Eötvös Loránd University

Faculty of Humanities

THESES OF DOCTORAL DISSERTATION

VALÉRIA KREPSZ

ACOUSTIC-PHONETIC PROPERTIES OF SPONTANEOUS SPEECH IN LANGUAGES ACQUISITION

Doctoral School of Linguistics

Head of Doctoral School: Prof. Dr. Tolcsvai Nagy Gábor, Member of the Academy

Applied Linguistics PhD Program

Head of Doctoral Program: Prof. Dr. Gósy Mária DSc

Members of the Assessment Committee

Chairperson of the Committee:	Prof. Dr. Keszler Borbála DSc
Opponents:	
	Dr. Adamikné Jászó Anna DSc
	Dr. Horváth Viktória PhD
Secretary of the Committee:	
	Dr. Bóna Judit PhD
Further members:	
	Dr. Bartha Krisztina PhD
	Dr. Dér Csilla Ilona PhD
	Dr. Antalné Szabó Ágnes, PhD

Supervisor

Prof. Dr. Gósy Mária DSc

Budapest, 2018

1. INTRODUCTION

Speech organs execute finely tuned movements during articulating speech. At the beginning of speech development, children carry out large and less precise articulatory movements during the speech act which are different from the adult's aim-configuration and aim-underachievement is often seen in their speech– due to their inexperience of motor-execution (Temple et al. 2002).

The acoustic characteristics of speech are defined simultaneously by the processing of the speech organs, their size and shape. Forming different types of vowel qualities, the acoustic structure of the speech sounds are modified by changing the shape and the size of the mouth cavity, since their own frequencies are changing, too (Gósy 2004). The same connection can be seen from another aspect: the acoustic characteristic of certain vowel sounds can be examined indirectly by the acoustic structure realized in the articulation. Further, modifications can be expected due to the physiological changes happening with aging.

From anatomic viewpoint, speech organs change in size and in functioning with aging (Vorperian et al. 2005). In mother tongue acquisition, the acoustic parameters of speech are influenced by physiological characteristics, mainly by the brain development, the growth of the lung capacity, the change in the shape and size of the larynx and the vocal tract; furthermore, the resilience and change of the proportion of the tissues, the modification of the shape, and the proportion of the size and the motility of the tongue (Kent–Vorperian 2005). The extent of the growth is not balanced in the period of first language development: faster and slower periods follow one another in the anatomic development which may induce significant temporary changes (Vorperian et al. 2005). The development in the speech of children is based partly on the macro-anatomic development and maturity which is mainly characterized by the length of the articulatory tract (Vorperian et al. 2009), although the development of the vocal tract – as in the cases of all other speech areas – is a very complex process.

As children grow, their speech gradually becomes similar to the adult samples in articulation as well as in its acoustic structure as a result of physiological changes (along with other factors, such as the age, gender characteristics and individual differences). Due to the maturation of the articular-motor and orofacial systems, the dynamic coordination of the speech organs becomes more consistent, therefore, the variability of the articular gestures decreases (Terband et al. 2011).

The process from cooing to speech is very complex. In mother tongue acquisition, continuous development can be seen in quality and in quantity at different levels of the language, for example, in phonetics, in phonology, morphology, syntax and semantic levels. From cooing to the complete ownership over the language, the stages (gradually and saccadic developing stages) gradually shift from one into another, although their timing may be different in each individuals to some extent (Gósy 2005; Lukács 2006).

Hungarian and international research results clearly support the statement that children over the age of 3 acquired the vowel set of the mother tongue, therefore, certain articulatory stability can be expected in their production, so the considerable changes may already be shown in speech production. In the same time, Hungarian research also disclosed that distinguishing on the bases of the phonological length can be expected only later on the areas of both production and perception (Gósy 2006; Bóna–Imre 2010).

Studies of the temporal and spectral structures of the vowels of Hungarian children appeared in the literature sporadically. These studies mostly examined the linguistic characteristics of younger or older children or they analysed smaller samples. Further, these examinations dealt separately with the processes of the production and the perception.

