

Grammatical and stylistic analysis of primary grades textbooks and children's literature quantitative grammatical research results

Kulahmet Moldabek^{1,*}, Otkir Irmetov², Rabiga Kenzhebekova², Gulnar Akilbaeva², Janna Botabaeva³, Sandugash Kulanova³

¹Doctor of Education, International Kazakh and Turkish University named after. H.A.Yasavi

²PhD in Philology, Kazakhstan Engineering and Pedagogical University of Peoples' Friendship

³PhD in Philology, Shymkent University

Abstract: A grammatical and stylistic analysis is conducted in the given article. Dynamics of junior students' vocabulary enrichment were researched based on the data of Kazakh children's literature and primary school textbooks. Unjustified fluctuations in the style of presentation, entailing changes of the quantitative-grammatical structure of educational texts, are not conducive to an involuntary memorization of necessary grammatical rules and structures.

Key words: Analysis; Quantitative analysis; Primary school; Primary school textbooks; Children's literature

1. Introduction

The research of child speech as an initial communication form, which acquaints the human being with the environment, has to be defined firstly by the fact that the mastering of the oral and then written speech during childhood is a general and most natural way of acquainting the developing human being and the entire generation with their national collective, united by its own language (Dzhubanov, 1987).

Mastery of the language in the childhood period during direct interaction with the national environment provides for, on the one hand, a spiritual cultural-esthetic and ideological-political formation of a future citizen, while on the other hand – the existence and unity of the nation.

Researching the formation and development of the child speech activity is important for both the solution of psycholinguistics and child psychology issues and the analysis of a number of important social studies and language nature theoretical questions. In particular, these researches are of primary importance to the human understanding of the ontogeny and, through it, the phylogeny of a number of lexical and grammatical language categories, its phonological system, as well as text generation mechanisms (Van der Lely, 2001).

The research of the child-child, adult-child, child-adult communication should follow two courses:

- The course of studying child language and speech (child oral speech recordings and their written "creative work");
- The course of researching texts intended for children.

Taking into consideration the fact that statistical study of child oral and written creative work faces the current lack of statistically sufficient material, the authors have chosen the second course – the research of the adult-child written communication. The research of child linguistic behavior is a supportive technique.

In the first case, necessary information regarding child language and speech structure is extracted directly from texts created by children. In the second case, this information is discovered indirectly – through the text of an adult, who is guided, wittingly or unwittingly, by the child's thesaurus and competence.

The obtained and analyzed extensive lingual statistical information of child linguistic exercises are of interest for typological, stylistic and informational-statistic researches;

- During theoretical development of modern Kazakh language lexicology, lexicography and morphology issues;

- During typological, stylistic and informational-statistic researches of Turkic languages;

- During the research of child speech ontogeny and its typology in terms of related (Turkic) and unrelated languages;

- During the definition of the theoretical and methodological base for textbook composition and during the selection of materials for the composition of Kazakh school elementary grade textbooks and educational lexicography;

- During the development of a specific method for teaching the Kazakh language in primary, basic and secondary schools;

* Corresponding Author.

- During the composition of educational dictionaries.

Thus, the quantitative research of child speech and speech communication with children in general in the Kazakh, general Turkic, Indo-European, Finno-Ugric and other linguistics is necessary to understand the ontogeny of the language system and human speech skills, as well as to discover lexical-grammatical peculiarities of child speech, which differentiate it from adult speech.

2. Literature review

Children's assimilation of the mother tongue is not a consistent mechanical explanation of the vocabulary and addition of grammatical forms and rules. However, it is thought of as a difficult mental and creative process (Leonard, 2007), an embodying and consolidating result of the child's knowledge of the outside world (Orphoepic dictionary of the Kazakh language, 2004).

Professor of Language and Communication in the Faculty of English Language and Literature at the University of Oxford Jean Aitchison states that during the process of language acquisition children will vary individually when they reach each stage, but there is little variation in the sentence of language learning. By the end, a child's language is in place and he or she has a basic lexicon of several thousand words. From now on, what is learned is retained and increasingly dependent upon experiences and environment – on opportunities to use language and hear it used, for a wide range of purposes and audiences in a wide range of contexts (Aitchison, 1998).

Not only the assimilation of new concepts and lexical items, but also a new calculation and interpretation of previously learned concepts and their lexical units' embodiment are performed during this process. These processes are accompanied by the formation of new associative (paradigmatic) relations and valence prognosis of extension of old and formation of new lexical microsystems and grammatical paradigms (Zhubanov, 2008).

