doi:10.5128/ERYa19.06

EESTI RAKENDUSLINGVISTIKA ÜHINGU AASTARAAMAT 19, 93–110

CHILDREN'S SENSITIVITY TO VISUAL INFORMATION: METHODOLOGICAL CONSIDERATIONS FOR NON-WORD REPETITION TESTING

Eglė Krivickaitė-Leišienė, Ineta Dabašinskienė

Abstract. This study introduces a Lithuanian non-word repetition test and explores methodological issues by comparing two stimuli presentation conditions (live mode vs. audio mode). The study aims to discuss methodological issues of the non-word repetition test and its potential consequences for the results. The study sample comprises 100 typically developing children (2 age groups: 4;00–4;11 and 5;00–5;11). The inquiry assessed the impact of presentation mode on the accuracy of the task performance, focusing on the effect of age.

The comparison of the different stimuli of the non-word repetition test indicates that the mode of presentation impacts the accuracy of the test results. The study showed that observing the target's visual articulation helped children identify the non-words: performance in the live presentation mode was more accurate than with audio-recorded stimuli, particularly by the group of older children. Regardless of the stimuli and age, an overall comparative analysis confirmed the tendency for non-word repetition accuracy to decline in longer and structurally more complex words.

Keywords: first language acquisition, non-word repetition test, speech perception, experiment, Lithuanian

1. Introduction

Language plays a significant role in human communication, but not absolute. Body language, facial expressions, and other physical expressions are all important and can convey an additional message to the interlocutor. More specifically, in face-to-face conversations, people are sensitive not only to acoustic cues in the speech signal but also to the visual cues present in a speaker's lip movements. Watching a speaker's facial movements can enhance the listener's ability to comprehend words,

especially in noisy environments (Hidalgo-Barnes, Massaro 2007, Okada, Hickok 2009). Using visual cues on the speaker's face to improve speech perception occurs automatically and implicitly, even when the auditory input is not impoverished. In addition to using audition for hearing another person's speech, visual cues such as lip, face, and body movements contribute to the audibility of speech and are integral to speech perception (Iarocci et al. 2010).

Speech is generally considered within the auditory domain, while visual information is also presumed to be very important for speech perception. Thus, speech is a multimodal phenomenon in which the speaker's articulatory movements produce correlated information in vision (i.e. lip movements) and audition (linguistic sounds). Indeed, the brain integrates both sources of information in order to decode the spoken message (Navarra, Soto-Faraco 2007: 4).

Sensitivity to the multimodality of speech develops early in infancy. Language acquisition during childhood is not a straightforward auditory-only process but rather a complex process influenced by non-auditory sources of information such as visual speech (Erdener, Burnham 2013).

Different methods exist for assessing children's phonological processing skills; however, non-word repetition tasks are known to be efficient and reliable. Moreover, such tasks are sensitive to individual variation in listening comprehension and decoding (Archibald 2008, Schwob et al. 2021).

The non-word repetition test is essential in monitoring child language development. It is an experimental method when the respondent is asked to repeat non-words[1]. In order to be able to repeat a word which is heard for the first time and does not have any meaning, linguistic-cognitive abilities are necessary: phonological processing, short-term memory, articulation abilities, etc. (Rispens, Parigger 2010). Each word the child heard for the first time some time ago sounded unusual and strange, similar to the words in this test (Chiat, Roy 2007).

The non-words can be presented in different ways: by the experimenter's live voice (with the visual cue) or by audio recording (without the visual cue). The choice of stimuli also depends on the children's age. Audio-recorded stimuli are especially appropriate for participants who are school-age and older. In contrast, live stimuli are more suitable for very young children (2–3 years old) and clinically referred children with poor attention; this population is more likely to produce responses if the experimenter actively engages with them. The decision on how to present non-words may depend on the aim of the research (Dollaghan, Campbell 1998, Chiat, Roy 2007).

The live presentation is more flexible and may increase engagement through greater interaction opportunities. However, it does mean that the presentation will be less consistent and accurate, and there is a greater likelihood of production errors by the experimenter (Polišenská, Kapalková 2014). Live stimuli can be presented in several ways. In the first type, the experimenter covers her mouth to avoid lip reading and eliminate all visual cues (Conti-Ramsden 2001, Botting 2001, Radeborg et al. 2006). Children are told that the experimenter will say some 'made-up words' and are asked to copy them exactly. The experimenter also explains that she will be covering her mouth with a piece of paper or her hand, e.g., "I am going to hold my hand in front of my mouth when I say the words. This is so you won't be looking

[1] Non-words are phonological sequences of sounds which correspond to the phonotactic rules of a specific language but do not have any meaning or function in a sentence.

at my mouth but only listen to what I say" (Raderborg et al. 2006: 188). In the second type, experimenters do not cover their mouths (see Chiat, Roy 2007, Hoff et al. 2008), and the procedure is the same as in the first type. Studies on 2-year-old children have used live presentation rather than audio-recorded stimuli, believing that live presentation will lead to increased engagement and, therefore, higher completion rates (Polišenská, Kapalková 2014). The third type – a non-word repetition test – is the most employed in relevant studies, especially with older children, when audio-recorded stimuli are used to eliminate visual cues.

One of the key reasons to use audio-recorded stimuli only is to avoid any inconsistency and external influence, as well as eliminate any visual cues. Furthermore, audio-recorded stimuli ensure uniformity of input, eliminating variations in rate, pitch, volume, and other phonetic and auditory features of the input that may occur when the experimenter delivers the stimuli that may enhance or depress children's performance. This way, the non-word repetition test might be presented in a game form to attract children. Puppets, animals, or aliens present non-words: the children are shown a picture of an alien/puppet/animal and told that it only speaks a foreign language, and they are asked to copy some words that the character is going to say (Chiat, Roy 2007, Kapalkova et al. 2013).

Results of the Lithuanian non-word repetition test have been extensively described (see Krivickaitė-Leišienė, Dabašinskienė 2022, Krivickaitė-Leišienė 2020, Krivickaitė 2017, 2016, 2014, Krivickaitė, Dabašinskienė 2013), but the comparison of the stimuli was not discussed in detail. The samples of the studies on the Lithuanian non-word repetition test were mainly collected using audio-recorded stimuli due to the large amount of data. However, for this study, we have collected additional data using live presentation as our primary aim was to compare the effect of different conditions (live mode vs. audio mode) on the accuracy of the performance. We followed the suggestion made by Chiat and Roy (2007: 432), which emphasizes that "comparison of live versus audio-recorded stimuli is clearly needed for evaluating the possible effects of such variations on rates and accuracy of response as well as the comparability and validity of the two methods".

