



Acoustic Features of Speech of Typically Developing Children Aged 5–16 Years

Alexey Grigorev, Olga Frolova, and Elena Lyakso^(✉)

Saint Petersburg State University, Saint Petersburg, Russia
lyakso@gmail.com

Abstract. The study is aimed at investigating the formation of acoustic features of speech in typically developing (TD) Russian-speaking children. The purpose of the study is to describe the dynamics of the temporal and spectral characteristics of the words of 5–16 years old children depending on their gender and age. The decrease of stressed and unstressed vowels duration from child's words to the age of 13 years is revealed. Pitch values of vowels from words significantly decrease to the age of 14 years in girls and to the age of 16 years in boys. Pitch values of vowels from girls' words are higher vs. corresponding features from boys' words. Differences in the pitch values and vowel articulation index in boys and girls in different ages are shown. The obtained data on the acoustic features of the speech of TD children can be used as a normative basis in artificial intelligence systems for teaching children, for creating alternative communication systems for children with atypical development, for automatic recognition of child speech.

Keywords: Child speech · Acoustic features of speech · Pitch
Vowel articulation index

1 Introduction

The data on the acoustic features of infant's vocalizations [1, 2], speech of children during the first years of life [3, 4] are obtained for different languages. Acoustic characteristics of speech of preschoolers [5–7], junior schoolchildren, and teenagers [8–12] are less studied. The focus of research is shifted to investigating child's speech disorders [13–15].

The works of Child speech research group of Saint Petersburg State University describe the acoustic features of vocalizations and speech of Russian typically developing (TD) infants [16, 17] compared with orphans and children with neurological disorders [18, 19], the dynamics of the acoustic features of vowels in vocalizations and words of TD children [20] and twins [21] from birth to the age of 7 years. The comparative study of speech formation in TD children and children with autism spectrum disorders (ASD) is carried out [22, 23]. However, in these studies, the speech of TD children was recorded in the model situations - the repetition of words [22, 23] and under the different emotional states [24]. Information on the acoustic characteristics of calm (neutral) speech of Russian-speaking TD children in the age range of 5–16

years is absent. Data on the acoustic characteristics of child's speech is necessary in clinical practice to clarify the diagnosis, for automatic speech recognition systems [25].

The aim of the study is to determine the age dynamics of temporal and spectral characteristics of vowels in words in 5–16 years old Russian children, depending on the gender and the age of the child.

2 Method

2.1 Data Collection

240 children aged 5–16 years (10 boys and 10 girls in each age) participated in the study. According pediatricians' conclusion all children developed normally, did not have diagnosed hearing and speech disorders. All children were born and have been living in St. Petersburg city with parents who were also born in St. Petersburg or have been living there for more than 10 years. For all the children the first language (L1) was Russian. At school, children were taught the second language (L2) English.

Audio records of child speech with parallel records of child's behavior in the situation of dialogue with the experimenter were made. The standard set of experimenter's questions addressed to the child was used. The experimenter began the dialogue with the request to say your name and age. Then the experimenter consistently asked questions:

- Do you like to go to school/kindergarten?
- What do you like in school/kindergarten (classes or play with friends)?
- What are your favorite tasks? Why?
- Do you have any hobbies?
- What are your favorite movies, cartoons, books, games (computer/desktop/mobile)?
- Do you have brothers or sisters?
- Do you have pets?
- Did you visit the zoo, circus, and museum?

This set of questions allowed obtaining the child's replicas containing similar and identical words. For example every child used in replicas the words: /like (nrAvitsya) – do not like (ne nrAvitsya)/, /know (znAyu) – do not know (ne znAyu)/, /Russian (rUsskiy)/, /bored (skUchno)/, /plays (Igr)/, /tiger (tIgr)/.

The duration of the dialogues was 5–10 min. The recordings were made by the “Marantz PMD660” recorder with a “SENNHEIZER e835S” external microphone and camera “SONY HDR-CX560E”. Speech files are stored in Windows PCM format, 44100 Hz, 16 bits per sample.

All procedures were approved by the Health and Human Research Ethics Committee and signed informed consent was obtained from parents of the child participant.

2.2 Data Analysis

Spectrographic analysis of speech was carried out in the “Cool Edit Pro” sound editor (Syntril. Soft. Corp. USA). We analyzed the duration of stressed and unstressed vowels

and the stationary part of vowels; pitch values, formants frequencies (F1, F2) for the stationary part of vowels.

