

Research article

UDC 811.512.122'42:81'255.2:37.016

DOI: <https://doi.org/10.18721/JHSS.17311>

EDN: <https://elibrary/HWTFPA>



KAZAKH FOLKLORE AND CHILDREN'S TEXTS: PRAGMATIC AND SEMANTIC TRANSFORMATIONS OF NARRATIVE IN DIDACTIC PRACTICE

B. Myrzabek , M. Mambetova

Al-Farabi Kazakh National University, Almaty, Republic of Kazakhstan

 bibikhanmm@mail.ru

Abstract. This article examines how structural-semantic transformations reorganize cohesion, causal coherence, event sequence, and narrative functions in Kazakh folklore-didactic and children's narratives. While level-based textual modification is commonly described in terms of lexical, syntactic, and quantitative simplification, less attention has been paid to how individual transformations affect referential continuity, causal recoverability, genre framing, characterization, and the overall pragmatic effect of a narrative. The study therefore aims to determine how these textual relations and functions change under two degrees of transformation, represented by A1 and A2 versions, and which components remain invariant in coherent transformed narratives. The corpus comprised eight Kazakh source narratives and sixteen purpose-built A1 and A2 versions, organized into eight aligned original–A1–A2 sequences. A total of 612 transformational segments were analyzed using a two-level coding model that distinguished six primary operations – structural compression, syntactic simplification, lexical substitution, semantic reduction, causal explication, and neutralization of imagery from their effects on textual organization. Preservation of the functional narrative core was treated as a control category of functional continuity. The findings show that A1 transformation relies more strongly on compression, syntactic redistribution, repeated reference, and a linear event sequence, whereas A2 retains a denser network of source-text relations and more frequently makes causal connections explicit. Syntactic splitting shifts cohesive relations from intra-sentential structures to inter-sentential linkage, while causal explication compensates for contextual relations weakened by compression. Genre framing, characterization, and conflict presentation may be reduced or recast without disrupting the central event structure. Macro-level coherence depends on the recoverability of the functional chain linking the principal participant, key event, motive or causal relation, plot resolution, and the narrative's overall pragmatic effect. The proposed model contributes to the linguistic analysis of transformed narratives by relating formal operations to their consequences for cohesion, causal coherence, event sequence, and narrative function.

Keywords: structural semantic transformation, Kazakh narrative, folklore-didactic text, children's narrative, textual cohesion, causal coherence, narrative function.

Citation: Myrzabek B., Mambetova M., Kazakh Folklore and Children's Texts: Pragmatic and Semantic Transformations of Narrative in Didactic Practice, *Terra Linguistica*, 17 (3) (2026) 181–206. DOI: 10.18721/JHSS.17311



КАЗАХСКИЕ ФОЛЬКЛОРНЫЕ И ДЕТСКИЕ ТЕКСТЫ: ПРАГМАТИЧЕСКИЕ И СЕМАНТИЧЕСКИЕ ТРАНСФОРМАЦИИ НАРРАТИВА В ДИДАКТИЧЕСКОЙ ПРАКТИКЕ

Б. Мырзабек , М. Мамбетова

Казахский национальный университет имени Аль-Фараби,
Алматы, Республика Казахстан

 bibikhanmm@mail.ru

Аннотация. В статье рассматривается, каким образом структурно-семантические трансформации реорганизуют когезию, каузальные отношения, последовательность событий и нарративные функции в казахских фольклорно-дидактических и детских нарративах. Хотя уровневая модификация текста обычно описывается с точки зрения лексического, синтаксического и количественного упрощения, значительно меньше внимания уделяется тому, как отдельные трансформации влияют на референциальную непрерывность, восстановимость причинно-следственных связей, жанровое оформление, характеризацию персонажей и общий прагматический эффект нарратива. В связи с этим цель исследования состоит в определении того, как изменяются данные текстовые отношения и функции при двух степенях трансформации, представленных версиями уровней A1 и A2, а также какие компоненты сохраняют инвариантность в когерентных трансформированных нарративах. Корпус исследования включал восемь исходных казахских нарративов и шестнадцать специально созданных версий уровней A1 и A2, организованных в восемь выровненных последовательностей «оригинал–A1–A2». В общей сложности было проанализировано 612 трансформационных сегментов с использованием двухуровневой модели кодирования, в которой шесть основных операций – структурная компрессия, синтаксическое упрощение, лексическая замена, семантическая редукция, каузальная экспликация и нейтрализация образности – рассматривались отдельно от их последствий для текстовой организации. Сохранение функционального нарративного ядра рассматривалось как контрольная категория функциональной непрерывности. Результаты показывают, что трансформация на уровне A1 в большей степени связана с компрессией, синтаксическим перераспределением, повторной номинацией и линейной организацией событий, тогда как версии A2 сохраняют более плотную сеть отношений исходного текста и чаще эксплицируют причинные связи. Синтаксическое расщепление переносит когезивные отношения из внутрипредложенческих структур на уровень межфразовой связи, тогда как каузальная экспликация компенсирует ослабление контекстуальных связей, вызванное компрессией. Жанровое оформление, характеристика персонажей и представление конфликта могут редуцироваться или перестраиваться без нарушения центральной событийной структуры. Макроуровневая когерентность зависит от восстановимости функциональной цепочки, связывающей главного участника, ключевое событие, мотив или причинную связь, сюжетную развязку и общий прагматический эффект нарратива. Предложенная модель позволяет соотнести формальные операции с их последствиями для когезии, каузальных отношений, последовательности событий и нарративной функции и тем самым уточняет лингвистическое описание структурно-семантической трансформации нарративного текста.

Ключевые слова: структурно-семантическая трансформация, казахский нарратив, фольклорно-дидактический текст, детский нарратив, текстовая когезия, каузальные отношения, нарративная функция.

