INTERSPEECH 2009 BRIGHTON

The Acoustic Characteristics of Russian Vowels in Children of 6 and 7 Years of Age

Elena E. Lyakso, Olga V. Frolova, Aleks S. Grigoriev

The Child Speech Research Group
Saint-Petersburg State University, Saint-Petersburg, Russia
lyakso@gmail.com

10.21437/Interspeech.2009-523

Abstract

The purpose of this investigation is to examine the process of acoustic features of vowels from child speech approaching corresponding values in the normal Russian adult speech. The vowels formants structure, pitch and vowels duration were examined. Word stress and palatal context influence on the formants structure of the vowels were taken into account. It was shown that the word stress is formed by 6 -7 years of age on the basis of the feature typical for Russian language. Formant structure of Russian vowels /u/ and /i/ is not formed by the age of 7 years. Native speakers recognize the meaning of 57-93% words in speech of 6 and 7-years-old children.

Index Terms: child speech, pitch, formant structure, child speech perception

1. Introduction

There are six vowel phonemes in Russian language: /a/, /o/, /u/, /i/, /e/, /ɨ/. The first two formants (F1, F2) are used to describe vowels from adult speech, but it is known that these characteristics are inadequate when applied to child speech vowels [1]. Because of the F0 high values the first two formants absolute values and their location on the two-formant plane are not sufficient to describe the vowel-like sounds in children of the first year of life. The F1 and F2 values of infants' vowel-like sounds in the first month of life are shown to occupy the two-formants space areas that do not correspond to these occupied by the corresponding vowels in the adults, lying in an area of significantly higher frequencies. The areas occupied by the sound pairs /o/, /u/ and /e/, /ɨ/ are almost joint together; those occupied by the vowel-like /i/, /a/ overlap in F1 values and partly overlap in F2 values. The difference between the formant values and their magnitudes may be used instead [1]. Nevertheless adults detect vowels in vocalizations of 3-month-old infants [2]. As we shown before native speakers start recognizing children's words meaning correctly (probability 0.75) at the beginning of the second year their life. The vocal phonemes from children's words and syllabic structure of words in children's utterances are recognized with the great success [3].

The meaning of 12-75% words of 3-year-old children is recognized by adults. 67-100 % vowels are detected by native speakers in words of 3-year-old children [4]. As compared to the end of the second year of life, by three years of age the amount of calmly pronounced words increases where pitch and formants values are close to these in adult speech. Still, F0 values are high (332±73 Hz) and make it impossible to describe the vowels /i/, /u/, /ɨ/ in terms of the first two formants absolute values, but these vowels are identified clearly in the words (independently of their stressed or unstressed position) [4]. Although the acoustic features of adult speech are not totally formed, it is shown that by the end of the third year of life some words pronounced by the children become comprehensible for the adults in absence of situational context. During the third year of life the vowel stress begins forming. In all the vowels the stressed vowel's duration tends to be longer than that of the unstressed one, but this difference is significant in some of the vowels only. The palatalized versus non-palatalized consonants opposition development begins, which is found in the characteristics of vowels following them [4].

At 4, 5 years of age the stressed vowel and its stationary part duration, as well as their difference, is higher in the stressed vowels than in unstressed ones. The formants characteristics do not correspond to those in adult speech yet, but may be used to distinguish between the vowels. The formants triangle for children's vowel-like sounds and vowels does not correspond to the basic allophones triangle in adult speech up to 5 years of age. Still the vowels differ by their characteristics from each other and usually they form a triangle, although its orientation for some children differs from that in adult speech [5].

Current study is the part of longitudinal investigation of Russian children speech development. The purpose of current study is to examine the further process of pitch (F0), first and second formants (F1, F2) and duration of the vowels from child speech approaching the corresponding values in the normal Russian adult speech.

2. Methods

A longitudinal audio recording of sound signals of ten normally developing Russian children during the sixth and seventh years of their life has been performed. Two 6-year-old boys Ar, Ni and three 6-year-old girls An, Li, Sf; two 7-year-old boys Sa, Eg and three 6-year-old girls Sn, Na, Da participated in the recording. The children were recorded while interacting freely with their mothers and the investigator. The recordings were made by the "Marantz PMD222" recorder with a "SENNHEIZER e835S" external microphone. The sounds were analyzed in the Cool Edit Pro (Syntrillium Soft. Corp. USA) and the Praat vers. 4.3.