As a result of the above-mentioned factors, the aim of the research is to investigate the speech production and the perception parallel, and administer an overall analysis using a standardized speech perception test. Further, the temporal and spectral characteristics of the vowels appearing in the spontaneous speech between the age groups 4 and 6 were also examined.(as a cross-sectional study, comparing the results to the adult.

2. THE STRUCTURE OF THE DISSERTATION

The first chapter is the Introduction, which first introduces the process of speech formation, with special emphasis on the formation of vowels. Subsequently the theories and stages of mother tongue acquisition are reviewed. The end of the chapter is a summary of national and international research results on the consequences of anatomic developmental level as well as physiological changes of speech organs.

Chapter 2 presents the research material, the methodology of the experiment and the characteristics of the individuals involved in the experiment.

Chapter 3, titled *Results,* is broken into six sections because of its length. The first section discusses the result of the speech perception experiment, sections 2-5 discuss the speech production examination. (Section 2 discusses various aspects of the time- and formant structure of 4-year-olds, section 3 the same for 5-year-olds, section 4 for 6-year-olds. Section

5 presents the comparison of results in the three age groups, and the group of adult speakers.) Section 6 is a new chapter, which compares the two aspects of the research (production and perception). Chapter 7 contains the conclusions and outlook.

3. RESEARCH MATERIAL, METHOD AND PARTICIPANTS

A total of 60 children were involved in the experiment from three age groups: 4, 5 and 6-year-olds. In every age group 10 boys and 10 girls were examined. In the production examination, recordings of the children's speech were compared with the recordings of adults' spontaneous speech, selected from the BEA database (Gósy et al. 2012). The gender ratio was the same for adults: recordings of 10 women and 10 men were chosen.

During the **perception** examinations, 9 subtests of GMP diagnostic tests (GMP 2, 3, 4, 5, 8, 9, 10, 12, 16) were administered to every child, the results were analysed within the age groups, and the results of the different age groups were also compared with each other. The mistakes made in each subtests were also analysed based on their quantity, frequency and type. The results of the tasks testing different areas of speech perception were also analysed. For statistical analysis, we performed 1. a correlation analysis, looking for connections between the results of the different subtests; 2. an analysis of the differences in performance between age groups using a general mixed model. Special attention was paid to individual differences during the analysis.

For the examination of **speech production**, recordings of spontaneous speech were used. For every recording, a sample of a 2-minute-speech (not counting breaks) was analysed. The sections of the recordings were manually annotated, then vowels were marked by using the Praat software (Boersma–Weenink 2014). With a script specifically created for this purpose, we automatically extracted the duration and the first two formant values of the vowels. Data were analysed using a general linear mixed model. Each vowel was marked in the recordings, their temporal and spectral structures, and the articulation tempo of the speakers were analysed. In the case of the temporal characteristics, the duration of vowels was examined according to 1. their position, 2. the number of syllables in the word, 3. the gender of the speaker and 4. the formation characteristics. During the examination of the spectral structure, the analysis was administered based on 1. the fundamental frequency of the speaker and the F1 and F2 values of vowels, 2. the vertical and 3. horizontal movements of the tongue, 4. lip function, 5. phonological length and 6. the gender of the speaker. In addition, the formation of the acoustic vowel space was examined in the different age groups and if the fundamental frequency showed difference, then the gender of the speakers was also considered.

4. RESULTS

4.1 Examination of speech perception

The examination of speech perception supports the hypothesized statement based on professional literature: age 4-6 is a highly sensitive period for speech development. The results showed significant development from age 4 to 5, and a more moderate development from age 5 to 6. Besides the group tendencies, every age group showed significant individual differences. Based on the examination of the different age groups, it was revealed that the sub-processes were becoming increasingly independent, therefore, the speech perception development is clearly verifiable in the examined kindergarten age.

4.2 Examination of speech production

4. 2. 1 The temporal characteristics of vowels in the speech of 4-year-old children

Based on the statistical analysis, the duration of vowels in the speech of 4-year-old children depends the most on the number of syllables, then on their position in the word, and then on vowel quality. The gender of the speaker and the interaction of the different factors did not play a significant role.