Для цитирования: Мырзабек Б., Мамбетова М. Казахские фольклорные и детские тексты: прагматические и семантические трансформации нарратива в дидактической практике // *Тегга Linguistica*. 2026. Т. 17. № 3. С. 181–206. DOI: 10.18721/JHSS.17311



Introduction

Textual transformation affects not only linguistic complexity but also the relations through which a text functions as a coherent whole. Compression may remove links between episodes, syntactic splitting may redistribute relations from clause-internal structures across separate sentences, lexical replacement may preserve reference while modifying evaluation or genre marking, and explicitation may make previously implicit relations overt. These effects are particularly consequential in narrative discourse, where continuity depends on participant tracking, event progression, motivational and causal relations, plot resolution, and the communicative effect produced by the completed narrative. Textual transformation therefore needs to be examined not simply as a reduction in linguistic complexity but as a reorganization of structural, semantic, and discourse relations.

Research on text simplification provides an important basis for analyzing such transformations while also revealing the limitations of a purely sentence-level perspective. F. Alva-Manchego et al. [1] show that simplification involve different rewriting operations, including lexical replacement, reordering and sentence splitting, with preservation of source meaning remaining a fundamental constraint. Moving beyond isolated sentences, L. Cripwell et al. [2] demonstrate that applying sentence-level simplification independently across a document may fail to preserve its discourse structure coherently. S. Agrawal et al. [3] likewise show that meaning preservation cannot be adequately assessed without considering the wider textual context in which simplified sentences occur. These findings are particularly relevant to narrative discourse, where a locally successful transformation may weaken a relation, whose function becomes apparent only across sentences or episodes.

This problem is closely connected with cohesion and coherence. M.C.J. Scholman et al. [4] treat causal relations as a theoretically and empirically distinguishable class of discourse coherence relations and emphasize the need for explicit analytical criteria. G. Mulder et al. [5] further show that causal coherence relations extend beyond surface linguistic form and are represented at higher levels of discourse interpretation. For narrative discourse, J. Cevasco et al. [6] review evidence demonstrating that causal connections are central to the construction of coherence. Taken together, these studies indicate that retaining an event does not necessarily preserve the relation that motivates it or connects it to its consequence. Compression may remove contextual support for such a relation, whereas causal explication may restore its recoverability through an overt connector or an explicit formulation of motivation or consequence.

Lexical transformation raises a comparable issue because denotative equivalence does not necessarily entail textual-functional equivalence. K. North et al. [7] characterize lexical simplification as the replacement of complex lexical material by simpler alternatives while maintaining the relevant source meaning. In narrative discourse, however, a denotatively close replacement may preserve participant reference or basic propositional content while weakening evaluation, emotional intensity, characterization, or genre marking. The analysis therefore distinguishes lexical substitution, in which reference and relevant textual function remain substantially stable, from neutralization of imagery, in which figurative, evaluative, emotional, characterizing, or genre-related meaning is weakened.

Kazakh folklore-didactic and children's narratives provide particularly suitable material for examining these relations because event-centered organization coexists with dialogue, repetition, evaluative vocabulary, culturally marked imagery, and conventional formulaic expressions. O. Tyshchenko et al. [8] distinguish formulae occurring in different structural positions of fairy-tale discourse and demonstrate their relevance to the organization and pragmatics of folklore texts. In the corpus examined here, folklore formulae comprise conventionalized expressions associated with narrative organization and genre framing, including opening and closing formulae where they occur, as well as recurrent fixed narrative expressions. Conversational discourse formulae are excluded from this category. Such units may remain peripheral to the central event progression while still contributing



substantially to genre framing; their removal can therefore leave the event sequence recoverable while reducing a distinct narrative function.

The A1 and A2 levels of the Common European Framework of Reference for Languages provide the empirical conditions under which different degrees of structural-semantic transformation can be compared. The Council of Europe [9] establish the reference framework for these proficiency levels, while Z. Iskakova et al. [10] specifically emphasize the role of lexical and grammatical selection in A2-level Kazakh-language materials. In this study, however, A1 and A2 are not treated as the principal theoretical object. They function as two controlled transformation conditions that allow changes within the same source narrative to be compared. The central linguistic question is therefore not whether a transformed text merely corresponds to a particular proficiency label, but how its internal textual organization changes when linguistic material is compressed, redistributed, substituted, explicated, or functionally recast.

Despite substantial progress in research on text simplification and discourse organization, the relation between particular transformation operations and their consequences for narrative structure remains insufficiently described. Existing studies account for a range of lexical and syntactic transformations and increasingly recognize the importance of document-level meaning preservation [1–3], while discourse research provides established approaches to coherence and causal relations [4–6]. Less clear is how an identifiable structural-semantic operation affects several interconnected properties within aligned versions of the same narrative: whether referential and lexical cohesion is retained or reconstructed; whether causal relations remain recoverable; whether the event sequence remains functionally connected; and whether genre framing, characterization, conflict presentation, plot resolution, and pragmatic closure are retained, reduced, or recast. These questions have received particularly limited attention in relation to transformed Kazakh folklore-didactic and children's narratives.

To address this problem, the study analyses eight aligned original–A1–A2 sequences comprising eight Kazakh source narratives and sixteen purpose-built transformed versions. The corpus contains 24 textual units and 612 transformational segments. Six primary operations – structural compression, syntactic simplification, lexical substitution, semantic reduction, causal explication, and neutralization of imagery – are examined in relation to three interconnected dimensions: cohesive reorganization, causal coherence and event sequence, and narrative function. Preservation of the functional narrative core is assessed separately as a control category of functional continuity. This core is identified through the recoverability of the relation linking the principal participant, key event, motive or causal relation, plot resolution, and pragmatic closure.