2. Theoretical background: the influence of visual cues on language acquisition

It is known that babies and young children learn the meaning of communicative signals by looking at adults' faces. Babies and young children see and hear communicative signals and learn to attach meanings to them through everyday interactions with their parents or caregivers (Weikum et al. 2007).

Studies show that from just two months of age, infants respond to the congruence between auditory and visual speech tokens (Knowland et al. 2016). At 4–8 months, infants show a heightened eye gaze pattern to speakers' mouths; at six months, they may use visual cues to help establish phonemic categories (Teinonen et al. 2008, Lewkowicz, Hansen-Tift 2012). The onset of lip-reading at this age corresponds with the onset of canonical babbling, suggesting that babies begin lip-reading because they become interested in speech and language (Lewkowicz et al. 2012, Pons et al. 2015).

Studies with older children show controversial results. In a study with children aged 5 to 6 years, visual speech stimuli had no effect in a sentence perception task, and this age group failed to use visual cues (Wightman et al. 2006). The opposite result came for single-word perception, as Erdener et al. (2010) found a benefit of visual cues for English-speaking 5-, 6-, 7- and 8-year-olds and adults. The results showed that the degree of visual influence and lip-reading ability increases from childhood (5, 6, 7, 8 years) to adulthood, lending support. Ross with colleagues (2011) tested 5–14-year-olds and adults (16–56 years). The results suggest that improvement in the ability to recognize speech-in-noise and audiovisual integration during speech perception continues quite late into the childhood years.

The McGurk effect is a classic and compelling demonstration of the visual signal's influence on auditory speech (McGurk, MacDonald 1976). The effect occurs when two stimuli – auditory and visual – are presented simultaneously to participants. The presented stimuli do not match, and results show that most listeners report hearing an illusory syllable /da/ (see Desai et al. 2008), etc. All the studies concluded that auditory–visual speech perception is already well-developed during the early months of infancy.

Research on the McGurk effect tasks revealed that language impaired (LI) 5-11-year-old children were less accurate than their typical language development (TD) peers but were able to make equivalent use of visual cues to boost performance accuracy as their TD peers (Knowland et al. 2016). Another study found that responses by the children with LI (4;0–5;8 years old) indicated less impact of visual processing on speech perception than was seen with their TD peers. These results demonstrate that LI children extend beyond the auditory-only modality and include auditory-visual processing as well (Norrix et al. 2007).

Visual influence on speech perception is broadly described by Massaro and colleagues (e.g., Hidalgo-Barnes, Massaro 2007, Massaro 1984, Massaro, Bosseler 2006, Massaro, Cohen 1996, Massaro et al. 1995). In most speech perception studies, Massaro used synthetic visible speech stimuli: participants were presented with a realistic computer-animated face (see Massaro et al. 1995, Massaro, Bosseler 2006). An experimental test was carried out using two experimental conditions: 1) training with a voice and a computer-animated face, and 2) training with a voice only. Five children with autism participated in the study, in which each child continuously learned sets of words with and without the face. Results showed that the learning rate was significantly faster, and the retention was better, with the face (Massaro, Bosseler 2006).

The diversity of methods (live (visual) vs. recorded (audio)) and population (younger vs. older children and adults, typically developing vs. language impaired children) demonstrate controversial results; however, the trend that a live presentation mode provides cues to perform the task more accurately was observed. The variance of the reported results inspired us to conduct a comparative study focusing on presentation mode and age factor. We assume that both presentation conditions (live mode and audio mode) might be equally exploited by younger children with no clear preference and more accurate performance. In contrast, older children might rely more on visual cues and prefer the live mode to perform the task more accurately.

3. Method

In order to investigate different stimuli conditions (live mode vs. audio mode) and their impact on children's performances on the non-word repetition test results, the participants, study procedure and scores are presented below.

3.1. Data and procedure

The study consists of 100 typically developing monolingual Lithuanian-speaking children of two age groups: 4;00–4;11 and 5;00–5;11. Children in each age group were randomly assigned to the two experimental conditions: live mode vs. audio mode (see Table 1). The data were collected in kindergartens in the city of Kaunas (Lithuania) and the Kaunas region. The children were tested individually in a quiet room.

Table 1. Participants: age and gender

Participants	Live stimuli		Audio-recorded stimuli	
	4;00–4;11	5;00–5;11	4;00–4;11	5;00–5;11
N	25	25	25	25
Age mean (month)	51.7	65.3	53.6	65.6
Gender (F/M)	11/14	12/13	14/11	10/15

The Lithuanian non-word repetition test[2] (Krivickaitė, Dabašinskienė 2013) was designed with regard to the structural characteristics of Lithuanian words (word length, vowel and consonant frequency, and syllable structure) (see Chiat 2015). The test consists of 24 non-words of different structures (8 non-words have two syllables (4–6 phonemes); 8 non-words have three syllables (6–7) phonemes; 8 non-words have four syllables (7–8) phonemes). Each group has two non-words without consonant clusters and six non-words with consonant clusters (see Table 2).

Table 2. Non-word items and their syllable structure (syllables are separated by dots)

Two-syllable non-words	Three-syllable non-words	Four-syllable non-words
kemu CV.CV	gelɔʃa CV.CV.CV	sulerɪte: CV.CV.CV.CV
dɔjæ CV.CV	ʃɪrutɑ CV.CV.CV	žɑdevɪnɑ CV.CV.CV.CV
skɪmo CCV.CV	ʃkulɪne: CCV.CV.CV	snɑlɪdɪnɑ CCV.CV.CV.CV
ʃvelɑ CCV.CV	plemutɑ CCV.CV.CV	spɪrɑtuʃɑ CCV.CV.CV.CV
gɑ:pre: CV.CCV	mɑ:spule: CV.CCV.CV	nɪspɑrɪmɑ CV.CCV.CV.CV
gɪtvɑ CV.CCV	lɑsmuvɪ CV.CCV.CV	mɑgvunɔle: CV.CCV.CV.CV
smɪntɔ CCVC.CV	spɑ:dəkɪ CCV.CV.CV	stɑlɪgɔsɑ CCV.CV.CV.CV
klestɑ CCV.CCV	pɑ:svʌpɪ CV.CCV.CV	gɔsɑklu:nɪ CV.CV.CCV.CV

[2] The Lithuanian non-word repetition test was developed participating in the project COST IS0804 Language Impairment in a Multilingual Society: Linguistics Patterns and the Road to Assessment (LITMUS, 2009–2013). https://www.bi-sli.org/nonword-repetition (18.4.2023).