There were significant differences in durations based on the grammatical structure, and there was statistically significant difference in 4 of the 6 long-short vowel pairs. Based on position, the longest durations were measured with one-syllable words, the second longest with phrase final syllables, and next with not phrase final syllables. The difference was significant between the three groups even with only slightly differing average values. Vowel durations gradually decreased as the number of syllables increased, which proved the existence of an equalization tendency. The vowels used in the longest words examined (6 syllables) were nearly 40 milliseconds shorter than the vowels in one-syllable words. There was also significant difference in the realization of tendencies based on vowel quality. No statistically significant difference was detected between genders, although most vowels had longer average duration times among girls. As for speech formation, only the tongue's vertical position had a statistically provable influence among 4-year-olds, while the rounding of lips and the horizontal movement of the tongue showed only tendential differences in the duration of vowels.

The values of the fundamental frequency in this age group showed statistically significant differences between genders: girls had significantly higher frequency values than boys. A significant difference could be detected in the first formant of vocals depending on the openness, even though there were overlaps among certain vowel ranges. There was also a

statistically significant difference in the second formant values of palatal and velar as well as labial and illabial vowels. The spectral comparison of long-short vowel pairs presented a more complex picture. The smallest overlap was detected in the *a-á* and *e-é* pairs, which are not only different in length. The formant structure of *i-í* and *ö-ő* pairs showed a bigger overlap than that of *o-ó* and *u-ú* pairs. We did not examine the vowel *ű* in this age group. Gender comparison did not show significant difference between boys and girls neither regarding F1, nor F2.

4. 2. 2 The temporal characteristics of vowels in the speech of 5-year-old children

According to the statistical model built on the basis of the duration of the speech sounds, duration were mostly defined by the number of syllables, then their position, and then vowel quality. Values were also significant looking at the interaction of the number of syllables and position, as well as gender and position. Other factors, or the interactions of factors played no significant role.

The longest duration values were detected in the spontaneous speech of 5-year-old children. The duration differences in long-short pairs of vowels — with the exception of *i-í* (*ü-ű* was not examined) – could be statistically verified.

Regarding position, significant difference was detected in the duration of vowels in one-word phrases, phrase final position and non-final position in the phrase. The interaction of position and vowel quality also proved significant, which shows that the duration of different vowels was different depending on their position. The number of syllables also significantly influenced the duration, despite the fact that the average duration of vowels in 1-4-syllable words showed only a smaller, 15 ms difference. Although there was no significant difference between genders, 12 out of the 14 examined vowels had a longer average duration among girls. As for speech formation, only the tongue's vertical position had a statistically provable influence among the 5-year-olds, while the rounding of lips and the frontal position of the tongue showed only tendential differences.

In the 5-year-old group, there was no significant difference detected between genders regarding the fundamental frequency. Both F1 and F2 values were higher than among 4-year-olds, and despite significant overlaps, vowel qualities could be statistically differentiated. The higher the vertical position of the tongue, the lower the first formant values became. F2 values were significantly higher for palatal vowels than for velar ones, and for illabial vowels than for labial ones. For long-short vowel pairs, there was a significant overlap for *i-í* and *ö-ő* pairs, a smaller overlap for *o-ó* and *u-ú* pairs, while *a-á* and *e-é* pairs were almost completely distinct.

4.2. 3 The temporal characteristics of vowels in the speech of 6-year-old children

The duration of vowels in the age group of 6-year-old children was statistically verified to be influenced by the number of syllables, position and vowel quality; while gender and the interaction of different factors did not have significant effect on the temporal characteristics. As for speech formation, the tongue's vertical position was a determining factor, while the rounding of lips and the frontal position of the tongue were not.

The durations of vowels in different positions were significantly different, we detected the shortest duration in this age group. As for vowel quality, the only case where the above mentioned tendency was not observed was with the vowel sound *o*. An equalization tendency could be detected as the number of syllables increased. For the four most common vowels, the above-mentioned tendency presented itself regardless of the speech sound quality. Only tendential differences could be observed between genders. As for the vertical position of the tongue, the longest durations were detected in the lowest position, with a gradual decrease of duration from the highest to the lowest position. The rounding of lips and the frontal position of the tongue had neither statistically significant nor tendential influence.