The acoustic features reflecting the basic physiological processes in the vocal tract during the articulation of the speech signal were chosen. Temporal features (vowels duration) are associated with the formation of speech breathing, the pitch values are the indicator of the frequency of oscillations of the vocal folds, the values of the two first formants reflect the processes occurring in the oral cavity and are acoustic keys for the identification of vowels.

Formant triangles with apexes corresponding to the vowels /a/, /u/, and /i/ in F1, F2 coordinates were plotted and their areas were compared. Vowels formant triangle areas [20] and vowel articulation index (VAI) [26] were calculated.

3 Result and Discussion

3.1 Duration of Vowels from the Words of 5–16 Years Old Children

We found dynamics of the duration of stressed and unstressed vowels in child's words. Duration of stressed vowels in girls' words significantly increases from the age of 5 years to 7 years (Mann–Whitney test), decreases during the age of 9–11 years ($p < 0.01$ Kruskal–Wallis test), and stabilizes at the age of 13–16 years ($p < 0.001$) (Fig. 1A).

The stressed vowels duration in the boys' words is significantly higher than the corresponding values of unstressed vowels duration for all ages except the age of 7 years. ($p < 0.001$ for child's age of 5 years, 8–9 years, 11 years and from 13 to 16 years; $p < 0.01$ for age of 6 years, 10 years and 12 years, Mann–Whitney test). The duration of stressed vowels in the words of boys reduces to the age of 13 years ($p < 0.05$) (Fig. 1B). The duration of unstressed vowels in the words of the girls reduces to the age of 13–16 years ($p < 0.05$) and to the age of 13 years in the words of the boys ($p < 0.05$).

Girls age correlates with the duration of stressed vowels $F(1.798) = 83.608$; $p < 0.000$ (Beta = -0.308 ; $R^2 = 0.095$) and the stationary part of stressed vowels $F(1.798) = 315.61$; $p < 0.000$ (Beta = -0.532 ; $R^2 = 0.283$); with the duration of unstressed vowels $F(1.1361) = 63.295$; $p < 0.000$ (Beta = -0.211 ; $R^2 = 0.044$) and the stationary part of unstressed vowels $F(1.1361) = 488.50$; $p < 0.000$ (Beta = -0.514 ; $R^2 = 0.264$) – Regression analysis.

Boys age correlates $F(1.760) = 119.26$; $p < 0.000$ with the duration of stressed vowels (Beta = -0.368 ; $R^2 = 0.136$) and the stationary part of stressed vowels $F(1.760) = 435.45$; $p < 0.000$ (Beta = -0.604 ; $R^2 = 0.364$).

The data on the stabilization of stressed and unstressed vowels duration to the age of 13 years may indicate the formation of speech breathing to this age.

According to the literature, speech breathing in children at the age of 7 years differs from speech breathing of adults [8]. These differences pass away by the age of 10 years, but developing some features of speech breathing (respiratory volume of lungs, sound pressure level) continues in adolescence [8]. Speech breathing features depend on the age of the informant, but not the informant's gender [27].