Two stimuli of the non-word repetition test were used:

1. **Live mode.** In all the data collection sessions, a 30-year-old female experimenter presented the children with the non-words, and she was the only one testing all the children with the live stimuli. During the procedure, the experimenter sat next to a child in a quiet room without other people around. The experimenter started: "I am going to say some words to you that I have made up. I want you to repeat each of them after me. The words do not mean anything. Do not be afraid, the words are really strange". The experimenter did not cover her mouth. All of the children's responses were recorded on a digital audio recorder.
2. **Audio mode.** The same non-words were adapted to be presented with audio-recorded stimuli and pictures on a laptop in PowerPoint. The non-words were audio-recorded by a 40-year-old female researcher. The test was presented as a game using *Ms PowerPoint* (see Figure 1) by the same experimenter as in the live stimuli mode. The child was introduced to the main game character, the monkey, who wanted to get some bananas and had to complete the tasks. The child was asked to help the monkey and do the tasks. In each step, the child heard a recorded non-word, which she had to repeat. All conversation was recorded; this way, imprecise answers were not marked in the child's presence. Children liked this task because it was playful, fun, and short. All of the children's responses were recorded on a digital audio recorder.

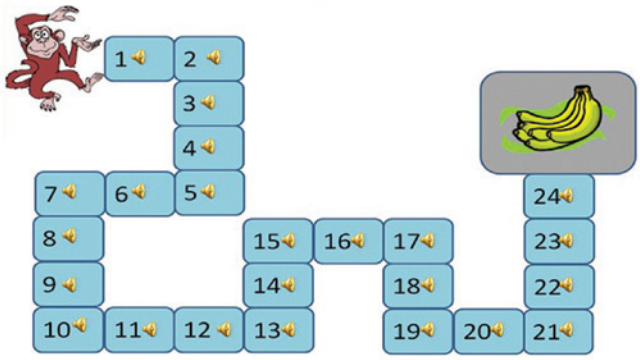

Figure 1. Visual of the non-word repetition test (designed by Kunnari 2011)

3.2. Scoring

Comparative research using non-word tests designed for different languages (e.g., Marshall et al. 2002, Gathercole 2006, Archibald, Gathercole 2006, Chiat, Roy 2007, Girbau, Schwartz 2007, Marshall, Lely 2009, Jones et al. 2010, Gutierrez-Clellen, Simon-Cereijido 2010; Kavitskaya et al. 2011, Williams et al. 2013, etc.) mostly employs three main types of scoring to assess the accuracy of non-word repetition.

Whole-item. Each item was scored as either correct or incorrect. When the sound was omitted, added etc., the answers were scored as incorrect. Each item

of two-, three- and four- syllables stimuli was scored as correct if it was repeated completely correctly (no syllable omission, addition, substitution, etc.). A correct answer was considered to be only the word repeated completely correctly.

The number of syllables. Each item of two-, three- and four-syllable stimuli was scored as correct if a child produced the same number of syllables as it was presented. The answers were regarded to be wrong if the word became one syllable shorter because of an omitted sound (Example 1); the answers were considered to be wrong if an entire syllable was omitted (2) or an additional syllable was added (3), etc. Other parameters were not considered, only the length.

(1) ʃ k u l n e: (= ʃ k u .l ɪ .n e:)[3]
(2) s p a r ɪ ma (= n ɪ . s p a. r ɪ .ma)
(3) d ɔ l u j æ (= d ɔ: .j æ)

Syllable structure.

A) Consonant cluster. Each item was scored as correct if the cluster was repeated completely correctly. Example (4) would be counted as a correct answer because of the correct production of the consonant cluster; example (5) would be counted as an incorrect answer because of the omission of the consonant cluster. The length of the non-word was not considered here.

(4) t ɑ: **p r** e: (= g ɑ: **p r** e:)
(5) k ɪ m o (= **s k** ɪ m o)

B) Consonant cluster position in the non-word. There are consonant clusters in the initial position (5) and in the medial position (6) in Lithuanian non-words.

(5) **s k** ɪ m o
(6) p ɑ: **s v** ʌ p ɪ

Each item was scored as correct if the consonant cluster was repeated completely correctly (7) or incorrect because of omission (8), substitution (9), addition (10) in the target position (initial vs. medial).

(7) **p l** e m u t ɑ; g ɔ s ɑ **k l** u: n ɪ;
(8) n ɪ **p** ɑ r ɪ m ɑ (= n ɪ **s p** ɑ r ɪ m ɑ); **s** ɑ l ɪ d ɪ n ɑ (= **s n** ɑ l ɪ d ɪ n ɑ)
(9) **s n** ɑ l ɪ g ɔ s ɑ (= **s t** ɑ l ɪ g ɔ s ɑ); p ɑ: **s l** ʌ p ɪ (= p ɑ: **s v** ʌ p ɪ)
(10) ʃ **k r** u l ɪ n e: (ʃ **k** u l ɪ n e:); m ɑ **n g v** u n ɔ l e: (= m ɑ **g v** u n ɔ l e:)

The data were coded manually and analyzed using the SPSS (*Statistical Package for the Social Sciences*) programme. In order to check and compare statistically significant differences, the Independent Sample T-test was used, and a standard 0.05 level of statistical significance was chosen.

[3] Syllables are separated by dots.

4. Results: comparison of live vs. audio-recorded stimuli

4.1. The whole-item analysis

The whole-item analysis shows that the non-words presented in live mode statistically were repeated significantly better ($t(95) = 3.632$, $p < 0.000$) than the non-words presented in audio-recording mode. The data show that in the live mode, non-words were repeated with an accuracy of 77%, while in the audio mode they were repeated with an accuracy of 70%.

Figure 2 shows that non-words in the live presentation were repeated correctly within the range of 50% to 92% by the youngest group and between 66% and 100% by the older group.