The average fundamental frequency of the 6-year-olds was 251 Hz, with a significant difference between boys and girls. The tendency of decreased F1 values with higher vertical positions of the tongue was also observed in this age group. The tendency based on formation characteristics was similarly observed here as in the other age groups: (1) the more frontal a vowel is formed, the higher the (average) F2 value is, and (2) labial vowels had a lower second formant value than illabial ones. Between genders, the statistical analysis did not reveal significant differences for the first formant while the second formant values were significantly lower for boys.

4.2.4 Correlation of speech production and speech perception

In the different age groups of children, speech perception showed strong correlation only with the duration of vowels. There was a weaker connection with articulation tempo and the distance of vowel formants relative to vowel space but no connection with other factors.

4.2.5 Summary

A gradual increase of the articulation tempo could be observed with aging, although the average difference was small. There was significant difference between the age groups in the duration of vowels, the longest duration being measured at age 5, second longest at age 4 and the shortest at age 6.

Phonemically long and short vowels were significantly distinct in all three age groups, but the differences in duration were also influenced by the age of the speaker and the vowel quality. Regarding other formation characteristics, the vertical position of the tongue proved to be an influencing factor for duration in the speech of both children and adults. Among adults – similarly to children – lip function did not make a difference, while the frontal position of the tongue – contrary to children – did.

The position of the vowels proved to be a determining factor in all three age groups. The longest durations were measured in one-syllable words. The phenomenon of phrase final lengthening was verified: phrase final vowels had significantly longer durations than vowels in any other position.

The tendency of decreasing vowel duration as the influence of the growing number of syllables (the equalization tendency) was observed to a different extent in the groups of children than among the adults. The length of the word was a statistically determining factor, but the degree of the reduction in duration was different for the different age groups.

The duration values showed only tendential differences in the cases of boys and girls: the average duration was longer among girls, but the difference did not prove to be statistically significant in any age group.

Among children, the fundamental frequency showed a small decrease with age, with significant difference found between genders among age groups 4 and 6 as well as among adults.

The size of vowel space grew from age 4 to 5, then decreased from age 5 to 6, at the same time becoming more „loaded", namely, more vowels were formed in the same area – this means that a certain centralisation tendency could be observed. The vowel space of adults was nearly two-thirds of the vowel space of the 4-year-olds.

The vertical position of the tongue, that is, the openness of vowels, had a statistically proven influence on the first formant values in all three age groups of children. The tendencies presented in professional literature, namely: 1. the F2 values of palatal vowels are higher than those of the velar ones, and 2. the F2 values of illabial vowels are higher than those of the labial ones, were verified in all three age groups of children as well as in the adult group: a significant difference was detected.

5. CONCLUSIONS

The aim of the present research was twofold: on the one hand, to examine the speech perception of 4-6-year-old children and on the other hand, to examine the production of the vowels in their spontaneous speech based on the temporal and spectral structure, compared

them with those of the adult speakers. The main questions to which the research was looking for answers were: (1) Can changes be measured in the articulation tempo of speakers in the examined age groups, (2) Can changes be measured in the duration and formant structure of vowels in the examined age groups, (3) Does the gender of the speaker have an effect on the duration and formant structure of vowels, (4) Do the formation characteristics of the vowels determine their temporal and spectral structure in the case of children – as they do in the case of adult speakers?

Based on the results, we can formulate the following theses:

1. The rate of the perception errors has decreased between the age of 4 and 6, the number of the sound errors decreased, while the number of the word errors increased, the different types of perceptual processes became independent from each other.
2. The articulation tempo of the 4-6-year-old children increased with age, however, individual differences played a significant role.
3. The duration of speech sounds gradually decreases with age, however, it is also significantly influenced by the position, the number of syllables and the quality of the vowel.
4. The phonemically long and short vowels were significantly distinct on the basis of their temporal values and the articulation of the short vowels are realized more firmly, while their long vowel pairs show greater variation.
5. The value of f_0 decreased regardless of gender, differences between genders were detected in the groups of 4 and 6-year-old children and adults, but not in the 5-year-old age group.
6. The size of the acoustic vowel space gradually decreased with age.
7. As for speech formation, only the vertical movement of the tongue had an effect on the temporal characteristics.
8. Formant values showed an increase from age 4 to 5, and a decrease from age 5 to 6.
9. Formant structure was influenced by both the vertical and horizontal movement of the tongue and the lip function.