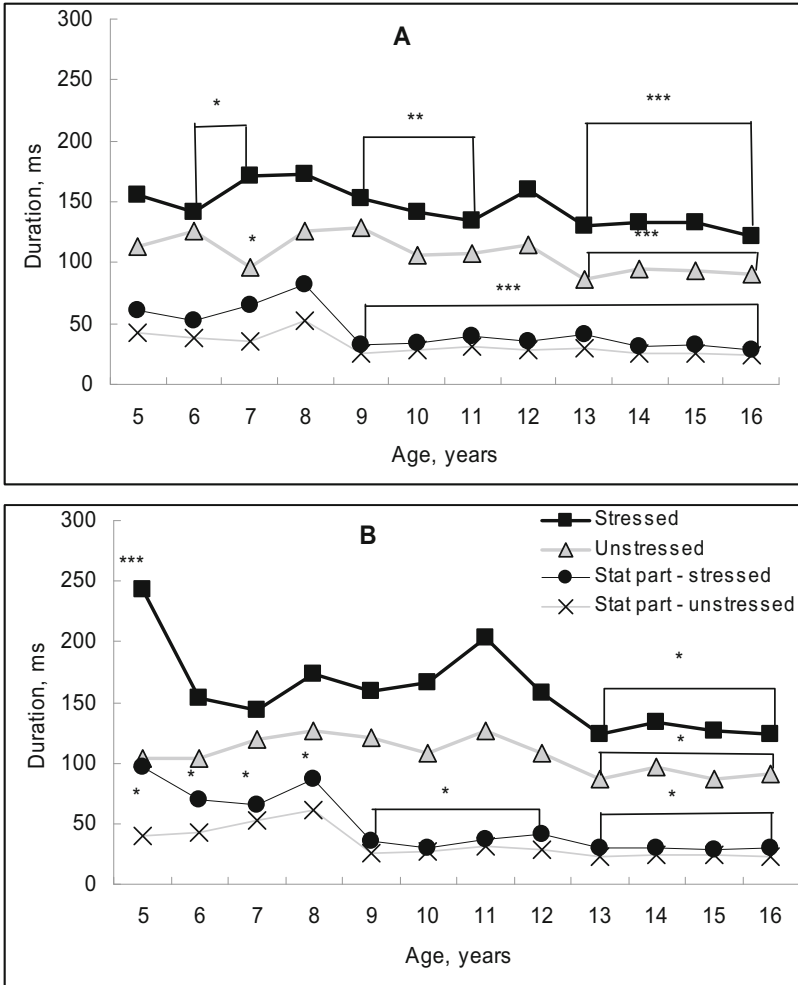


Fig. 1. The duration of stressed and unstressed vowels from the words of girls (A) and boys (B) aged 5–16 years. Square marker – stressed vowels, triangle marker - unstressed vowels, round marker – stationary part of stressed vowels, X-marker – stationary part of unstressed vowels. * - $p < 0.05$; ** - $p < 0.01$; *** - $p < 0.001$, Kruskal–Wallis test. Horizontal axis – age, years; vertical axis – duration, ms.

Our data on the stabilization of stressed and unstressed vowels duration up to the age of 13 years may point at the end of speech breathing developing at this age.

Thus, we revealed the main trends of the vowels duration. The identification of specific correlation between sex, age, and nonlinear variation in the vowels duration will be the subject of further work.

3.2 Pitch Values of Vowels from the Words of 5–16 Years Old Children

Stressed and unstressed vowels pitch values dynamics with child’s age is traced. Pitch values of stressed vowels in girls’ words reach maximum at the age of 5 years and decrease till 9 years and remain stable at the age of 9–13 years, decrease to the age of 14–16 years (Fig. 2A, Tables 1, 2).