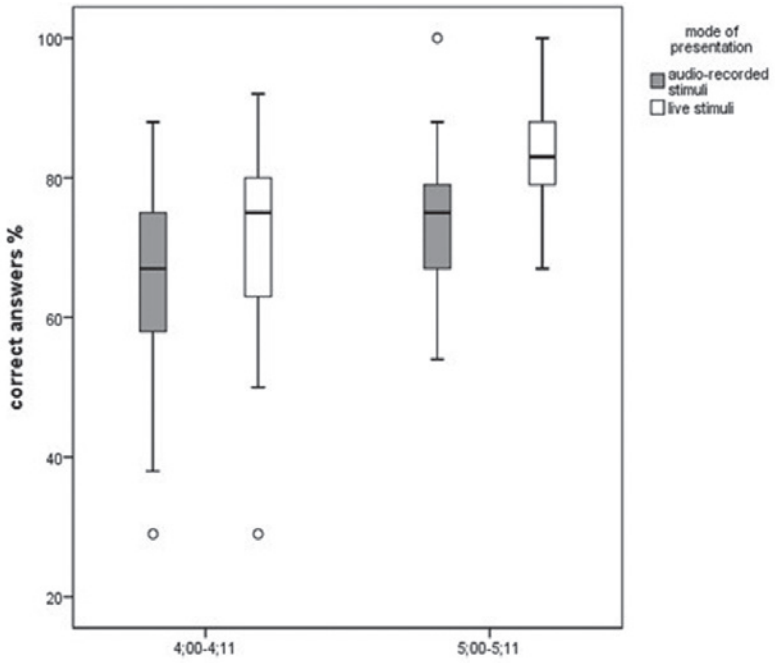

Figure 2. The distribution of the highest and lowest value of the non-word repetition test (live vs. audio-recorded stimuli) by two age groups

The non-words in the audio-recorded mode were correctly repeated within the 38% to 88% range by the 4;00–4;11 age group, whereas the older group repeated non-words correctly between 55% and 87%. A few cases of outliers were identified, and they were removed from the statistical analysis. The statistical analysis has revealed the importance of age for the live mode: 5-year children repeated non-words significantly better than 4-year-olds ($t(47) = 3.581$, $p < 0.001$), while the audio-recorded stimuli were repeated similarly by both age groups.

The length of the word is another important factor. Figure 3 shows the repetition scores of both the *live* and the *audio-recorded* stimuli of two-, three- and

four-syllable non-words. The tendency is clear: the longer the word, the more difficult it is to pronounce it correctly for both age groups and in both modes of presentation; however, the results in favour of live presentation are visible.

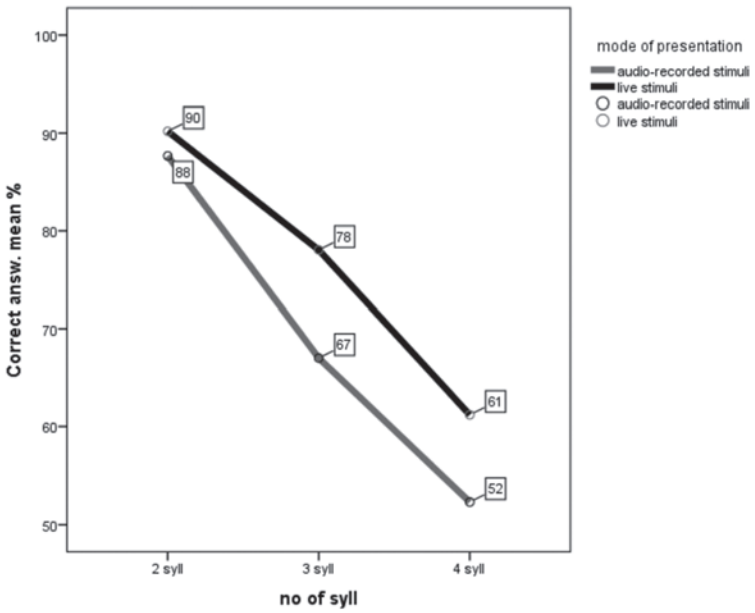

Figure 3. The general accuracy of two-, three- and four-syllable non-words scores (live vs. audio-recorded stimuli)

The analysis of two-syllable non-words shows that the highest values of both presentation modes are the same (99% accuracy), but the lowest value of the audio mode is much lower than that of the live mode (50% vs. 75%) (see Figure 4).

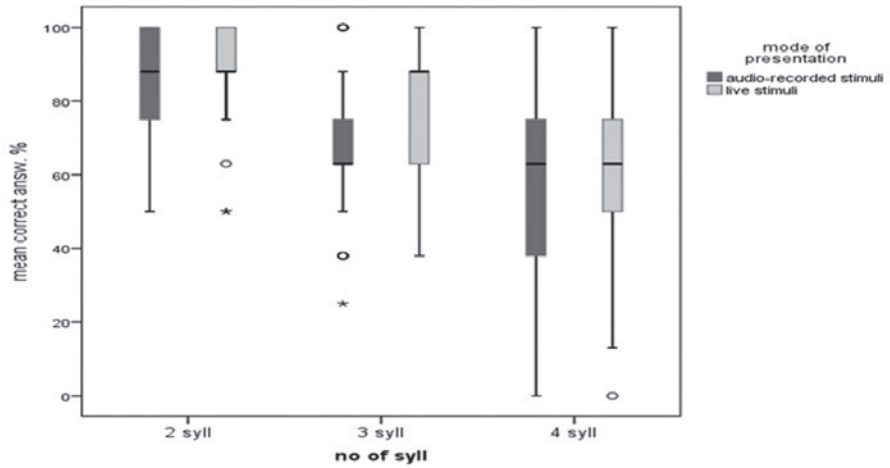

Figure 4. The distribution of highest and lowest values in the production of two-, three- and four-syllable non-words (live vs. audio-recorded stimuli)

The analysis of three-syllable non-words reveals that the range of accuracy is much more extensive in live mode than in audio mode: the lowest value in the live mode is 38%, the highest 100%; the lowest value in the audio mode is 49%, the highest 88%. The analysis of four-syllable non-words shows the largest range of accuracy in both presentation modes. The lowest value in the live mode is 13%, and the highest is 100%; the lowest value in the audio mode is 0 % (no correct answers), and the highest is 100% (all answers are correct).

The analysis[4] has revealed that two-syllable non-words were repeated significantly better in the live mode than the audio mode ($t(95) = 2.118$, $p < 0.037$), as were three-syllable ($t(91) = 3.261$, $p < 0.002$), and four-syllable non-words ($t(97) = 2.418$, $p < 0.017$).

The the whole-item accuracy measure confirms the significance of visual cues in encoding sound sequences in the non-words for both age groups of children. The results demonstrate that the live mode presentation impacted the generally better performance of the non-words regardless of their length and complexity.

4.2. Number of syllables in the word

The analysis of word length has shown that the live stimuli were repeated with an accuracy of 97%, while the audio-recorded stimuli were produced with an accuracy of 99%.

Table 3. The mean percentage for two-, three- and four-syllable non-words

Stimuli	Two-syllable	Three-syllable	Four-syllable
Live	99%	98%	93%
Audio-recorded	100%	100%	94%

However, the detailed analysis has shown that the results of the live and the audio-recorded stimuli of two-, three-, and four-syllable non-words were repeated very similarly, with no significant difference.

The inquiry has demonstrated that both age groups in both modes of presentation repeated the two-syllable non-words at a similar rate, with an accuracy of 99–100%. Three-syllable non-words in the audio mode were repeated with an accuracy of 100% by both age groups; three-syllable non-words in the live mode were repeated with an accuracy of 97–100% (the result of the younger group was lower) (see Figure 5).

