rebenka rannego vozrasta [Early stages of phonetic and lexical development of Russian child]. SPb.: Izd-vo RPGU im. A. I. Gercena 2008.

[8] Levelt CC, Schiller NO, Levelt WJ. The acquisition of syllable types. Language Acquisition 2000; 8(3): 237-264. https://doi.org/10.1207/S15327817LA0803_2

[9] Vihman MM. Later phonogical development. In: Bernthal JE, Bankson NW, Eds. Articulation and Phonological Disorders. Prentice Hall 1988; pp. 110-144.

[10] Levelt WJ, Wheeldon L. Do speakers have access to a mental syllabary? Cognition 1994; 50(1-3): 239-269. https://doi.org/10.1016/0010-0277(94)90030-2

[11] Stoel-Gammon C. Relationships between lexical and phonological development in young children. Journal of Child Language 2011; 38: 1-34. https://doi.org/10.1017/S0305000910000425

[12] Blom E, Unsworth S. Experimental Methods in Language Acquisition Research. John Benjamins Publishing Company 2010. https://doi.org/10.1075/lllt.27

[13] Ivanova KA. Imperative in child-adult communication. Acta Linguistica Petropolitana 2015; 11(1): 565-584.

[14] Kornev AN, Rakhlin N, Grigorenko EL. Dyslexia from a cross-linguistic and cross-cultural perspective: The case of Russian and Russia. Learning Disabilities: A Contemporary Journal 2010; 8(1): 51-78.

[15] Bates E, Marchman V, Thal D *et al.* Developmental and stylistic variation in the composition of early vocabulary. Journal of Child Language 1994; 21(1): 85-123. https://doi.org/10.1017/S0305000900008680

Received on 02-12-2019 Accepted on 12-12-2019 Published on 17-12-2019

DOI: https://doi.org/10.12970/2311-1917.2019.07.04