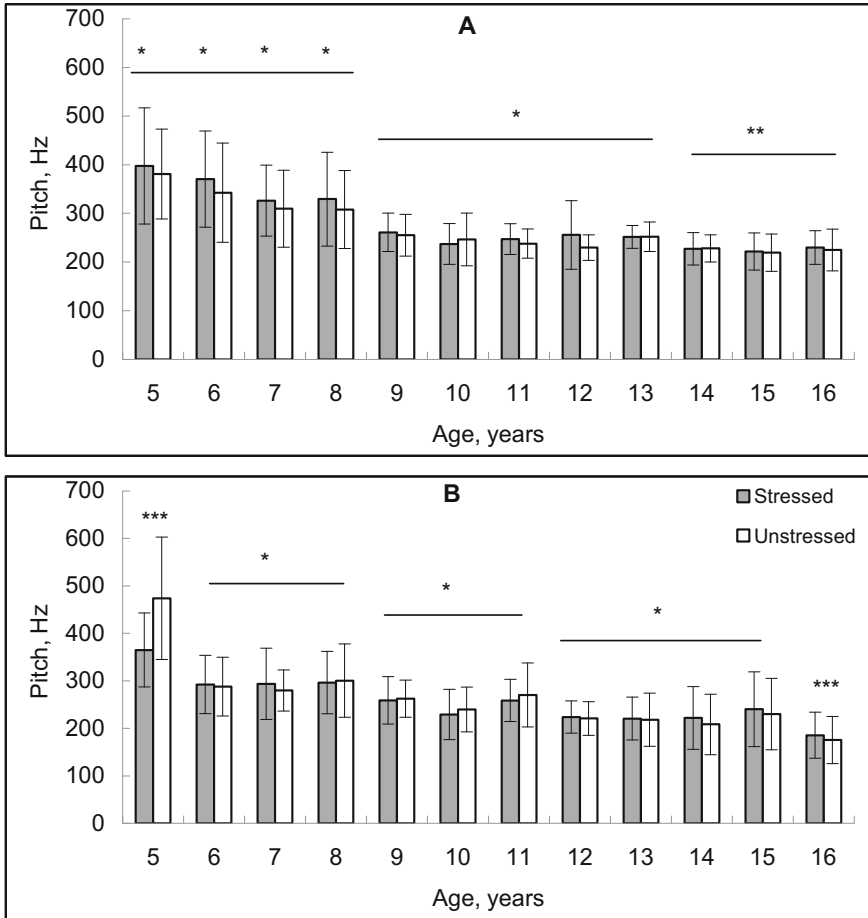


Fig. 2. Pitch values of vowels from 5–16 years olds’ words. A – girls, B – boys, * - $p < 0.05$; ** - $p < 0.01$; *** - $p < 0.001$ – Mann–Whitney test. Horizontal axis – age, years; vertical axis – pitch values, Hz.

Figure 2B shows the same data for boys. Maximal values of pitch of stressed vowels are revealed in the words of 5 years old boys, after pitch decreases to the age of 6–8 years, 9–11 years, 12 years. Minimal values of pitch are revealed at the boy’s age of 16 years. Tables 1, 2 represents accurate age and gender data of pitch values of child’s vowels in words.

Table 1. Pitch values of stressed vowels, Hz.

Age, years	Boys	Girls
5	365 ± 78	398 ± 119
6	292 ± 61	370 ± 99
7	294 ± 75	326 ± 73
8	296 ± 66	329 ± 96
9	259 ± 50	261 ± 39
10	229 ± 53	237 ± 42
11	259 ± 44	247 ± 31
12	224 ± 34	256 ± 71
13	221 ± 45	252 ± 24
14	222 ± 66	227 ± 33
15	240 ± 79	222 ± 38
16	186 ± 48	230 ± 35

Table 2. Pitch values of unstressed vowels, Hz.

Age, years	Boys	Girls
5	474 ± 129	381 ± 92
6	288 ± 62	343 ± 102
7	280 ± 43	310 ± 79
8	301 ± 77	308 ± 80
9	263 ± 39	256 ± 43
10	240 ± 47	247 ± 54
11	271 ± 67	238 ± 30
12	221 ± 35	230 ± 26
13	218 ± 56	252 ± 30
14	208 ± 64	228 ± 28
15	230 ± 75	219 ± 38
16	175 ± 50	225 ± 43

Pitch values of stressed vowels from girls' words are significantly higher vs. corresponding values of vowels from boys' words at the age of 6 years ($p < 0.001$; Mann–Whitney test), 7–8 years ($p < 0.05$), 13 years ($p < 0.01$), and 16 years ($p < 0.001$). Pitch values of unstressed vowels from girls' words are significantly higher vs. corresponding values of vowels from boys' words at the age of 6 years ($p < 0.01$), 10 years ($p < 0.05$), 13–14 years ($p < 0.001$), and 16 years ($p < 0.001$). Pitch values of unstressed vowels from boys' words are significantly higher than corresponding values of vowels from girls' words at the age of 5 years and 11 years ($p < 0.05$).

Child's gender correlates $F(6.1797) = 5.155$; $p < 0.0000$ with pitch values of stressed vowels (Wilks'Lambda 0.965; $p = 0.0000$); $F(6.2669)$; $p < 0.0000$ with pitch values of unstressed vowels (Wilks'Lambda 0.974; $p < 0.0000$) – Discriminant analysis.

Girls' age correlates with pitch values of stressed vowels $F(1.892) = 458.99$; $p < 0.000$ (Beta = -0.583 ; $R^2 = 0.34$) and pitch values of unstressed vowels $F(1.1361) = 643.32$; $p < 0.000$ (Beta = -0.566 ; $R^2 = 0.321$) – Regression analysis. Boys' age correlates with pitch values of stressed vowels $F(1.808) = 174.49$; $p < 0.000$ (Beta = -0.421 ; $R^2 = 0.178$).