Noam Chomsky believes that children are born with an inherited ability to learn any human language. He claims that certain linguistic structures, which children use so accurately, must be already imprinted on the child's mind. Chomsky believes that every child has a "language acquisition device" or LAD, which encodes the major principles of a language and its grammatical structures into the child's brain. Children have then only to learn new vocabulary and apply the syntactic structures from the LAD to form sentences. Chomsky points out that a child could not possibly learn a language through imitation alone because the language spoken around them is highly irregular – adult speech is often broken up and even sometimes ungrammatical. Chomsky's theory applies to all languages as they all contain nouns, verbs, consonants and vowels and children appear to be 'hard-wired' to acquire the

grammar. Every language is extremely complex, often with subtle distinctions, which even native speakers are unaware of. However, all children, regardless of their intellectual ability, become fluent in their native language within five or six years (Chomsky, 1975).

Chomsky's ideas have had a strong influence on researchers of children's language acquisition, although many researchers in this area, such as Elizabeth Bates (1991) and Michael Tomasello (2008) argue very strongly against Chomsky's theories, and instead advocate emergentist or connectionist theories, explaining the language with a number of general processing mechanisms in the brain that interact with the extensive and complex social environment in which language is used and learned.

Sylvana Sofkova Hashemi concerns herself with the analysis of grammar errors in Swedish texts written by primary school children and the development of a finite state system for finding such errors. Grammar errors are more frequent for this group of writers than for adults and the distribution of the error types is different in children's texts. In addition, other writing errors above word-level are discussed here, including punctuation and spelling errors resulting in existing words (Sofkova Hashemi, 2003).

C. C. Anderson tested the standard assumptions about differences in language usage in adult and children's literature by analyzing parallel passages from the works of four authors – Nathaniel Hawthorne, George MacDonald, Oscar Wilde, and John Gardner, who each wrote prose fantasy for both audiences. A computer program and syntactic code based on those used by York University in Toronto provided a statistical analysis of the 20,000 words of selected text (Johnson and Fey, 2006). The scientist found that the passages from the children's books had much shorter paragraphs, and slightly shorter sentences, T-units, clauses, and words. T-units were the most consistently and notably reduced elements. The children's books also had more lexical repetition and fewer abstract and Latinate words and tended towards a verbal style (Jokel et al., 2010). These characteristics support some of the common assumptions about children's literature, but the differences were slighter than anticipated. In the area of syntax, the assumption that coordination would increase and subordination decrease markedly in the children's stories did not prove true. Coordination was only marginally more frequent in the children's passages, and subordination nearly equal in both sets. The reduction of prepositions in the juvenile samples seems of more significance syntactically (Johnson and Morris, 2007). In the children's passages there are large increases of the amount of dialogue and in the use of Germanic based words. The general conclusion of the researcher is that the differences in the children's passages reflect a stronger tendency towards everyday speech, that children's authors borrow more conventions from

conversation and from oral traditions when writing for a child audience (Anderson, 1984).

3. Discussion & results

The analysis of grammatical quantitative data obtained from statistical surveys, by means of computer-assisted educational texts and children's literature texts, shows that grammatical features of the speech children, entering school, is best reflected in the "ABC Book". Despite the fact that the distribution of word classes in the "ABC Book" is generally quite similar to their distribution in the texts of other books, children's, belles-lettres and public texts, the "ABC Book" language is characterized by the following features:

The predominance of simple unextended and nominative sentences, cf. ai – "moon", ayna – "mirror", Mayra – "proper name", etc.

Anau - ak kaz – "This is a white goose", Zada men Zere - okushy – "Zada and Zere are students", etc.

Simplified form of construction; It primarily becomes evident in:

a) the predominance of original root forms and stems without a binding morpheme in the text of the "ABC Book", while in the texts of other textbooks and children's literary texts, the ratio of roots and stems of nouns, used as independent word forms, fluctuates between 33 and 45%; among indirect case forms, the guide-dative form is predominant in the "ABC Book", while in other textbooks, these forms are usually almost as frequent as the genitive form;

b) the prevalence among nominal word forms with inflexion of simple two-membered structures of the R + , N + 0 type (81%), the very small ratio of three units and four-membered structures (not more than 19%), which results in about 87% of nominal word forms formation schemes, possible in terms of the grammatical system and norms of Kazakh language, missing from the "ABC Book";

c) the use of pronouns, which is considerably more frequent in the "ABC Book" (about 23%) than in other textbooks and literary texts for children, with an the average 10% use in other texts, wherein the use of first person possessive endings is typical (approximately 64.4%, compared to 0-1.42% in the mathematics textbook and 17.11-48.62% in the "Kazakh language" textbook). This means that for children, who entered school, the "ABC Book" is the first textbook, with which he masters the orthography and the initial explanatory information regarding the grammar and vocabulary of their native language. Therefore, it is natural that the text of this textbook should be very similar to their native language's lexical-grammatical information and rules, which have already been mastered during the preschool period.