The study proposes a two-level model that separates the formal identification of a structural-semantic operation from its consequences for textual organization. The model establishes whether transformation retains, weakens, or reconstructs cohesive relations; whether causal relations remain recoverable, become explicit, or are disrupted across the event sequence; and whether narrative functions are retained, reduced, or recast. This distinction makes it possible to account for cases in which extensive compression preserves macro-level coherence because functionally necessary relations remain recoverable, as well as cases in which a comparatively limited local transformation produces semantic distortion by disrupting a relation required for interpreting subsequent events, plot resolution, or pragmatic closure.

Accordingly, the study aims to determine how structural-semantic operations reorganize cohesion, causal coherence, event sequence, and narrative functions in Kazakh folklore-didactic and children's narratives under two transformation conditions and to identify the functional relations that remain invariant in coherent transformed versions. The following research questions are addressed:

1. How do structural-semantic operations reconfigure local and inter-sentential cohesion in aligned original–A1–A2 sequences of Kazakh narratives?



2. How do compression, syntactic restructuring, and causal explication affect causal coherence and the recoverability of relations across the event sequence?

3. Which narrative functions are retained, reduced, or recast across transformation, and which relations constitute the functional narrative invariant required for macro-level coherence?

4. How do the profiles of cohesive reorganization, causal explicitness, event-sequence preservation, and functional continuity differ between the A1 and A2 versions of the same source narratives?

The study thus treats A1 and A2 as comparative conditions of structural-semantic reorganization rather than as two quantitative degrees of shortening. Its central linguistic concern is how local transformations redistribute cohesive resources, alter the recoverability of causal and event relations, and preserve or recast the functions through which narrative continuity, plot resolution, and pragmatic closure are established.

Literature review

Current research increasingly treats text simplification as a set of identifiable linguistic operations rather than as a uniform reduction in length or difficulty. R. Cardon et al. [11] organize the field around the notion of *operations*, understood as modifications applied to an input text during transformation. This perspective is particularly useful when the analytical task is to distinguish a formal change from its consequences for the text as a whole. I. Espinosa-Zaragoza et al. [24], reviewing research-based simplification tools across languages, target populations, and linguistic levels, likewise demonstrate considerable variation in both the phenomena subjected to simplification and the criteria used to evaluate transformed texts. Such diversity makes a purely quantitative conception of transformation inadequate. N.G. Burmakina [16], for example, reports that automatically simplified texts may continue to contain syntactically complex structures and unjustified lexical substitutions and may distort content despite improvements in selected formal indicators. Taken together, these studies support an operation-based approach in which compression, restructuring, substitution, and explication are evaluated not only as formal procedures but also in terms of their effects on textual organization.

The shift from sentence-level to document-level analysis makes this distinction especially important. L. Vázquez-Rodríguez et al. [12] show that coherence requires separate consideration when simplification extends from isolated sentences to multi-sentence texts; their coherence-aware evaluation demonstrates that document-level transformation introduces problems that sentence-level metrics do not adequately capture. A. Grubišić et al. [13] reach a related conclusion in their analysis of split-and-rephrase procedures: shorter and lexically simpler output may facilitate processing, yet restructuring can weaken coherence, textual unity, or stylistic continuity when relations across sentences are not sufficiently controlled. Narrative material intensifies this problem because relevant dependencies often extend across several clauses, sentences, and episodes. T. Schomacker et al. [14], working with aligned German narrative documents, identify narrative simplification as a particularly underdeveloped area and demonstrate the limitations of existing generation approaches when applied to longer narrative material. A. Farajidizaji et al. [15] add that larger shifts in readability tend to be accompanied by lower semantic and lexical similarity to the source. The degree of transformation therefore needs to be considered together with the preservation of relations required for interpreting the transformed text as a coherent whole.

One such relation is referential cohesion. A. Díaz-Negrillo et al. [17], analyzing written and spoken narratives, show that narrative discourse configuration is associated with the selection of referring expressions and the maintenance of reference across connected structures. Although their study concerns L2 English rather than simplified texts, its relevance lies in the textual mechanism it describes: continuity is not carried by individual lexical items in isolation but by the distribution of referring expressions across discourse. This becomes particularly consequential after syntactic splitting. When a participant previously maintained within a complex sentence must be tracked across two or more



independent sentences, cohesive work may shift to repeated nomination, pronouns, or other explicit forms of reference. Syntactic splitting therefore does more than reduce structural complexity; it may redistribute the linguistic resources through which participant continuity is maintained.

Narrative coherence, however, depends not only on participant continuity but also on the relations connecting events. R.A. Zwaan [23], revisiting the event-indexing model of narrative comprehension, describes narrative events as linked across several dimensions, including entities or protagonists, goals, time, space, and physical causation. This multidimensional account is particularly relevant to transformed narratives because chronological preservation does not necessarily entail preservation of the relations that make successive events intelligible. In the analytical framework adopted here, **event sequence** refers to the ordered progression and functional connection of events within a narrative, whereas **causal coherence** concerns the recoverability of motivational and cause–consequence relations between those events. The two dimensions are related but not interchangeable. Two episodes may remain in their original order while the motive, goal, or causal relation connecting them has been removed; conversely, a compressed version may retain a coherent event configuration when an implicit relation is made linguistically recoverable. The relevant question is therefore not simply whether an event remains present after transformation, but whether its relation to preceding and subsequent events can still be interpreted. From this perspective, causal explication may compensate for contextual information lost through compression or restructuring rather than merely increase the explicitness of the transformed text.