The results on the decrease of pitch values with child's age correspond to the data for other languages [10, 12] and reflect the general patterns of voice formation in ontogenesis [28–32]. In our study sharp changes in the pitch values in boys at the ages of 6 years, 9 years, 12 years and 16 years were revealed. More linear decrease of pitch values with child's age was described in girls vs. boys. Age-related anatomical changes in the vocal tract, in particular, changes in its length could be used as an explanation of these data. Differences in the length of the vocal tract between boys and girls after the age of 12 years revealed by MRI data caused changes in the pitch values between girls and boys [29]. The authors concluded based on the fMRI that boys and girls have different age dynamics in the vocal tract length, causing various dynamics of pitch [32]. Two age periods with sharp decreases in the pitch values 6–8 years and 12–15 years are revealed in boys, all pitch changes in girls are more linear without sudden changes [32].

3.3 Formant Characteristics of Vowels from the Words of 5–16 Years Old Children

Figure 3 presents the formant triangles with apexes corresponding to values of the first and second formants of stressed vowels /a/, /i/, /u/ from girls (A) and boys (B) words. The first formant values of stressed vowel /a/ from girls' words are maximal at the age of 5 years and decrease to the age of 12 years ($p < 0.05$; Kruskal–Wallis test). The first formant values of stressed vowels /a/, /u/, /i/ from boys' words are maximal at the age of 5 years and decrease with child's age. The second formant values of stressed vowels from girls' and boys' words show tendency to decrease with child age.

The formant triangles of stressed vowels shift to the low-frequency region on the two-formant coordinate plane by the values of the first formant – up to the age of 14 years for girls and up to the age of 16 years for boys (Fig. 3).

VAI of stressed vowels from girls' and boys' words changes with children's age non-linearly. The end of the preschool period is characterized by the maximum values of VAI (for boys at the age of 6 years, for girls at the age of 7 years), which can be explained by preparing children for schooling and the need for active use of verbal communication in the learning process. At the age of 8 years VAI values are high and similar for boys and girls. A further decrease in VAI may be due to the increase in fluency of speech and the termination of the articulation skills mastering.

VAI of stressed vowels from girls' words is 0.75 (conventional units) at the age of 5 years and 0.82 at the age of 16 years. VAI of stressed vowels from boy's words is 0.61 at the age of 5 years and 0.87 at the age of 16 years (Fig. 4).

The values of formant triangle areas of stressed vowels from girls' words are higher vs. boys' words except the ages of 10, 11, 12 years (Fig. 5).