Four-syllable non-words were repeated with an accuracy of 92–97% (the result of the younger group was lower) in the audio mode, and with an accuracy of 90–96% (the result of the younger group was lower, $t(106.473) = 2.839$, $p < 0.005$) in the live mode. The results within the younger age group show different performance in repeating four- vs. two- and three-syllable non-words in both presentation modes ($t(60.590) = 4.465$, $p < 0.000$).

[4] Figure 4 also demonstrates some cases of outliers, mostly of poorer repetitions, but they were removed from the statistical analysis.

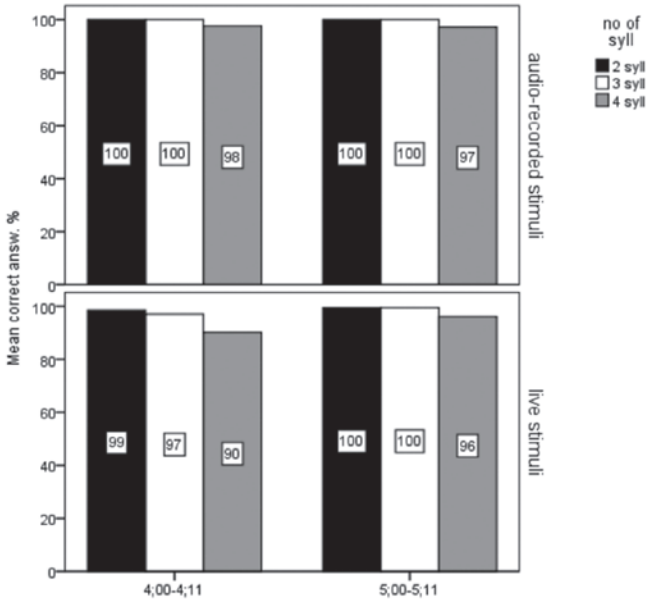

Figure 5. Word length and age groups in *live* vs. *audio* presentation modes

4.3. Syllable structure

The analysis of consonant clusters continues to demonstrate the trend favouring the live stimuli: consonant clusters were repeated with 87% accuracy in the live mode. In comparison, in the audio-recorded mode, they were repeated with 82% accuracy ($t(285) = 2.474$, $p < 0.014$).

Both age groups demonstrate preference towards the live mode: the repetition of the live stimuli with an accuracy of 84–89% ($t(141) = 2.452$, $p < 0.015$), while the audio-recorded stimuli were repeated with an accuracy of 81–84% (see Figure 6).

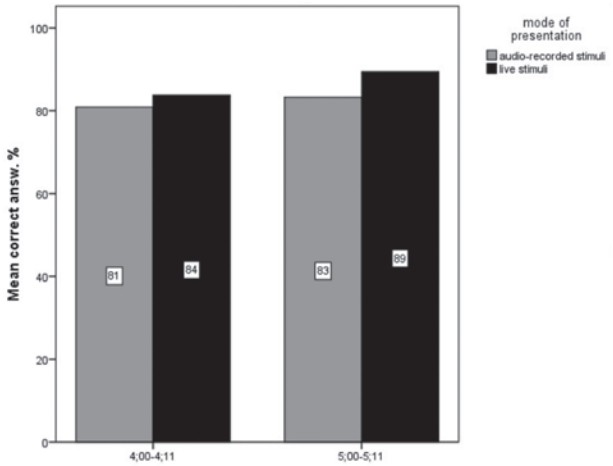

Figure 6. Syllable structure (consonant clusters) and age groups in *live* vs. *audio* presentation modes

4;00–4;11 age group repeated non-words with 84% accuracy in the live presentation mode, whereas the accuracy in the audio presentation was 81%. The older group repeated consonant clusters in the live mode with 89% accuracy, compared to 83% in the audio mode ($t(144) = 2.533$, $p < 0.012$).

The results demonstrate that regardless of the age and stimuli mode applied, children have difficulty accurately producing consonant clusters in non-words; however, the live stimuli show better performance in general, even more significantly in the older group.

Position of the cluster in the non-words: initial vs. medial. The data show that in both stimuli presentation modes, the non-words with consonant clusters in *the initial* position were repeated similarly (87% vs. 90%). In contrast, results in the *medial* position demonstrate statistically significant worse performance for the audio presentation mode (74% vs. 81%) ($t(293) = 2.387$, $p < 0.018$).

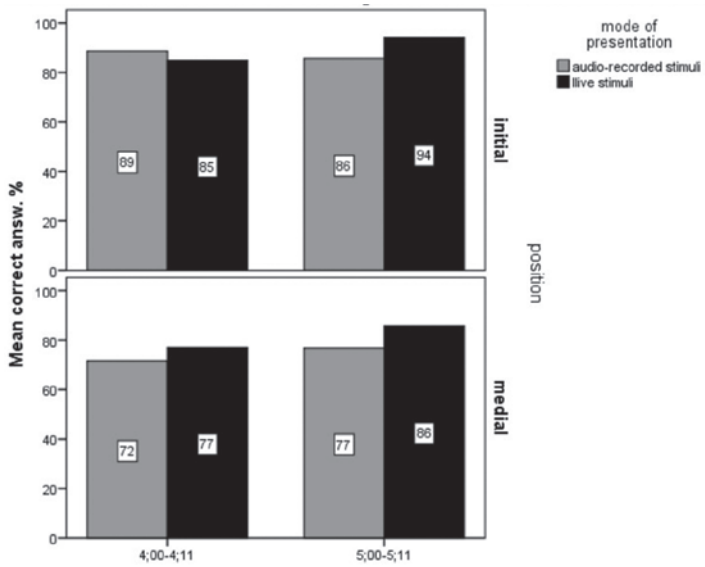

Figure 7. Consonant cluster position and age groups in *live* vs. *audio* presentation modes

Initial position. Both age groups did not demonstrate any statistical difference in repeating consonant clusters with audio-recorded stimuli: 4;00–4;11 year-olds performed with 89% accuracy, and 5;00–5;11 year-olds repeated with 86% accuracy. In the live mode, the consonant clusters were repeated with 85% accuracy by younger children and with 94% accuracy by older children ($t(108.855) = 2.948$, $p<0.004$).

The results within the age groups show different preferences for the presentation mode: the younger group demonstrates better results with audio stimuli, whereas for the older group live stimuli (94% vs. 86%, $t(124.820) = 3.032$, $p < 0.003$) work better.