100 words were selected from each child's speech for further analysis. The pitch and the first two formants as well as the vowel duration were measured. To consider the word stress development, the vowel duration in the stressed versus the unstressed vowels were compared, as well as the pitch and formants values in the stationary parts of the vowels.

The same parameters were compared in /a/, /i/ and /u/ after the following consonants: /k/ and /d/ for /a/, /b/ and /g/ for /u/ and /t'/ for /i/. These consonants cause the minimal

articulatory and hence acoustic influence on the corresponding vowels in Russian.

Vowel triangles were constructed for stressed vowels from the speech of each child at 6 and 7 years of age. The vowels were selected with the pitch up to 350 Hz. For /a/, /o/ and /u/ the vowels only in non-palatalized context were included in the graphs, as palatalization causes these vowels to differ a lot from the base allophone. We did not consider the vowel /ɨ/ because of its complex acoustical nature (it is not quite comparable to other vowels). We considered [e] as the basic allophone of the phoneme /e/, as it is the more common one. Thus we analyzed the rest of the phonemes, namely /i/ and /e/, in palatalized context only.

In every child 30 words from phrases were chosen and taken to form test sequences. The test sequences were presented to Russian native speaking adults. The amount of native speakers was 100 auditors (age – 19 years, women – n=55; men – n=45). The aim of the perception analysis was to investigate how the meaning of words produced by the 6 - 7-years-old children was recognized by native speakers. Each word was presented three times in the tests, the pause between presentations being 5 s, while the pause between diverse words was 15 s. The signals sequence was presented to the native-speakers. Mann – Whitney, Kruskal - Wallis test were used.

3. Results

In all the children at 6 and 7 years of age the vowels duration is higher in the stressed vowels than in unstressed ones (fig. 1). The median values of vowels pitch (F0) did not differ significantly between 6 and 7-years-old children. The vowel context and quality were not taken into account when comparing the F0 in stressed versus unstressed vowels. F0 in stressed versus unstressed vowels differed only in one 7-year-old child Sn (fig. 2). It was found that pitch value is individual characteristic of each child.

The arrangement of the formants values on a two-formant plot is different in every child at every age. In the Figure 3 an example is given of formant triangles of vowels /a/, /u/ and /i/ in two children An and Eg, compared to those in adult speech (fig. 3). The stressed vowel /a/ in one 7-year-old child Eg chosen taking into consideration the context and pitch (see methods) coincides with the basic allophone from adult speech. The shape and orientation on two coordinate plot of these children's stressed vowels triangle where the vowels were chosen with regard to factors of context and pitch (triangle 1) correspond to the adult stressed vowels triangle, but the former is moved up and right.

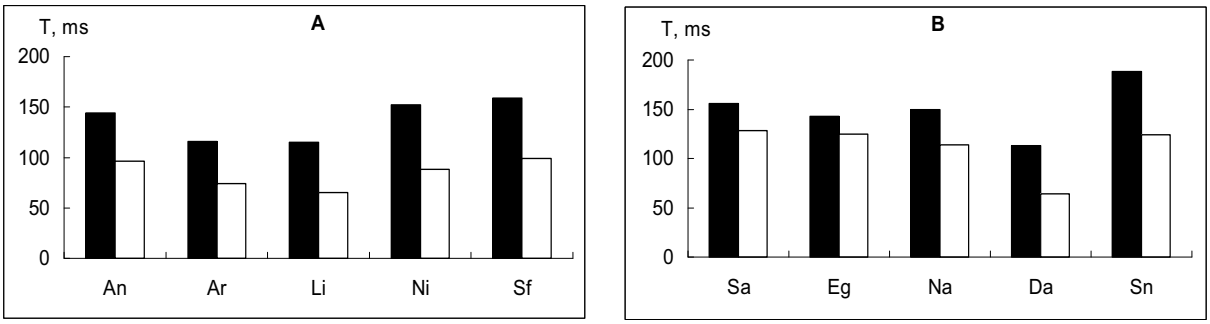

Figure 1. *Vowel duration at 6 (A) and 7 (B) years of age (median values). Black column – stressed vowels; white column– unstressed vowels. Vertical axis values are duration - T, ms.*