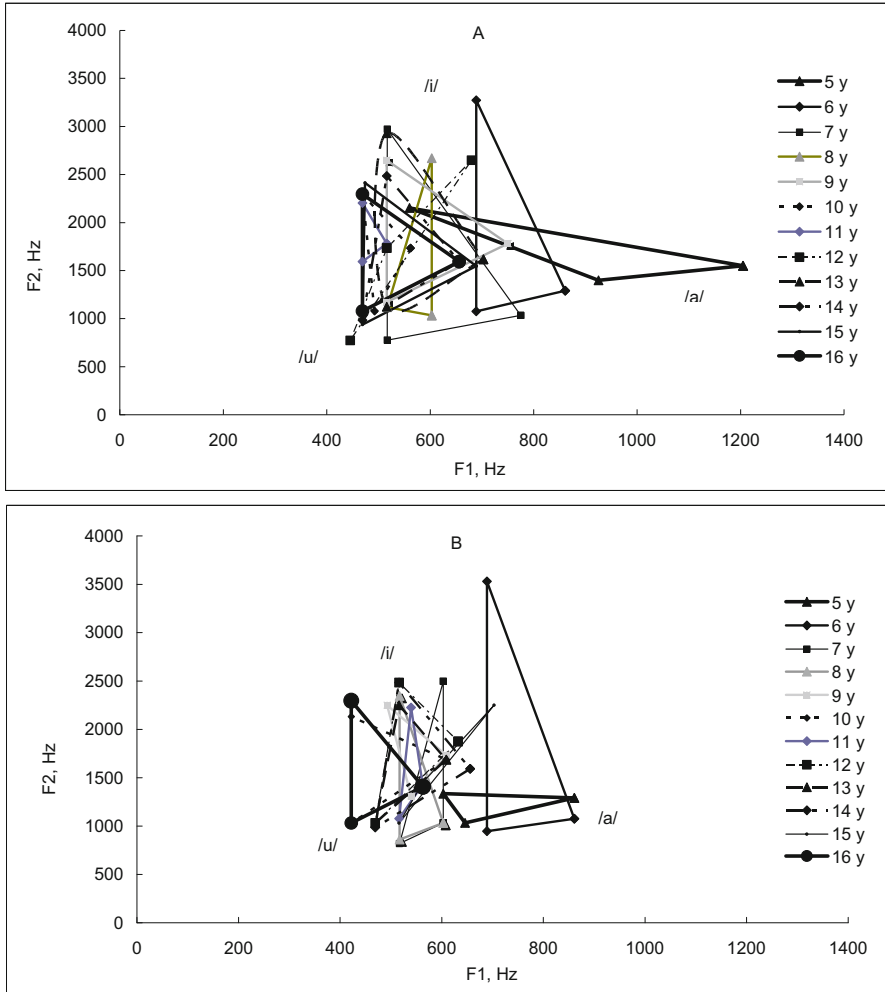


Fig. 3. The stressed vowels formant triangles with apexes /a/, /u/, /i/ from words of girls (A) and boys (B). Horizontal axis values are F1, Hz, vertical axis values are F2, Hz. Bold lines indicate the data for 5 and 16 years old children (boundaries of the analyzed age range).

Revealed differences in the values of VAI and formant triangle areas with the age of boys and girls are indirectly confirmed by the data on the material of the Chinese language about the significant decrease in the values of the first two formants in the speech of children in the age range 3–18 years [33]. The authors associate differences between boys and girls on the base of the vowels' formant frequencies with differences in the volume of the pharynx of children [33].

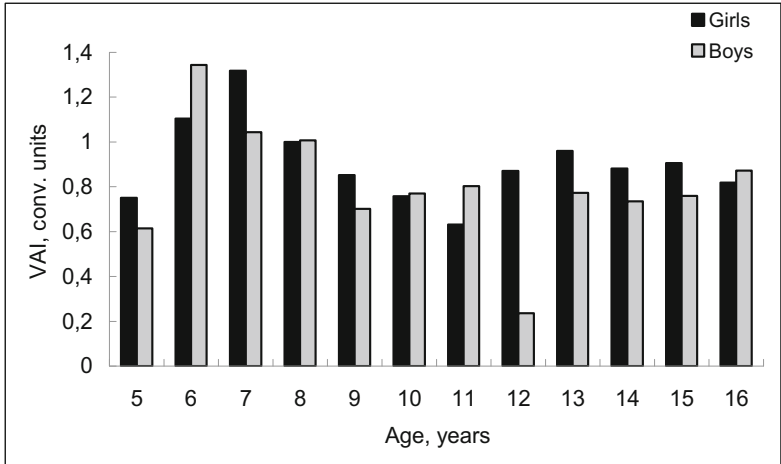


Fig. 4. Articulation index of stressed vowels. Black columns – data for girls, grey columns – data for boys. Horizontal axis – age of child, years, vertical axis - vowel articulation index values, conventional units.

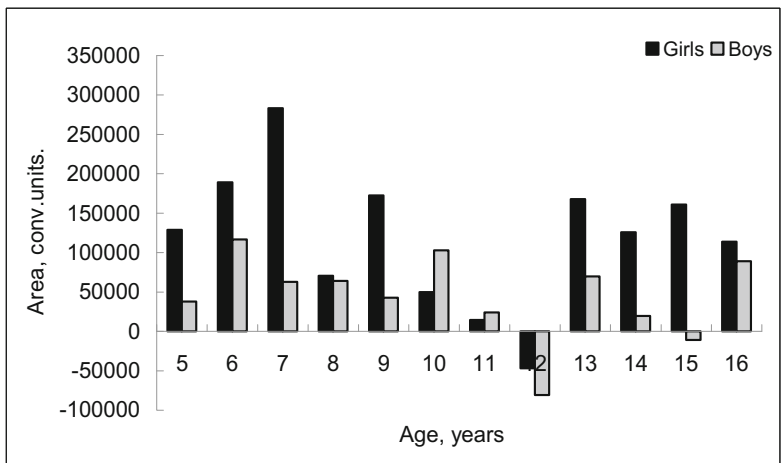


Fig. 5. Areas of vowels formant triangles. Black columns – data for girls, grey columns – data for boys. Horizontal axis – age of child, years, vertical axis - areas of vowels formant triangles values, conventional units.