Medial position. In the audio mode, the consonant clusters were repeated similarly by both age groups: with 72% accuracy by 4;00–4;11 year-olds and 77% accuracy by 5;00–5;11 year-olds. In the live mode, the repetition of consonant clusters scored 77% by 4;00–4;11 year-olds and 86% by 5;00–5;11 year-olds ($t(143) = 2.105$, $p < 0.037$).

The results within the age groups demonstrate a clear preference for the live mode presentation in both age groups; however, this tendency is particularly strong for the older group ($t(245.153) = 3.374$, $p < 0.001$).

The position of the cluster in a word is an essential indicator in assessing accuracy. The clusters in the *initial* position were repeated more accurately than in the *medial* position in both presentation modes and both age groups. A distinct preference for live vs. audio presentation mode is manifested in the older group.

5. Conclusion and discussion

The human face presents visual information during speech production, which is essential for effective communication, particularly in early language acquisition. While the voice alone is usually adequate for communication, visual information from movements of the lips, tongue, and jaws enhances the intelligibility of the message (Cosi et al. 2002).

Research has shown that infants and children are sensitive to the congruence between auditory and lip-read information (Aldridge et al. 1999, Hidalgo-Barnes, Massaro 2007, Teinonen et al. 2008, Okada, Hickok 2009, Lewkowicz, Hansen-Tift 2012, Pons et al. 2015). The already described McGurk effect is one dramatic proof that speech perception is not solely dependent upon the auditory modality (McGurk, MacDonald 1976). The effect has been found to affect infants, preschoolers, and adults (Kuhl, Meltzoff 1982, Desai et al. 2008). To assess whether visual cues work, comparative experimental studies of at least two presentation conditions are usually performed: audition-only stimuli and visual-only stimuli; the third condition might employ audiovisual stimuli (e.g., Norrix et al. 2007, Sekiyama, Burnham 2008, Ross et al. 2011, Erdener et al. 2010, etc.).

In this study, a Lithuanian non-word repetition test was used to assess if different stimuli modes (live vs. audio) impact the accuracy of the performance in two age groups. We have used three scoring types: 1) whole-item accuracy, 2) the length of the non-word, and 3) the structure of the non-word. The general conclusion is that non-word repetition accuracy declines with the increasing number of syllables, i. e. two- and even three-syllable non-words were repeated more accurately than four-syllable non-words. A similar tendency has been noticed for the word structure parameter: longer non-words with consonant clusters, especially in the medial position, were repeated less accurately. As expected, the younger group exhibited lower accuracy for most parameters. Despite these difficulties, children of the studied age were able to produce many longer and more complex non-words correctly.

The comparison of the performance results in live vs. audio-recorded stimuli suggests that the mode of presentation impacts the accuracy of the test results. The study showed the older group's strong preference towards the live presentation mode, especially for the word structure parameter (initial and medial consonant clusters).[5] The clear trend to rely on the live mode by older children might be related to their more extended experience using visual cues when encountering difficulties.

[5] The unpublished results of the non-word repetition test in adult population show a strong tendency towards the live presentation mode ($t(97.871) = 4.178$, $p < 0.000$) (Krivickaitė-Leišienė 2023).

Other parameters did not reveal specific results, although a tendency for higher accuracy in live mode presentation was observed in both age groups.

The research findings suggest that visual information can help a child decode auditory information more easily. We did not consider including interactive elements for a more playful atmosphere (e.g., use of a puppet or figurine) in live mode presentation, as these interactive strategies might help make non-word repetition tasks developmentally more appropriate, thus enhancing young children's attention better than only listening to and repeating non-words (Chiat, Roy 2007, Stokes, Klee 2009). We noticed that our children were more interested in audio-recorded stimuli when the non-words were presented in a recorded audio and picture on a laptop screen. However, this mode has its disadvantage related to the children's more enthusiastic engagement in the computer *game*, which could have resulted in poorer performance compared to live stimuli results, particularly for younger children. Obviously, the computer game approach had its positive outcomes, as children were eager to try the task. Given that we did not control if this engagement resulted in a linguistic outcome, we consider this methodological issue a limitation of the study, as we did not apply the same "game-like" procedure with both presentation modes. The second limitation regards the age groups. Although we had two age groups of children and the sample is sufficient, the results showing the age effect are not straightforward for all the parameters, as both groups are still too close. More distant age groups of children should be considered for future studies, including younger (3-year-old) and older (primary school-age) children.

The study indicates important implications for the assessment process. Speech therapists who assess young children's performance should be aware of the methodological differences and prefer using live stimuli with young TD and LI children.

References

Aldridge, Michelle A.; Braga, Erika S.; Walton, Gail E.; Bower, T. G. R. 1999. The intermodal representation of speech in newborns. – Developmental Science, 2 (1), 42–46. https://doi.org/10.1111/1467-7687.00052

Archibald, Lisa M. D. 2008. The promise of nonword repetition as a clinical tool. – Revue canadienne d'orthophonie et d'audiologie, 32 (1), 24–28.

Archibald, Lisa M. D.; Gathercole, Susan 2006. Nonword repetition: A comparison of tests. – Journal of Speech, Language, and Hearing, 49 (5), 970–983. https://doi.org/10.1044/1092-4388(2006/070)

Botting, Nicola; Conti-Ramsden, Gina 2001. Non-word repetition and language development in children with specific language impairment (SLI). – International Journal of Language & Communication Disorders, 36 (4), 421–432. https://doi.org/10.1080/13682820110074971

Chiat, Shula 2015. Non-word repetition. – Sharon Armon-Lotem, Jan de Jong, Natalia Meir (Eds.), Methods for Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment. Communication Disorders Across Languages. Bristol: Multilingual Matters, 125–150. https://doi.org/10.21832/9781783093137-008

Chiat, Shula; Roy, Penny 2007. The preschool repetition test: An evaluation of performance in typically developing and clinically referred children. – Journal of Speech, Language, and Hearing Research, 50 (2), 429–443. https://doi.org/10.1044/1092-4388(2007/030)

Cosi, Piero; Cohen, Michael; Massaro, Dom 2002. Baldini: Baldi speaks Italian! – 7th International Conference on Spoken Language Processing (ICSLP'02), 2349–2352. https://doi.org/10.21437/ICSLP.2002-224

Desai, Sheetal; Stickney, Ginger; Zeng, Fan-Gang 2008. Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. – The Journal of the Acoustical Society of America, 123, 428–400. https://doi.org/10.1121/1.2816573

Dollaghan, Chris; Campbell, Thomas F. 1998. Nonword repetition and child language impairment. – Journal of Speech, Language, and Hearing, 41 (5), 1136–1146. https://doi.org/10.1044/jslhr.4105.1136

Erdener, Dogu; Burnham, Denis 2013. The relationship between auditory-visual speech perception and language-specific speech perception at the onset of reading instruction in English-speaking children. – Journal of Experimental Child Psychology, 116 (2), 120–138. https://doi.org/10.1016/j.jecp.2013.03.003

Erdener, Dogu; Sekiyama, Kaoru; Burnham, Denis 2010. The development of auditory-visual speech perception across languages and age. – Proceedings of 20th International Congress on Acoustics (ICA 2010), 1–10.