Lexical transformation presents a parallel problem. C.-O. Truică et al. [18] formulate lexical simplification as a reduction in lexical difficulty that should not entail loss of meaning or nuance, while E. Graichen et al. [19] emphasize the importance of contextual information in selecting appropriate lexical replacements. J.A. Ortiz-Zambrano et al. [20] further demonstrate that lexical complexity is associated with a combination of linguistic and contextual characteristics rather than with a single surface feature; their analysis incorporates lexical, semantic, and sequence-related information. These findings make it necessary to distinguish a replacement that preserves the relevant textual function of an expression from one that alters its semantic or discourse contribution. In narrative discourse, a simpler lexical unit may remain denotatively appropriate while losing evaluative force, figurative meaning, characterizing potential, or genre-related associations. Lexical substitution and neutralization of imagery therefore cannot be collapsed into a single replacement category: the distinction depends on what remains functionally available in the transformed narrative, not merely on whether one lexical form has been exchanged for another.

This functional distinction is especially important in folklore and children's narratives, where lexical and formulaic material contributes not only to propositional content but also to genre organization, characterization, evaluation, and plot development. L.S. Tokpayeva et al. [21], in a comparative study of Kazakh, Russian, and English children's folklore, show that lexical and phraseological organization varies across folklore genres and includes both historically transmitted and nationally marked elements. Their findings caution against treating culturally or genre-marked expressions as dispensable ornament. On Kazakhstani short-story material, G.K. Tileubay et al. [22] similarly demonstrate that textual elements may participate simultaneously in plot organization, character construction, comparison, personification, and broader semantic contexts. Consequently, an element may remain referentially recoverable after transformation while part of its narrative or semantic function has nevertheless been reduced or recast.

Formulaic material requires particularly careful treatment because its formal presence cannot be equated with preservation of the narrative's overall communicative organization. E. Karagöz [26], analyzing tale endings in the Tatar folklore tradition, shows that closing formulae constitute a differentiated class of conventional narrative elements and distinguishes several functional types, including instructive closing formulae. This is relevant to Turkic folklore material because it demonstrates that



tale-final formulae have identifiable structural and genre-related status. A folklore formula, however, is not equivalent to pragmatic closure. A closing formula is a conventional linguistic unit occupying a particular position within the textual organization of a narrative, whereas pragmatic closure concerns the communicative effect produced by the completed narrative as a whole. This distinction is also related to plot resolution, which denotes the outcome through which the central narrative problem or conflict reaches its conclusion. Plot resolution and pragmatic closure are therefore functionally connected but analytically distinct: the former concerns what happens to the central problem or conflict, whereas the latter concerns how that outcome, together with the preceding event structure, supports an evaluative, didactic, or other communicative interpretation. A narrative may retain pragmatic closure after a conventional closing formula has been reduced, while preservation of such a formula alone cannot guarantee pragmatic closure if the preceding conflict–action–outcome relation has been disrupted.

The broader research landscape also reveals an important language-specific limitation. I. Espinosa-Zaragoza et al. [24] identify a marked need for simplification resources beyond English, while Y. Veitsman et al. [25] characterize Kazakh, together with other Central Asian Turkic languages, as comparatively under-resourced in contemporary language-processing research and note continuing limitations in linguistic datasets and resources. These studies address computational resources rather than narrative transformation itself, but they help to locate the present linguistic problem: methods developed and evaluated predominantly for high-resource languages cannot substitute for direct analysis of how Kazakh texts reorganize cohesion, lexical-semantic relations, and narrative structure under transformation. At the same time, the Kazakh-related research represented here illuminates individual lexical, folkloric, or text-semantic properties [21, 22] rather than aligned transformations of the same narrative across controlled versions.

The reviewed literature therefore identifies several components of the problem without yet integrating them into a single analytical account. Research on simplification describes transformation operations and the risks associated with increasingly extensive modification [11–16]; studies of cohesion show that referential continuity is distributed across discourse [17]; lexical research establishes contextual and semantic constraints on replacement [18–20]; narrative research conceptualizes events as relationally rather than merely sequentially organized [23]; and studies of children’s folklore and narrative semantics demonstrate that linguistic elements may simultaneously serve referential, evaluative, genre-related, and plot functions [21, 22, 26]. What remains insufficiently described is how these dimensions interact within aligned source and transformed versions of Kazakh narratives. In particular, no segment-level model has yet systematically linked a formally identified transformation operation to its consequences for cohesion, causal coherence, event sequence, and narrative function, nor has it been established which relations may be reduced or recast and which must remain recoverable for a transformed narrative to preserve its functional core. The present study addresses this gap by separating the primary transformation operation from its textual effect and tracing both across aligned original–A1–A2 sequences.

Materials and methods

Research material and corpus formation

The study used a purpose-built corpus of Kazakh folklore-didactic and children’s narratives together with two systematically transformed versions of each source text. Eight source narratives were selected from the open Kazakh-language resources *Balalar Alemi*, *Ertegiler.kz*, and *Adebportal*. Corpus formation was guided by the need to trace structural-semantic transformations within narratives possessing a sufficiently stable event structure to permit segment-level alignment between the source text and its transformed versions.

A text was included if it:



- 1) was written in Kazakh;
- 2) represented a folklore-didactic or children's narrative;
- 3) contained an identifiable progression from an initial situation through event development and a problem or conflict episode to a plot resolution;
- 4) allowed linguistic transformation without reconstruction of the central plot;
- 5) contained features relevant to the analysis, including dialogue, figurative or evaluative vocabulary, conventional folklore formulae, culturally marked units, or syntactically extended constructions.

In accordance with the terminology adopted in this study, folklore formulae refer to conventionalized expressions associated with the organization and genre framing of folklore narrative, including opening or closing formulae where present and other fixed narrative expressions represented in the corpus; conversational discourse formulae are not included in this category. Narratives with highly branched plots, numerous interacting characters, predominantly static description, or strong dependence on an extended mythological context were excluded because such features would reduce the reliability of segment alignment and make it difficult to attribute observed changes to individual structural-semantic operations. The eight source narratives and the linguistic and narrative features that motivated their inclusion in the corpus are summarized in Table 1.