The slower articulation tempo and the gradually decreasing duration of vowels in cases of younger children can be explained by the slower movement of their speech organs, a lower level of speech routine, less practice and lower level of the automatization of the motor functions, and lower memory capacity. It is known that due to the smaller size but at the same time, the larger proportion of the speech organs of children make the aim configuration based

on adult patterns more difficult and less purposeful, also because of the lower level of motor experience.

As for the differences in the fundamental frequency, no exact anatomical measurements were performed during the examination (due to the lack of availability of the tools), thus it can only be assumed that physiological attributes are also responsible for the differences between speakers according to their gender. It is also important to note, as professional literature also points it out, that girls and boys develop at a different pace (see Whiteside 2001; Tóth 2017), which can also contribute to the differences between the two groups.

The proportion of the „loaded" part of the vowel space gradually increased. This means that children made increasingly better in using the articulation space available to them. However, they still do not reach the peripheral areas but rather stay centralized.

Based on data obtained from professional literature we also know that coordination of speech organs does not develop at an even pace, jaw movements mature sooner than lip movements, and tongue movements come last. The reason for this is that precise control of tongue movements is a very complex function, and happening mostly inside the oral cavity, it is extremely difficult to imitate the adult pattern (Markó et al. 2016). This may explain why the formation characteristics influence the duration and the formant values of the vowels.

6. REFERENCES

Boersma, P. – Weenink, D. 2014. Praat: doing phonetics by computer. http://www.fon.hum.uva.nul/ (Downloaded: 2018. 03. 15.)

Bóna J. – Imre A. 2010. A rövid-hosszú magánhangzók óvodás és kisiskolás gyermekek beszédprodukciójában. In Navracsics J. (szerk.): Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I. Segédkönyvek a nyelvészet tanulmányozásához 107. Tinta Könyvkiadó, Budapest. 49–56.

Gósy M. 2004. Fonetika, a beszéd tudománya. Osiris Kiadó, Budapest.

Gósy M. 2005. Pszicholingvisztika. Osiris Kiadó, Budapest.

Gósy M. 2006. A beszédhangok megkülönböztetésének fejlődése. Beszédkutatás 2006. 147–159.

Gósy M. 2012. Multifunkcionális beszélt nyelvi adatbázis – BEA. Általános Nyelvészeti Tanulmányok XXIV. 329–349.

Kent, R. D. – Vorperian, H. K. 1995. Anatomic development of the craniofacial-oral-laryngeal systems: A review. Journal of Medical Speech-Language Pathology 3:145–190.

Lukács Á. 2006. Nyelvelsajátítás. In: Kovács I. – Szamarasz V. Z. (szerk.) Látás, nyelv, emlékezet. Typotex Kiadó, Budapest. 95–111.

Temple, E. C. – Hutchinson, I. – Laing, D. G. – Jinks, A. L. 2002. Taste development: differential growth rates of tongue regions in humans. Developmental brain research 135(1): 65–70.

Terband H. – Maassen B. – van Lieshout P. – Nijland L. 2011. Stability and composition of functional synergies for speech movements in children with developmental speech disorders. Journal of Communication Disorders 44(1):59–74.

Vorperian, H. K. – Kent, R. D. – Lindstrom M. J. – Gentry L. R. – Yandell B. S. 2005. Development of vocal tract length during early childhood: a magnetic resonance imaging study. Journal of Acoustical Society of America 117(1): 338–350.

Vorperian, H. K. – Wang, S. – Chung, Moo K. – Schimek, E. M. – Durtschi, R. B. – Kent, R. D. 2009. Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. Journal of Acoustical Society of America 125(3):1666–1678.