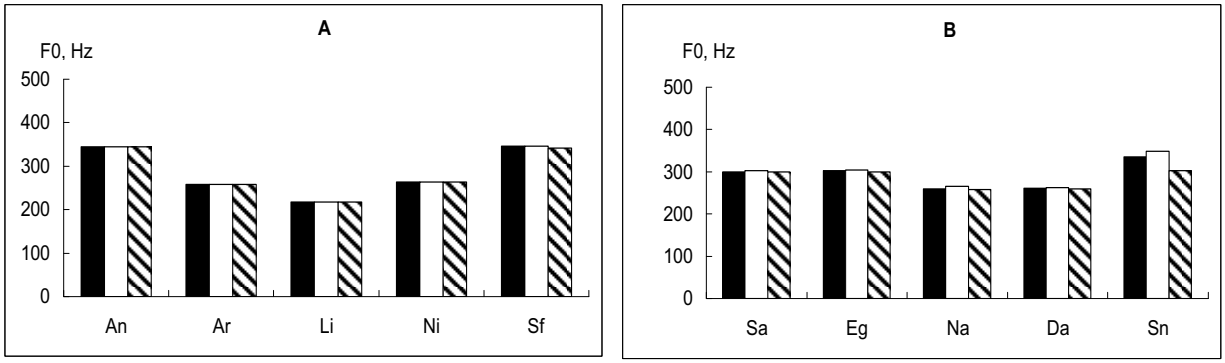

Figure 2. *The pitch median values in the vowels from the words pronounced by children of 6 (A) and 7 (B) years of age. Black column – F0 values of all the vowels; white column – F0 values in stressed vowels; shaded column – that of the unstressed vowels. Vertical axis values are F0, Hz.*

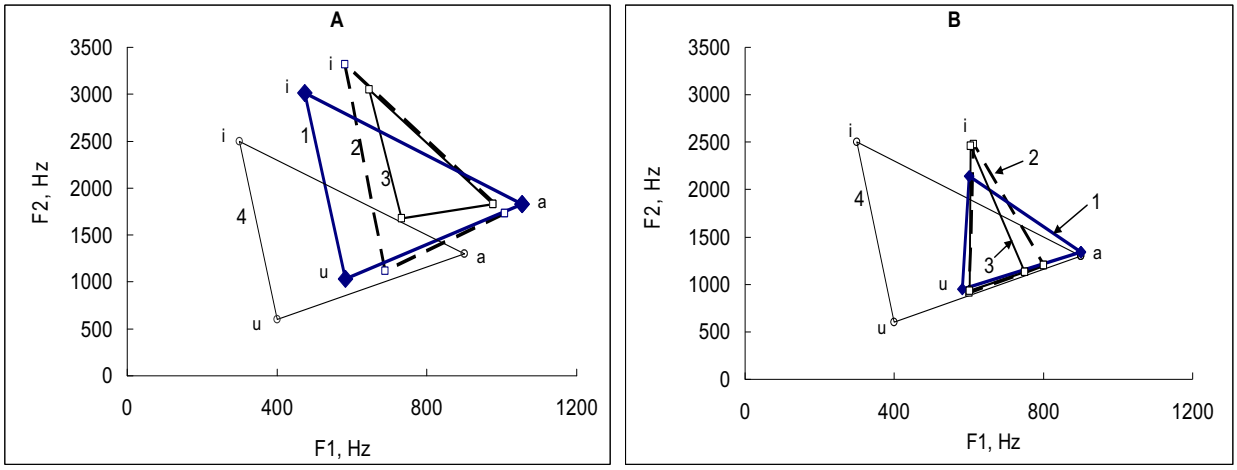

Figure 3. *Vowel formant triangles with apexes /a/, /u/, /i/ in words pronounced by children An, Eg and in adult speech. A – data for 6-year-old An; B – data for 7-year-old Eg. 1 (bold lines) – the triangle of stressed vowels chosen taking into consideration the context and pitch (see methods); 2 (interrupted lines) – triangle of stressed vowels not taking into consideration the above factors; 3 - the triangle of unstressed vowels not taking into consideration context and pitch; 4 – the triangle of basic allophones in adult speech [6]. Horizontal axis values are F1, Hz, vertical axis values are F2, Hz.*