Table 1. Corpus of original Kazakh narrative texts

No.	Source text	Source of the original text	Basis for inclusion in the corpus
1	“Tapqyr qoian”	Balalar Alemi	Dialogue, sequential event development, and a plot resolution with an evaluative implication
2	“Qoian men tasbaqa”	Balalar Alemi	Conflict-based event sequence and contrast between character actions
3	“Änshi qoian”	Balalar Alemi	Characterization, pursuit dynamics, and expressive narration
4	“Maqta qyz ben mysyq”	Adebportal	Chain composition, dialogue, repetition, and a culturally recognizable plot
5	“Ür, toqpaq”	Adebportal / Ertegiler.kz	Conventional folklore formulae, a magical object, and relations between event sequence and causality
6	“Altyn saqa”	Ertegiler.kz	Culturally marked imagery, plot tension, and expressive vocabulary
7	“Tülki, tasbaqa jāne kene”	Ertegiler.kz	Contrastive character actions, dialogic organization, and an evaluative narrative orientation
8	“Jaqsylыq pen Jamandyq”	Ertegiler.kz	Semantic opposition, sequential event development, and a plot resolution with an evaluative implication

As Table 1 indicates, the corpus was deliberately heterogeneous with respect to dialogue, imagery, formulaic language, characterization, and event configuration, while all eight narratives retained an identifiable central event structure. This combination made it possible to compare the effects of the same general transformation categories across texts in which particular linguistic and narrative functions were realized differently.

For each source narrative, one A1 and one A2 version were produced according to the same transformation protocol. The resulting corpus therefore comprised 24 textual units: eight source texts, eight A1 versions, and eight A2 versions. Its organization as eight aligned original–A1–A2 sequences made it possible to trace changes within the same narrative rather than compare unrelated texts. Alignment was used to examine changes in syntactic organization, lexical representation, referential and lexical cohesion, event sequence, causal relations, imagery, and narrative function.

The A1 and A2 versions were created specifically for the study rather than taken from textbooks or previously published graded materials. *Tilqural.kz* was consulted only as an auxiliary reference for the



general organization of Kazakh-language materials by proficiency level and was not used as a source of corpus texts or transformed versions. A1 and A2 were treated analytically as two comparative conditions of structural-semantic transformation. The A1 condition involved a greater degree of formal restructuring and reduction, whereas the A2 condition permitted denser retention of source-text relations and narrative organization. Accordingly, the proficiency labels functioned as conditions under which different transformation profiles could be compared; they were not themselves the primary object of the linguistic analysis.

Transformation protocol and procedure for structural-semantic analysis

The transformation protocol distinguished between variable linguistic form and the functional narrative core that had to remain recoverable for a transformed version to preserve macro-level coherence. The functional narrative core comprised the principal participant or participants, the key event, the motive or causal relation necessary for interpreting that event, the plot resolution, and the pragmatic closure of the narrative. These components were treated as related but analytically distinct. Plot resolution refers to the outcome through which the central narrative problem or conflict is resolved. Pragmatic closure refers to the overall communicative effect of the completed narrative, including an evaluative or didactic implication where such an implication is recoverable from the preceding event structure. The presence of a final evaluative statement was therefore not, by itself, treated as evidence of pragmatic closure if its relation to the preceding events had been disrupted.

Peripheral description, individual dialogic turns, expressive detail, imagery, genre-framing material, and secondary episodes could be compressed, reformulated, or omitted only where their modification did not make a functionally necessary relation within the central narrative structure unrecoverable. This principle distinguished reduction of peripheral material from transformation that disrupted participant continuity, causal motivation, event progression, plot resolution, or the overall communicative effect of the narrative.

The A1 versions employed shorter syntactic units, predominantly direct word order, more frequent lexical items, repeated key nominations, restricted synonymic variation, reduced descriptive material, and limited figurative language. The A2 versions retained a denser event sequence, selected dialogue, evaluative vocabulary, and a greater proportion of genre-related and expressive material. These characteristics were treated as transformation parameters rather than as independent measures of proficiency. The contrast between A1 and A2 was therefore examined not simply through differences in text length, but through the extent to which cohesive resources, causal relations, event connections, and narrative functions were retained, reconstructed, made explicit, reduced, or recast.

The basic unit of analysis was the transformational segment, defined as a lexeme, phrase, clause, sentence, dialogic utterance, or plot micro-fragment that underwent an identifiable change between the source text and at least one transformed version. Each narrative was first segmented according to its functional organization, including exposition, event development, dialogue, description, culmination, and plot resolution. Corresponding fragments were then aligned across the original, A1, and A2 versions. Comparison proceeded in three directions: original–A1, original–A2, and A1–A2. Across the eight aligned narrative sequences, 612 transformational segments were identified.

The analysis proceeded at two levels. At the first level, each transformational segment was assigned one primary code: structural compression, syntactic simplification, lexical substitution, semantic reduction, causal explication, or neutralization of imagery. These six categories describe operations performed on textual material. Preservation of the functional narrative core was recorded separately as a control category of functional continuity rather than as a seventh transformation operation. This distinction ensured that preservation of a narrative relation was not treated as formally equivalent to an operation that modified the source text.

At the second level, the textual consequence of each transformation was assessed across three interconnected dimensions. First, cohesive organization was coded as retained, weakened, or reconstructed



with reference to referential continuity and lexical linkage. Second, causal coherence and event sequence were assessed as distinct dimensions. Causal coherence concerned whether a motive or cause–consequence relation remained recoverable, was made explicit or reconstructed, or was weakened by transformation. Event sequence concerned whether the ordered and functionally relevant progression of events was preserved, compressed, or disrupted. This distinction was necessary because chronological continuity could survive even where the causal relation connecting adjacent events had become less recoverable. Third, narrative functions, including genre framing, characterization, conflict presentation, event progression, plot resolution, and pragmatic closure, were classified as retained, reduced, or recast.