3.4 Phonetic Data

The phonetic analysis revealed that children use normative phonemes for Russian language in words at the age of 8–16 years. Some differences in pronunciation related to individual features of the child and they do not affect the lexical meaning of the word. Unformed phonemes /l/, /lʲ/, /lʲʲ/, /tʲsʲ/, /r/, /ʐ/, /s/, /sʲ/ were described in 5–7 years old

children. Phonetic analysis revealed sound replacements not affecting the lexical meaning of the word: changes and replacements of phonemes /r'/ and /r/, change // to /s/, and /tS'/ to /t'/, replacements of group of consonants with one phoneme.

4 Conclusions

Age dynamics of the duration, pitch values, and spectral features of stressed and unstressed vowels from words of Russian-speaking boys and girls aged 5–16 years was described. Age dynamic of the duration of stressed and unstressed vowels and their stationary parts was shown. Duration of stressed and unstressed vowels from child's words significantly decrease to the age of 13 years in children of both genders. Significant decrease of pitch values of vowels from words with child's age was shown. The difference in pitch values between girls and boys was defined. These data reflect general patterns of the voice formation in ontogenesis, taking into account the gender of the informant. The clarity of articulation of Russian-speaking 5–16 years old children was described by the vowels articulation index.

The obtained data on the acoustic features of the speech of TD children can be used as a normative basis in artificial intelligence systems for teaching children, for creating alternative communication systems for children with atypical development, for automatic recognition of child speech [34], for teaching children the clarity of pronunciation based on the use of acoustic feedback.

Acknowledgements. This study is financially supported by Russian Science Foundation (Project № 18-18-00063).

References

1. Kuhl, P.K., Meltzoff, A.N.: Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* **100**(4), 2425–2438 (1996)
2. Nathani, S., Ertmer, D.J., Stark, R.E.: Assessing vocal development in infants and toddlers. *Clin. Linguist. Phon.* **20**(5), 351–369 (2006)
3. Lyakso, E., Gromova, A., Frolova, O., Romanova, O.: Acoustic aspect of the formation of speech in children in the third year of life. *Neurosci. Behav. Physiol.* **35**(6), 573–583 (2005)
4. Liu, S., Jin, Y., Yu, H., Yang, L.: Study on the acoustic characteristics of speech and physiological development of vocal organs for two-year-old children. In: Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control, pp. 576–579. IEEE, Qinquangdao (2015)
5. Ballard, K.J., Djaja, D., Arciuli, J., James, D.J.H., van Doorn, J.: Developmental trajectory for production of prosody: lexical stress contrastivity in children 3 to 7 years and adults. *J. Speech Lang. Hear. Res.* **55**, 1716–1735 (2012)
6. Brehm, S.B., Weinrich, B.D., Sprouse, D.S., May, S.K., Hughes, M.R.: An examination of elicitation method on fundamental frequency and repeatability of average airflow measures in children age 4:0–5:11 years. *J. Voice* **26**(6), 721–725 (2012)