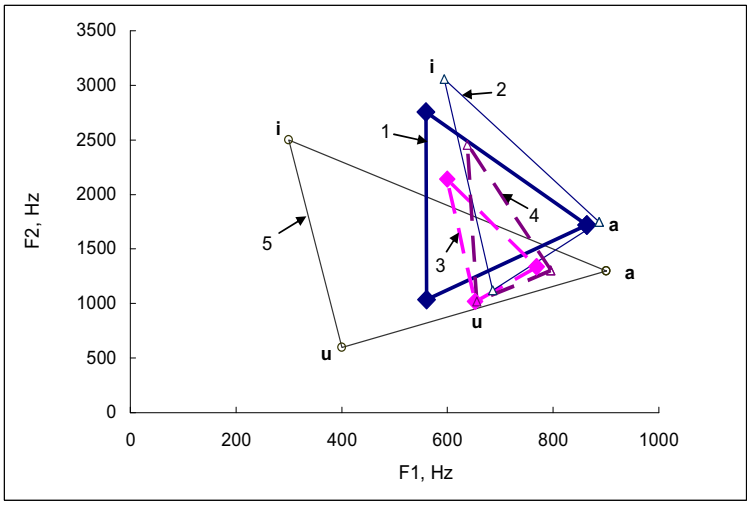

Figure 4. *Vowel formant triangles with apexes /a/, /u/, /i/ for stressed vowels from words pronounced by children of 6 and 7 years of age and adults. 1 (bold lines) –the triangle of vowels in 6-year-old children speech chosen taking into consideration the context and pitch (see methods); 2- the triangle of vowels in 6-year-old children speech not taking into consideration the above factors; 3 - (interrupted lines) – the triangle of vowels in 7-year-old children speech chosen taking into consideration the context and pitch; 4 (interrupted lines) –the triangle of vowels in 7-year-old children speech not taking into consideration these two factors; 5- the triangle of basic allophones in adult speech [6]. - Horizontal axis values are F1, Hz, vertical axis values are F2, Hz.*

F1, F2 median values of vowels /a/ were chosen with regard to factors of context and pitch are situated in the area of two-formant plot typical for F1, F2 values of adult's /a/ (fig. 4), but they are not coincide with the basic allophone from adult speech [6]. F1 median values of /i/ and /u/ vowels are higher than in adult speech.

The unstressed vowels triangle's shape and orientation differs from the stressed vowels triangles for all children. The tendency to decrease of unstressed vowels' formant triangle area on the two-formant plot versus stressed one was revealed in 6-year-old children: Sf, Li, Ni and An (fig. 3) and in 7-year-old children: Sn, Sa, Na and Eg (fig. 3).

The shape and orientation of the stressed vowels triangle where the vowels are chosen without regard to context and pitch is not correspond completely to the stressed vowels triangle where the vowels were chosen with regard to these factors mainly in 6 years old children (fig. 4). 6-years-old children's stressed vowels triangle where the vowels are chosen with regard to context and pitch occupies larger area on the two-formant plot then the same triangle built for vowels of 7-year-old children (fig. 4).

Adults recognize the words meaning of 6-year-old children slightly better then words meaning of 7-year-old children. Native speakers recognize correctly (with probability 0.75) 80% (57 - 93%) words of 6-year-old children and 74 % (57 - 93 %) words of 7-year-old children (fig. 5). The words of 6-year-old children are recognized correctly by 70% (52 - 83%) of native speakers; the words of 7-year-old children are recognized by 62% (50 - 72%) of native speakers.

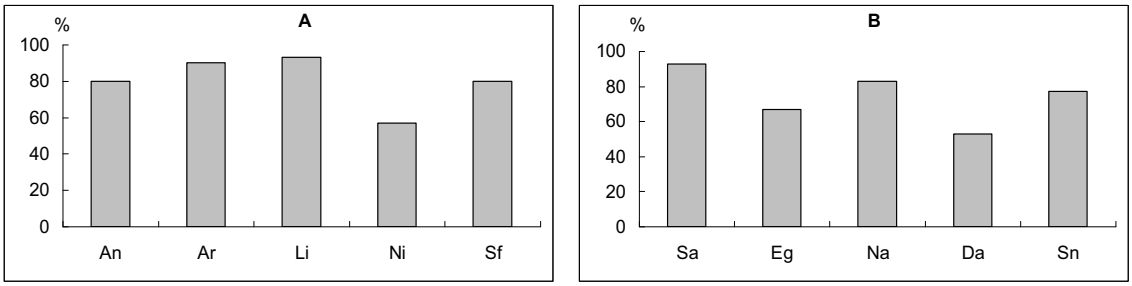

Figure 5. *The amount of children words recognized correctly by native speakers (probability 0.75), %. A - data for 6- year-old children; B - data for 7- year- old children. Horizontal axis values are % of recognized words.*