The distinction between the two analytical levels was essential to the coding procedure because formally similar operations could produce different textual consequences depending on their position and function. Syntactic splitting, for example, could preserve propositional content while transferring referential linkage from an intra-sentential structure to repeated nomination or pronominal reference across sentences. Causal explication could reconstruct a motive or cause–consequence relation that had become less inferable after compression. The removal of a conventional folklore formula could reduce genre-framing function without disrupting the central event sequence, whereas deletion of a motivational episode could leave chronological ordering intact while weakening causal coherence. Similarly, preservation of a closing formula was not automatically interpreted as preservation of pragmatic closure, because the latter depended on the functional relation between the preceding event structure, plot resolution, and the communicative effect of the completed narrative.

Quantitative indicators – word and sentence counts, average sentence length, numbers of retained and deleted segments, and frequencies of the six transformation operations – were used to establish the scale and distribution of change. These measures were interpreted together with segment-level evidence concerning cohesive organization, causal coherence, event sequence, and narrative function. The procedure therefore treated transformation as a multilevel reorganization of textual relations rather than as linear simplification measured primarily by reduction in length or sentence complexity.

Coding matrix and parameters for comparing textual transformations

A unified coding matrix was used to record both the linguistic operation applied to each transformational segment and its consequence for textual organization. Segments aligned across the original, A1, and A2 versions received one primary operation code. The matrix therefore separated the formal type of change from its effects on cohesive organization, causal coherence, event sequence, and narrative function. The operational criteria, boundary rules, and text-organizational effects used in coding are presented in Table 2.

As Table 2 shows, the coding scheme was designed to prevent formally similar changes from being treated as functionally equivalent. The primary code identified what operation occurred, whereas the secondary assessment established what relation or narrative function was affected.

Preservation of the functional narrative core was assessed separately from the six primary transformation operations and was excluded from operation-frequency counts. For each aligned narrative sequence, the analysis established whether the principal participant or participants, key event, motive or causal relation, plot resolution, and pragmatic closure remained recoverable after transformation. Plot resolution was treated as the outcome of the central narrative problem or conflict, whereas pragmatic closure referred to the overall communicative effect recoverable from the completed event structure. Preservation of these relations indicated macro-level functional continuity; disruption of a relation necessary for interpreting the narrative was treated as evidence of semantic distortion.

Coding consistency and verification of analytical decisions

All aligned sequences were coded and subsequently reviewed using the same matrix. Borderline cases were resolved through a fixed decision sequence: complete removal of a fragment; syntactic restructuring with the principal proposition retained; denotatively equivalent lexical replacement; loss



Table 2. Coding matrix of structural-semantic transformations

Type	Mandatory criterion	Boundary and priority rule	Text-organizational effect
Structural compression	Complete removal or condensation of a descriptive, explanatory, dialogic, or secondary event fragment, including an independent proposition or plot micro-episode	If the segment remains represented but loses only a semantic component, the case is coded as semantic reduction. If complete removal also eliminates an expressive or genre-framing function, structural compression remains the primary code and the functional loss is recorded as a secondary effect	Reduces episodic density and may weaken cohesive or causal relations or compress the event sequence while leaving the central narrative structure recoverable
Syntactic simplification	Splitting, restructuring, or linearization of a clause or sentence while its principal propositional content is retained	Loss of an independent proposition excludes this code and is classified as structural compression. Loss of a semantic component within a retained proposition is classified as semantic reduction	Redistributes temporal, referential, or logical relations across clauses or sentences and may require reconstruction of inter-sentential cohesion
Lexical substitution	Replacement of a lexical item by a denotatively equivalent or near-equivalent unit while reference, evaluation, and textual function remain substantially stable	If replacement weakens figurative, evaluative, emotional, characterizing, or genre-marking meaning, the case is coded as neutralization of imagery	Retains or reconstructs lexical and referential continuity, often through reduced lexical variation or more transparent nomination
Semantic reduction	Loss of a semantic feature, qualification, or meaning component within a segment whose principal proposition and textual position remain represented	Complete removal of an independent proposition, utterance, or functional fragment is coded as structural compression. Syntactic restructuring without semantic loss is coded as syntactic simplification	Reduces semantic density and may weaken local motivation, characterization, evaluation, or explanatory completeness
Causal explication	A previously implicit motive or cause–consequence relation becomes overtly encoded through a connector, causal construction, or explicit formulation	Rewording, sentence splitting, or chronological sequencing without increased causal or motivational explicitness does not qualify for this code	Reconstructs or strengthens causal coherence where contextual support has been weakened by compression or restructuring
Neutralization of imagery	A retained segment undergoes lexical or syntactic reformulation that weakens figurative, evaluative, emotional, characterizing, or genre-marking meaning	Complete deletion of the relevant fragment is coded as structural compression, with functional loss recorded as a secondary effect. If expressive and textual functions remain substantially stable, the case is coded as lexical substitution	Reduces or recasts genre framing, characterization, emotional intensity, conflict presentation, or another expressive narrative function

of a semantic component within a retained segment; explicit encoding of a previously implicit causal relation; and weakening of figurative, evaluative, characterizing, or genre-framing meaning. The first applicable criterion determined the primary operation, except where lexical replacement demonstrably reduced an expressive or genre-related function, in which case neutralization of imagery took precedence over lexical substitution.

Only one primary operation entered the frequency counts. Additional consequences were recorded as secondary textual effects and were not counted as separate operations. Thus, sentence splitting could be coded as syntactic simplification while also being marked as reconstructing referential cohesion through repeated nomination. This procedure prevented double counting while preserving information about multifunctional textual effects. The second review concentrated on the two principal boundary areas: structural compression versus semantic reduction and lexical substitution versus neutralization of imagery. Quantitative distributions were then checked against the segment-level interpretations to ensure that frequency patterns did not contradict the qualitative coding decisions.