Gathercole, Susan E. 2006. Nonword repetition and word learning: The nature of the relationship. – Applied Psycholinguistics, 27 (4), 513–543. https://doi.org/10.1017/S0142716406060383

Girbau, Dolors; Schwartz, Richard 2007. Non-word repetition in Spanish-speaking children with Specific Language Impairment (SLI). – International Journal of Language & Communication Disorders, 42 (1), 59–75. https://doi.org/10.1080/13682820600783210

Gutierrez-Clellen, Vera; Simon-Cereijido, Gabriela 2010. Using nonword repetition tasks for the identification of language impairment in Spanish-English speaking children: Does the language of assessment matter? – Learning Disabilities Research and Practice, 25 (1), 45–58. https://doi.org/10.1111/j.1540-5826.2009.00300.x

Hidalgo-Barnes, Miguel; Massaro, Dominic W. 2007. Read my lips: An animated face helps communicate musical lyrics. – Psychomusicology, 19 (2), 3–12. https://doi.org/10.1037/h0094037

Hoff, Erika; Core, Cynthia; Bridges, Kelly 2008. Non-word repetition assesses phonological memory and is related to vocabulary development in 20- to 24-month-old. – Journal of Child Language, 35 (4), 903–916. https://doi.org/10.1017/S0305000908008751

Iarocci, Grace; Rombough, Andrienne; Yager, Jodi; Weeks, Daniel J.; Chua, Romeo 2010. Visual influences on speech perception in children with autism. – Autism, 14 (4), 1362–3613. https://doi.org/10.1177/1362361309353615

Jones, Gary; Tamburelli, Marco; Watson, Sarah; Gobet, Fernand; Pine, Julian 2010. Lexicality and frequency in specific language impairment: Accuracy and error data from two nonword repetition tests. – Journal of Speech, Language, and Hearing Research, 53 (6), 1642–1655. https://doi.org/10.1044/1092-4388(2010/09-0222)

Kapalkova, Svetana; Polišenska, Kamila; Vicenova, Zuzana 2013. Non-word repetition performance in Slovak-speaking children with and without SLI: Novel scoring methods. – International Journal of Language and Communication Disorders, 48 (1), 78–89. https://doi.org/10.1111/j.1460-6984.2012.00189.x

Knowland, Victoria C. P.; Evans, Sam; Snell, Caroline; Rosen, Stuart 2016. Visual speech perception in children with language learning impairments. – Journal of Speech, Language, and Hearing Research, 59 (1), 1–14. https://doi.org/10.1044/2015_JSLHR-S-14-0269

Krivickaitė, Eglė 2014. The research of phonotactics acquisition of typically developing children and children with specific language impairment by non-word repetition test. – Special Education, 2 (31), 105–127.

Krivickaitė, Eglė 2016. Vaikų kalbos fonotaktikos raida ['Phonotactics Development of Children's Language']. PhD. Kaunas: Vytauto Didžiojo universitetas.

Krivickaitė, Eglė 2017. Vaikų taikomos žodžių tarimo paprastinimo strategijos: Eksperimentinis tyrimas [‘Pronunciation simplification strategies in child language: Experimental study’]. – Taikomoji kalbotyra, 9, 46–72. https://doi.org/10.15388/TK.2017.17447
Krivickaitė-Leišienė, Eglė 2020. Tarmės įtaka garsų ir jų junginių įsisavinimui: eksperimentinis tyrimas išgalvotų žodžių pakartojimo testu [‘Pronunciation simplification strategies in child language: Experimental study’]. – Taikomoji kalbotyra, 14, 133–147. https://doi.org/10.15388/Taikalbot.2020.14.10
Krivickaitė-Leišienė, Eglė 2023. Išgalvotų žodžių pakartojimo testo metodologiniai skirtumai [‘Methodoloogical issues of the non-word repetition test’]. Taikomoji kalbotyra (in progress).
Krivickaitė, Eglė; Dabašinskienė, Ineta 2013. Lietuvių kalbos fonologijos įsisavinimas: netikrų žodžių kartojimo eksperimentinis tyrimas [‘Acquisition of Lithuanian phonology: Nonword repetition experimental study’]. – Kalbos kultūra, 86, 237–255.
Krivickaitė-Leišienė, Eglė; Dabašinskienė, Ineta 2022. The bilingual advantage: Performing the non-word repetition test. – Philologia Estonica Tallinnensis, 7, 239–262. https://doi.org/10.22601/PET.2022.07.09
Kuhl, Patricia; Meltzoff, Andrew N. 1982. The bimodal perception of speech in infancy. – Science, 218 (4577), 1138–1141. https://doi.org/10.1126/science.7146899
Kunnari, Sari; 2011. Powerpoint presentation for nonword repetition test in COST IS0804 meeting (Unpubliched presentation) http://www.bi-sli.org
Lewkowicz, David J.; Hansen-Tift, Amy 2012. Infants deploy selective attention to the mouth of a talking face when learning speech. – Proceedings of the National Academy of Sciences of the United States of America, 109, 1431–1436. https://doi.org/10.1073/pnas.1114783109
Marshall, Chloe; Ebbels, Susan; Harris, John; Lely, Heather van der 2002. Investigating the impact of prosodic complexity on the speech of children with Specific Language Impairment. – UCL Working Papers in Linguistics, 14, 44–66.
Marshall, Chloe; Lely, Heather van der 2009. Effects of word position and stress on onset cluster production: Evidence from typical development, specific language impairment, and dyslexia. – Language, 85 (1), 39–57.
Massaro, Dominic W. 1984. Children’s perception of visual and auditory speech. – Children Development, 55 (5), 1777–1788. https://doi.org/10.2307/1129925
Massaro, Dominic W.; Bosseler, Alexis 2006. The importance of the face in a computer-animated tutor for vocabulary learning by children with autism. – Autism, 10 (5), 495–510. https://doi.org/10.1177/1362361306066599
Massaro, Dominic W.; Cohen, Michael M. 1996. Perceiving speech from inverted faces. – Perception & Psychophysics, 58 (7), 1047–1065. https://doi.org/10.3758/BF03206832
Massaro, Dominic W.; Cohen, Michael M.; Smeele, Paula M. T. 1995. Cross-linguistic comparisons in the integration of visual and auditory speech. – Memory & Cognition, 23 (1), 113–131. https://doi.org/10.3758/BF03210561
McGurk, Harry; MacDonald, John 1976. Hearing lips and seeing voices. – Nature, 264, 746–748. https://doi.org/10.1038/264746a0
Navarra, Hidalgo Jordi; Soto-Faraco, Salvador 2007. Hearing lips in a second language: Visual articulatory information enables the perception of second language sounds. – Psychological Research, 71 (1), 4–12. https://doi.org/10.1007/s00426-005-0031-5
Norrix, Linda; Plante, Elena; Vance, Revecca; Boliek, Corol 2007. Auditory-visual integration for speech by children with and without specific language impairment. – Journal of Speech, Language, and Hearing Research, 50 (6), 1639–1651. https://doi.org/10.1044/1092-4388(2007/111)
Okada, Kayoko; Hickok, Gregory 2009. Two cortical mechanisms support the integration of visual and auditory speech: A hypothesis and preliminary data. – Neuroscience Letters, 452, 219–223. https://doi.org/10.1016/j.neulet.2009.01.060