4. Discussion

We showed that the word stress is formed by 4 years of age: stressed vowel differs from an unstressed one by a longer duration which is normal in Russian. At the same time fundamental frequency values differ from stressed to unstressed vowels at 4 and 5 years of age [5] witch is unusual for Russian language. The results of current study show that pitch values are not differ in stressed versus unstressed vowel in speech of 6-year-old and 7-year-old children. The unstressed vowels triangle's shape and orientation differs from the stressed vowels triangles for all children. This might be an evidence of a strong context's influence on the unstressed vowels characteristics which changes them as compared to those of the basic allophones. The same happens in Russian adult speech [7]. So the word stress is formed completely at 6 years of age.

F1 values of /i/ and /u/ vowels are higher than in adult speech, perhaps as a result of the higher pitch values.

The area of 6-year-old children vowels triangle is more then area of 7-years-old children vowel triangle. The improvement in word's recognition by native speakers from the age of 4-5 years to the age of 7 years was not revealed. It was shown previously that 75 % (68-85%) words of 4-5-years-old children are determined by adults (probability 0.75) [8]. At the same time word's recognition by native speakers slightly decreased from children's age of 6 to the age of 7. Children's lexicon changes to more complex words and complicated utterances. Children appears not to try to articulate stressed vowels in words clearly. They are aligned mainly to pass the meaning of message by their phrases.

5. Conclusion

A stressed vowel in words pronounced by children of 6 and 7 years of age differs from an unstressed one by a longer duration which is normal in Russian.

The pitch values remain high. It is an individual characteristic of each child. There are no differences in a pitch values between stressed and unstressed vowels.

Formant structure of Russian vowels is not formed completely by the age of 7 years. F1 values of /i/ and /u/ vowels are higher than F1 values in adult speech. The environment of vowels bring contribution to the shape and orientation of vowels triangles on two-formant plot.

Native speakers determine (probability 0.75) 57-93% words of 6 – 7-years-old children.

6. Acknowledgements

The study was performed with the financial support from the Russian Fund for Humanities (projects № 08-06-90604 a/B), Russian Fund for Basic Investigations (project № 09-06-00338a).

7. References

[1] Galunov V. I., Lyakso E. E. Formation the acoustic image of signals in the early ontogenesis, XI session of the Russian acoustic Society Proceedings, Moscow, 2001, vol. 3, p. 20-24.

[2] Lyakso E. E., Petrikova N. A., Chelibanova O. V., Ostrouhov A. A., Razumihin D. V. Zvuki russkih detey pervogo goda zhizni ii h vospriyatie vzroslymi (Sounds of Russian children in the first year of life and their recognition by adults), Child speech: Psycholinguistic researches, Eds: T. N. Ushakova, N. V. Ufimceva, Moscow, PER SE, 2001, p. 65-87.

[3] Lyakso E. E., Petrikova N. A., Chelibanova O. V. Osobennosti vospriyatiya russkimi auditorami zvukov detey vtorogo goda zhizni (Features of Russian auditors perception of sounds in children of the second year of life), Physiology journal, 2003, № 4, p. 456-472.

[4] Lyakso E., Gromova A., Frolova O., Romanova O. Acoustic aspect of the formation of speech in children in the third year of life, Neuroscience and Behavioral Physiology, 2005, Vol. 35, N 6, p. 573-583.

[5] Lyakso E., Gromova A. The acoustic characteristics of Russian vowels in children of 4 and 5 years of age, Psychology of Language and Communication, 2005, Vol. 9, № 2, p. 5 - 14.

[6] Derkach M.F., Gumetskii R.Ya., Guba B.M., Chaban M.E. Dinamicheskie spektry rechevykh signalov (The dynamic spectrograms of speech signals), Lvov, 1983.

[7] Bondarko L.V. Fonetica sovremennogo russkogo yazyks: Uchebnoe posobie (The modern Russian phonetics: A tutorial), St-Petersburg, 1998.

[8] Lyakso E., Kurazova A, Gromova A., Ostrouxov A. Recognition of words and phrases of 4-5-years-old children by adults, Speech and Computer, XI International conference, SPb, 2006, p. 567 - 570.