Acceptability of a transformed segment was evaluated against the recoverability of the functional narrative core. A transformation was considered functionally continuous when the relations linking participant, key event, motive or causality, plot resolution, and pragmatic closure remained interpretable in the transformed narrative. Disruption of a necessary motivational or causal relation, breakdown of the event sequence, alteration of the evaluative orientation, or detachment of the final outcome from the preceding event logic was treated as semantic distortion. Importantly, preservation of a final evaluative or formulaic statement alone was not regarded as evidence of pragmatic closure if its functional relation to the preceding events had been lost.

The procedure thus converted the six primary transformation operations and the separate control category of functional narrative-core preservation into a reproducible analytical model linking local linguistic change to cohesive organization, causal coherence, event sequence, and narrative function.

Results

Formal-structural reconfiguration and cohesive continuity in the original, A1, and A2 narratives

The first stage of the analysis examined how measurable formal changes were related to the reorganization of narrative structure. Across the eight aligned original–A1–A2 sequences, total text length and average sentence length decreased substantially, whereas sentence count declined less sharply. This asymmetry indicates that transformation involved not only compression but also the redistribution of relations previously encoded within extended syntactic units across shorter sentences. Table 3 presents the formal-structural profile of the corpus.

Table 3. Formal-structural parameters of the original, A1, and A2 versions of the narrative texts

Parameter	Original texts	A1 versions	A2 versions	Identified pattern
Total length, words	5184	2248	3496	Strongest compression in A1
Average length of one text, words	648.0	281.0	437.0	A2 retains a larger proportion of the narrative
Total number of sentences	232	184	216	Textual reduction is accompanied by syntactic redistribution
Average sentence length, words	22.3	12.2	16.2	Strongest reduction of intra-sentential complexity in A1
Retained segments, %	100.0	56.4	73.1	A2 preserves more source-text relations and episodes
Deleted segments, %	0.0	43.6	26.9	A1 removes more peripheral and functionally secondary material

Measured by word count, the A1 corpus was reduced by 56.6% and the A2 corpus by 32.6% relative to the source texts. These values should be distinguished from the segment-retention percentages in Table 3: the former describe reduction in textual volume, whereas the latter indicate the proportion of aligned source segments retained or deleted. The two transformation conditions therefore cannot be treated simply as proportional degrees of shortening. A1 removed more description, secondary repetition, dialogic material, and expressive detail, producing a more restricted event sequence centered on the principal participants, key event, and plot resolution. A2 retained more linking episodes, causal relations, and dialogue, thereby preserving a denser network of relations supporting both event sequence and causal coherence.

The relation between word count and sentence count further clarifies the structural effect of transformation. Average sentence length decreased from 22.3 words in the source narratives to 12.2 in A1 and 16.2 in A2, while the total number of sentences declined from 232 to 184 and 216, respectively. The comparatively smaller reduction in sentence count reflects syntactic splitting and restructuring:

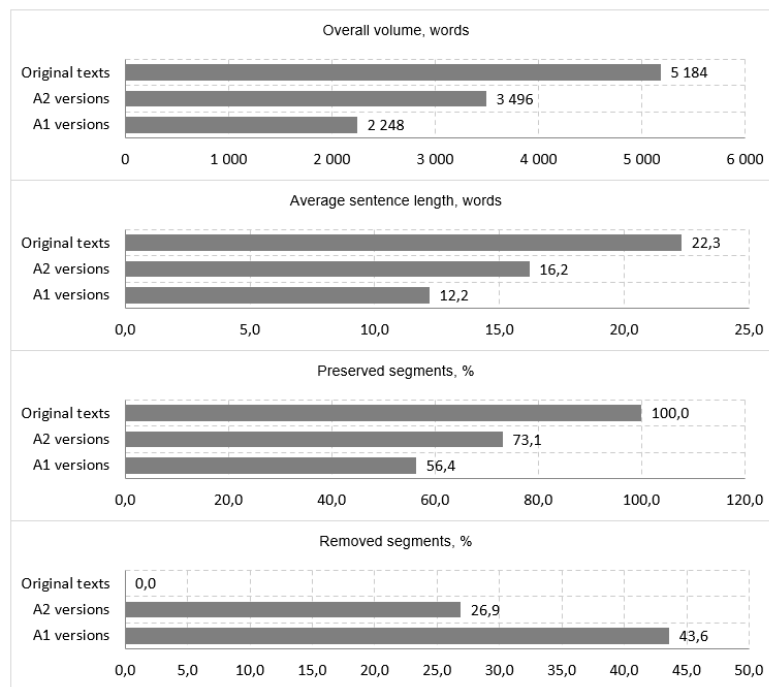


Fig. 1. Compression profile of the original, A1, and A2 versions of the narrative texts

relations previously integrated within extended constructions were redistributed across shorter sentences. This tendency was strongest in A1, where repeated key nominations frequently maintained referential continuity after clause-internal links had been redistributed. Syntactic simplification therefore shifted part of the cohesive load from intra-sentential structure to inter-sentential reference rather than merely reducing grammatical complexity.

The segment data reinforce this distinction. A1 retained 56.4% of source segments and deleted 43.6%, whereas A2 retained 73.1% and deleted 26.9%. More extensive compression in A1 consequently required greater reconstruction of textual continuity through direct succession, repeated nomination, and explicit participant tracking. Because A2 preserved a larger proportion of source segments, more of the original connections between participants, actions, and adjacent episodes remained textually available. The contrast is therefore not only quantitative: the two conditions differ in how cohesive continuity is maintained after structural reduction.

To show how these differences converge across several formal indicators, Fig. 1 compares overall textual volume, average sentence length, and the proportions of retained and deleted segments in the original, A1, and A2 versions.