Polišenská, Kamila; Kapalková, Svetlana 2014. Improving child compliance on a computer-administered nonword repetition task. – Journal of Speech, Language, and Hearing Research, 57 (3), 1060–1068. https://doi.org/10.1044/1092-4388(2013/13-0014)
Pons, Ferran; Bosch, Laura; Lewkowicz, David J. 2015. Bilingualism modulates infants' selective attention to the mouth of a talking face. – Psychological Science, 26 (4), 490–498. https://doi.org/10.1177/0956797614568320
Radeborg, Karl; Barthelom, Eleonora; Sjorberg, Monika; Sahlen, Birgitta 2006. A Swedish non-word repetition test for preschool children. – Scandinavian Journal of Psychology, 47 (3), 187–192. https://doi.org/10.1111/j.1467-9450.2006.00506.x
Rispens, Judith; Parigger, Esther 2010. Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. – British Journal of Developmental Psychology, 28 (1), 177–188. https://doi.org/10.1348/026151009X482633
Ross, Lars A.; Molholm, Sophie; Blanco, Daniella; Gomez-Ramirez, Manuel; Saint-Amour, Dave; Foxe, John J. 2011. The development of multisensory speech perception continues into the late childhood years. – European Journal of Neuroscience, 33 (12), 2329–2337. https://doi.org/10.1111/j.1460-9568.2011.07685.x
Schwob, Salome; Edde, Laurane; Jacuin, Laure; Leboulanger, Megane; Picard, Margot; Oliveira, Patricia Ramos; Skoruppa, Katrin 2021. Using nonword repetition to identify developmental language disorder in monolingual and bilingual children: A systematic review and meta-analysis. – Journal of Speech, Language, and Hearing Research, 64 (9), 3578–3593. https://doi.org/10.1044/2021_JSLHR-20-00552
Sekiyama, Kaoru; Burnham, Denis 2008. Impact of language on development of auditory-visual speech perception. – Developmental Science, 11 (2), 306–320. https://doi.org/10.1111/j.1467-7687.2008.00677.x
Stokes, Stephania F.; Klee, Thomas 2009. The diagnostic accuracy of a new test of early nonword repetition for differentiating late talking and typically developing children. – Journal of Speech, Language, and Hearing Research, 52 (4), 872–882. https://doi.org/10.1044/1092-4388(2009/08-0030)
Teinonen, Tuomas; Aslin, Richard N.; Alku, Paavo; Csibra, Gergely 2008. Visual speech contributes to phonetic learning in 6-month-old infants. – Cognition, 108 (3), 850–855. https://doi.org/10.1016/j.cognition.2008.05.009
Weikum, Whitney; Vouloumanos, Athena; Navarra, Jordi; Soto-Faraco, Salvador; Sebastian-Galles, Nuria; Werker, Janet 2007. Visual language discrimination in infancy. – Science, 316 (5828), 1159. https://doi.org/10.1126/science.1137686
Wightman, Frederic; Kistler, Doris; Brungart, Douglas 2006. Informational masking of speech in children: Auditory-visual integration. – The Journal of the Acoustical Society of America, 119, 3940–3949. https://doi.org/10.1121/1.2195121
Williams, David; Payne, Heather; Marshall, Chloe 2013. Non-word repetition impairment in autism and specific language impairment: Evidence of distinct underlying cognitive causes. – Journal of Autism Development Disorders, 43, 404–417. https://doi.org/10.1007/s10803-012-1579-8

LASTE TUNDLIKKUS VISUAALSE TEABE SUHTES: MITTESÕNA KORDAMISE TESTI METODOLOOGILISED KAALUTLUSED

Eglė Krivickaitė-Leišienė, Ineta Dabašinskienė
Vytautas Magnuse Ülikool

Artikkel tutvustab leedukeelse mittesõnade kordamise testi ja uurib metodoloogilisi küsimusi kahe erineva esituslaadiga (reaalajas vs. helisalvestus) stiimuli võrdlemise kaudu. Uuringu eesmärk on välja selgitada metoodika mõju tulemustele. Valimis on 100 eakohase arenguga last (kaks vanuserühma: 4;00–4;11 ja 5;00–5;11). Uurimuses hinnati esituslaadi mõju ülesande täitmise täpsusele, keskendudes vanuselistele erinevustele.

Erinevate stiimulite võrdlus näitab, et esituslaad mõjutab mittesõna kordamise testi tulemuste täpsust. Uurimus näitas, et artikulatsiooni visuaalne jälgimine aitab lastel tuvastada mittesõnad: reaalajas esitatud mittesõnade tuvastamine oli täpsem kui helisalvestatud stiimuliga, eriti vanemate laste rühmas. Üldine võrdlev analüüs kinnitas täpsuse kalduvust langeda pikemate ja struktuurselt keerukamate mittesõnade kordamisel sõltumata stiimulist ja vanusest.

Võtmesõnad: keeleomandamine, mittesõna kordamise test, kõnetaju, katsemeetod, leedu keel

Eglė Krivickiatė-Leišienė (Vytautas Magnus University). Her research interests focus on child language acquisition, early bilingualism/multilingualism, and child language disorders.
K. Donelaicio 58, 44248 Kaunas, Lithuania
egle.krivickaite-leisiene@vdu.lt

Ineta Dabašinskienė (Vytautas Magnus University). Her research interests include child first and second language acquisition, heritage language, language policy and multilingualism, grammar and pragmatics of spoken language.
K. Donelaicio 58, 44248 Kaunas, Lithuania
ineta.dabasinskiene@vdu.lt