Extension of the native language grammatical system core, formed in the child at the time of his or her admission to school, is provided alongside the process of teaching grammar, other textbooks and children's literature. The authors emphasize that a solution of the given problem is provided by not only

literature textbooks ("ABC Book", "Kazakh language", "Mother Tongue"), but also textbooks on natural and mathematical sciences ("Mathematics", "Nature study"). This process particularly demonstrates the interdisciplinary connections in secondary school.

Statistical analysis of children's literature and textbooks texts shows that in terms of text grammatical organization, different textbooks are aimed at different literary language types. For example, the grammar of the "Mother Tongue" texts, as well children's fiction texts, focuses on grammatical rules of fiction, as is evident from:

- The similar parts of speech ratio;

- The increase of sentence constructing schemes complexity and diversity by the 3rd grade. While in the "ABC Book" and other 1st grade textbooks, the simple unextended sentence is the most typical syntactic scheme, 3rd grade textbooks often feature sentences with such complicated structure types and word forms as: Kusht zhel topyrakty koter p, eg st ng tub n ashyp keted – "A strong wind is blowing and shifting the soil, leaving bare crop roots";

- The case forms distribution percentage, similar to that of the fictional style;

- The similar to the fictional style distribution of possessive forms in the texts of "Mother Tongue" for the 2nd and 3rd grades (the 1st grade "Mother Tongue" text is similar to the grammar of the "ABC Book" and speech of preschoolers in terms of its grammatical organization).

In other words, the "Mother Tongue" text is recognized as a grammatical norm, peculiar to the literary language, for developing junior Kazakh students.

In contrast, the grammatical organization of the "Kazakh language", "Mathematics" and "Mother Tongue" textbooks' exposition is focused on the socio-political, scientific and technical diversity of the Kazakh literary language. This is evidenced, in particular, by:

- A high noun use percentage (up to 46.55%) in the above-mentioned textbooks and in Kazakh newspaper texts, which is associated with the nominative form of presentation, usually typical for the scientific-technical and semi-journalistic style of different languages;

- A minimal use of word forms, in comparison with the texts of fiction, children's literature and texts of the "Mother Tongue", which are rich in grammatical patterns;

- A common use of third person forms (99%).

Thus, the textbook and children's literature texts consolidated grammatical forms through involuntary memory, metabolized by active grammar learning, which develops the native language stylistic flair of Kazakh junior students. By reading and developing these texts, children will learn to speak and use grammatical techniques, specific to colloquial, belles and then journalistic and scientific-technical speech.

The formation and expansion dynamics of the Kazakh children mother tongue grammatical structure in primary school is determined by a

sequential complexity of statements, due to the ever-expanding activities of the child in terms of mastering of the world, his language and mental competence growth and operating activity (Zhantekeyeva, 2006).

The results of this process are easily detected by comparing the standard phrases of the older preschooler and a first grader novice, on the one hand, and a third-grader, on the other hand. This is how this process is reflected in the grammatical statistics of children's literature and textbook texts.

The analysis shows that the percentage of concrete grammatical construction schemes of nominative word forms, not recorded in educational texts, gradually drops, from grade to grade. For example, if the "ABC Book" text covers only 13% of all practically used nominative grammatical schemes (while almost 87% of such schemes are not reflected in this text), 3rd grade textbooks take into account at least 30% of nominal word forms construction schemes (even more than 50% in the "Mother Tongue"). In the children's literature texts, the ratio of accounted name structure schemes is 61%, which is virtually identical to the data regarding children's prose (55%).

The structure of nominal word forms is gradually complicated. Indeed, while four-component schemes, except for (N + K + T + S (bala.lar.y.m.a) "on my kids", are rare in 1st grade book texts, five-component schemes are absent altogether. In 2nd grade books, four-component schemes are more frequently encountered. In addition, in 3rd grade educational texts, alongside four-component word forms - ki.e.t.n.m.n (literally: "clothed"), at.ta.gy.dan (literally: "from sitting on a horse"), even six-component word forms appear - oy.yl.a.tyn.y.n (literally: "supply, supplied"), which also occur in children's literature texts - suret.te.g.s.n.dey (literally: "as in the picture (figure))".

The use of the third person possessive form in book texts increases from grade to grade. It helps introduce children to literary written speech norms, which, unlike oral and colloquial speech used by seniors, preschoolers and first grade students, is focused not on the first and second person, but on the third person of the Kazakh noun and verb. As a result, 3rd grade book texts (apart from the "Kazakh language" textbook), in terms of statistics regarding the use of the third person possessive form, are not that different from the texts of children's literature, fiction and journalism.