7. Lyakso, E.E., Frolova, O.V., Grigorev, A.S.: Infant vocalizations at the first year of life predict speech development at 2–7 years: longitudinal study. *Child Psychol.* **5**(12), 1433–1445 (2014)
8. Hoit, J., Hixon, T., Watson, P., Morgan, W.: Speech breathing in children and adolescents. *J. Speech Hear. Res.* **33**, 51–69 (1990)
9. Vorperian, H., Kent, R.: Vowel acoustic space development in children: a synthesis of acoustic and anatomic data. *J. Speech Lang. Hear. Res.* **50**(6), 1510–1545 (2007)
10. Lee, S., Iverson, G.K.: The development of monophthongal vowels in Korean: age and sex differences. *Clin. Linguist. Phon.* **22**(7), 523–536 (2008)
11. Harries, M., Hawkins, S., Hacking, J., Hughes, I.: Changes in the male voice at puberty: vocal fold length and its relationship to the fundamental frequency of the voice. *J. Laryngol. Otol.* **112**(5), 451–454 (1998)
12. Lee, S., Potamianos, A., Narayanan, S.: Acoustics of children’s speech: developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.* **105**, 1455–1468 (1999)
13. Bonnef, Y.S., Levanon, Y., Dean-Pardo, O., Lossos, L., Adini, Y.: Abnormal speech spectrum and increased pitch variability in young autistic children. *Front. Hum. Neurosci.* **4** (237), 1–7 (2011)
14. Kent, R.D., Vorperian, H.K.: Speech impairment in down syndrome: a review. *J. Speech Lang. Hear. Res.* **56**(1), 178–210 (2013)
15. Fortunato-Tavares, T., Andrade, C.R.F., Befi-Lopes, D., Limongi, S.O., Fernandes, F.D.M., Schwartz, R.G.: Syntactic comprehension and working memory in children with specific language impairment, autism or down syndrome. *Clin. Linguist. Phon.* **29**(7), 499–522 (2015)
16. Lyakso, E.: Characteristics of infant’s vocalizations during the first year of life. *Int. J. Psychophysiol.* **30**, 150–151 (1998)
17. Lyakso, E.E.: Phonological developments of Russian children during the first postnatal year. In: Joint Conference of the IX International Congress for the Study of Child Language and the Symposium on Research in Child Language Disorders, p. 171. SRCLD-IASCL, Madison, USA (2002)
18. Lyakso, E., Frolova, O.: Russian vowels system acoustic features development in ontogenesis. In: Interspeech 2007, pp. 2309–2313. Antwerp, Belgium (2007)
19. Lyakso, E.E., Frolova, O.V., Stolyarova, E.I.: Russian child speech development in mother–child interaction: basis rules and individual features. *Int. J. Psychophysiol.* **69**(3), 306–307 (2008)
20. Lyakso, E.E., Grigor’ev, A.S.: Dynamics of the duration and frequency characteristics of vowels during the first seven years of life in children. *Neurosci. Behav. Physiol.* **45**(5), 558–567 (2015)
21. Kurazova, A.V., Lyakso, E.E.: Speech development of children and vocal-speech interaction in the triads “mother-twins”: longitudinal study. *Bull. St. Petersburg State Univ.* **3**, 93–103 (2012). (in Russian)
22. Lyakso, E., Frolova, O., Grigorev, A.: A comparison of acoustic features of speech of typically developing children and children with autism spectrum disorders. *Lect. Notes Comput. Sci.* **9811**, 43–50 (2016)
23. Lyakso, E., Frolova, O., Grigorev, A.: Perception and acoustic features of speech of children with autism spectrum disorders. *Lect. Notes Comput. Sci.* **10458**, 602–612 (2017)
24. Lyakso, E.E., Frolova, O.V., Grigor’ev, A.S., Sokolova, V.D., Yarotskaya, K.A.: Recognition by adults of emotional state in typically developing children and children with autism spectrum disorders. *Neurosci. Behav. Physiol.* **47**(9), 1051–1059 (2017)
25. Fernando, S., et al.: Automatic recognition of child speech for robotic applications in noisy environments. *Comput. Speech Lang.* 1–28 (2016). [arXiv:1611.02695v1](https://arxiv.org/abs/1611.02695v1) [cs.CL]

26. Roy, N., Nissen, S.L., Dromey, C., Sapir, S.: Articulatory changes in muscle tension dysphonia: evidence of vowel space expansion following manual circumlaryngeal therapy. *J. Commun. Disord.* **42**(2), 124–135 (2009)
27. Boliek, C.A., Hixon, T.J., Watson, P.J., Jones, P.B.: Refinement of speech breathing in healthy 4- to 6-year-old children. *J. Speech Lang. Hear. Res.* **52**, 990–1007 (2009)
28. Hacki, T., Heitmuller, S.: Development of the child's voice: premutation, mutation. *Int. J. Pediatr. Otorhinolaryngol.* **49**(1), 141–144 (1999)
29. Fitch, W.T., Giedd, J.: Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J. Acoust. Soc. Am.* **106**, 1511–1522 (1999)
30. Hollien, H., Green, R., Massey, K.: Longitudinal research on adolescent voice change in males. *J. Acoust. Soc. Am.* **96**, 2646–2654 (1994)
31. Perry, T.L., Ohde, R.N., Ashmead, D.H.: The acoustic bases for gender identification from children's voices. *J. Acoust. Soc. Am.* **109**, 2988–2998 (2001)
32. Markova, D., et al.: Age- and sex-related variations in vocal-tract morphology and voice acoustic during adolescence. *Horm. Behav.* **81**, 84–96 (2016)
33. Wan, P., Huang, Z., Zheng, Q.: Acoustic elementary research on voice resonance of Chinese population. *Chin. J. Surg.* **24**(6), 250–252 (2010)
34. Kaya, H., Salah, A.A., Karpov, A., Frolova, O., Grigorev, A., Lyakso, E.: Emotion, age, and gender classification in children's speech by humans and machines. *Comput. Speech Lang.* **46**, 268–283 (2017)