Of course, not every quantitative change in a commonly used grammatical category is evident of a stable and consciously controlled development and expansion dynamics of the grammatical system of primary school children. For example, the decline in the use of conjunctions in textbooks on mathematics (20% compared to 5.3% in the same 2nd grade textbook and 5.7% in the 3rd grade textbook) is devoid of any internal or psycholinguistic linguo-didactic meaning, although random changes in the mathematics textbook presentation style also affirmed noticeable fluctuations in the accusative

form use ratio (1st grade - 23%, 2nd grade - 6.7%, 3rd grade - 15%). Similarly, there is a lack of stability in the use of cases in the "Kazakh language" textbook.

Thus, the accusative case use ratio is reduced: 34.7% in the 1st grade, approximately 25% in the 2nd grade, and 20.5% in the 3rd grade), while the genitive case ratio is increasing (16.6% in the 1st grade, 25.8% in the 2nd grade, and 27% in the 3rd grade). While being incidental from the perspective of psycholinguistic and linguo-didactic dynamics, there is an obvious radical increase in the number of first person possessive forms in the text of the "Kazakh language" for the 3rd grade.

4. Conclusion

Unjustified fluctuations in the style of presentation, entailing changes of the quantitative-grammatical structure of educational texts, are not conducive to an involuntary memorization of necessary grammatical rules and structures. Therefore, when writing any textbook, any changes of the presentation manner, the quantitative-grammatical (and lexical) text structures of this book, must be consciously regulated according to psycholinguistic and linguo-didactic considerations, taking into account the interests of the formation and expansion of the child's native language grammatical system, as well as the prospects for language teaching optimization.

Formation dynamics of the Kazakh junior students' native language grammatical system is achieved by extending and expanding the grammatical structure core, which formed in the child at the first stage of language ontogeny, attributed to the age of 2-7.

This final formation of the native language grammatical system happens in accordance with the Turkic grammatical stereotypes, as demonstrated by the psycholinguistic test and machine statistics survey of primary school texts of textbooks and children's literature. During this period, the child masters the syntactic mechanisms with a view to constructing a simple unextended sentence, as well as to generate ordinary nominal forms, consisting of one, two and sometimes three morphemes.

As shown by the quantitative-linguistic experiment (analysis), these initial grammatical patterns are fixed in the text of the "ABC Book", which is the first textbook for a child entering school. This book helps him or her to develop orthography and native language grammar explication principles. The initial grammatical system core is developed alongside teaching grammatical rules during lessons, as well as teaching texts of such books as "Kazakh language", "Mother tongue", reading children's fiction, as well as texts on mathematics and natural study.

References

- Aitchison J. (1998), *The Articulate Mammal: An Introduction to Psycholinguistics*. 4th edition, New York: Routledge.
- Anderson C.C. (1984), *Style in children's literature: a comparison of passages from books for adults and for children*. 231 p.
- Bates E. (1991) *From First Words to Grammar: Individual differences and Dissociable mechanisms*.
- Chomsky N. (1975), *Reflections on Language*, New York: Pantheon Books.
- Dzhubanov A. Kh. (1987), *Quantitative structure of Kazakh texts*, (linguistic analysis of computer experiment). Alma-Ata, p. 147.
- Johnson, B. W., Fey, M. E. (2006), "Interaction of lexical and grammatical aspect in toddlers' language". *Journal of Child Language*, 33, p. 419-435.
- Johnson, B. W., Morris, S. R. (2007), "Clinical implications of the effects of lexical aspect and phonology on children's production of the regular past tense". *Child Language Teaching and Therapy*, 23, p. 287-306.
- Jokel, R., Rochon, E., Anderson, N. (2010), "Errorless learning of computer-generated words in a patient with semantic dementia". *Neuropsychological Rehabilitation*, 20, p. 16-41.
- Leonard, L. B. (2007), "Lexical aspect and the use of verb morphology by children with specific language impairment". *Journal of Speech, Language, and Hearing Research*, 50, p. 759-777.
- Orphoepic dictionary of the Kazakh language (2004), Almaty: Aryss.
- Sofkova Hashemi S. (2003), *Automatic Detection of Grammar Errors in Primary School Children's Texts*. Humanisten: Göteborg University, 234 p.
- Tomasello M. (2008) *Origins of Human communication*. 393 p.
- Van der Lely, H., & Ullman, M. (2001), "Past tense morphology in specifically language impaired children and normally developing children". *Language and Cognitive Processes*, 16, p. 177-217.
- Zhantekeyeva Z., Nakysbekov O. (2006), *Kazakh-Russian and Russian-Kazakh dictionary for students*. Astana, p. 208
- Zhubanov A. (2008) *Issues of applied linguistics*, Almaty: Aryss, p. 636