Taken together, the indicators in Fig. 1 reveal two distinct profiles of structural reorganization. A1 combines stronger compression with greater syntactic redistribution and a lower proportion of retained source segments, making continuity more dependent on reconstructed inter-sentential links. A2 remains formally reduced relative to the source texts but preserves a denser network of source-text relations. The difference between the two conditions thus concerns not simply how much linguistic material is retained, but how cohesion, causal coherence, and event sequence are maintained under different degrees of structural-semantic transformation.

Distribution and text-organizational effects of structural-semantic transformations in A1 and A2 versions

The second stage of the analysis examined the distribution of the coded operations and their consequences for textual organization. Across the eight aligned original–A1–A2 sequences, the dataset comprised 612 coded segment records: 356 in the A1 versions and 256 in A2. Within these records,

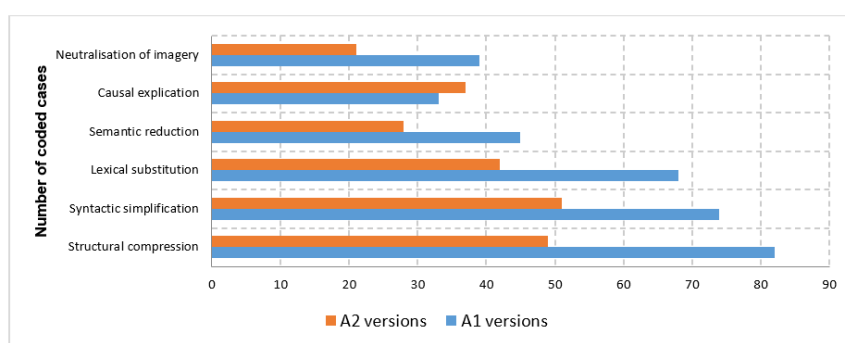


Fig. 2. Frequency of the six primary structural-semantic transformation operations in the A1 and A2 versions

the six primary transformation operations accounted for 341 A1 and 228 A2 cases, while preservation of the functional narrative core was registered separately in 15 and 28 control cases, respectively. This distinction is essential because the control category records functional continuity rather than an operation performed on textual material.

Structural compression, syntactic simplification, and lexical substitution were the most frequent primary operations in both transformation conditions and occurred more often in A1. Causal explication showed the opposite pattern, with a slightly higher absolute frequency in A2 despite the smaller number of coded records at that level. To keep operation frequencies analytically distinct from the control assessment, Fig. 2 presents only the six primary structural-semantic operations.

Structural compression accounted for 82 A1 cases and 49 A2 cases. In A1, it most often affected peripheral description, secondary repetition, parts of dialogue, and episodic detail, reducing the number of intermediate links between events. The event sequence remained recoverable, but continuity more often depended on direct succession and repeated nomination. In A2, less extensive compression left more episodic links available to connect actions and successive stages of the narrative. Structural compression therefore differed between the two conditions not only in frequency but also in the amount of cohesive reconstruction required after deletion.

Syntactic simplification was recorded in 74 A1 and 51 A2 cases. A1 more frequently redistributed propositions from extended constructions across two or more shorter sentences, shifting part of the cohesive load from clause-internal syntax to inter-sentential reference. A2 more often retained sequential actions within a single construction and required less extensive redistribution of their syntactic relations. The operation thus affected the location of cohesive relations within the text rather than merely reducing sentence-level complexity.

Lexical substitution occurred in 68 A1 and 42 A2 cases. Under the coding criteria, this category was restricted to replacements in which denotative reference and relevant textual function remained substantially stable. In A1, more transparent and recurrent nominations frequently supported referential continuity by reducing lexical variation across adjacent sentences. Replacements that weakened figurative, evaluative, emotional, characterizing, or genre-marking meaning were not included in this category but were coded as neutralization of imagery. The lower frequency of lexical substitution in A2 corresponded to greater retention of source-text nominations.

Semantic reduction was identified in 45 A1 and 28 A2 cases. Unlike structural compression, it involved the loss of a semantic component within a segment whose principal proposition remained represented. Reduced material most often concerned additional character attributes, descriptive qualifications, or explanatory detail. In A1, the cumulative effect was lower semantic density and a more linear representation of events; in A2, semantic reduction was more selective. The functional consequence depended on the status of the omitted component: peripheral detail could be reduced without



disrupting narrative continuity, whereas loss of a motivational component could weaken causal coherence by making the relation between action and consequence less recoverable.

Causal explication was the only primary operation more frequent in A2, with 37 cases compared with 33 in A1. Although A2 underwent fewer coded changes overall, it more often made an implicit motive or cause–consequence relation linguistically explicit. In A1, event sequence was frequently maintained through direct linear succession; where a causal result was explicitly marked, connectors such as *sondyqtan* (‘therefore’, ‘for that reason’) provided an overt causal cue. By contrast, *sodan keiin* (‘after that’) marks temporal succession and therefore contributes primarily to explicit event sequencing rather than, by itself, to causal coherence. This distinction is important because chronological ordering alone was not coded as causal explication. Across the corpus, causal explication functioned as a compensatory mechanism where compression or restructuring had weakened the contextual basis from which a motivational or cause–consequence relation could otherwise be inferred.

Neutralization of imagery was recorded in 39 A1 and 21 A2 cases. These transformations involved a demonstrable weakening of figurative, evaluative, emotional, characterizing, or genre-marking meaning rather than denotatively equivalent lexical replacement. In A1, expressive verbs, emotionally intensified utterances, and folklore formulae were more frequently neutralized. Such changes generally left the event sequence recoverable but reduced genre framing, characterization, conflict intensity, or another expressive narrative function. A2 retained these functions more often, particularly where expressive material contributed simultaneously to event representation and characterization.

Preservation of the functional narrative core was assessed separately from the six operations and was recorded in 15 A1 and 28 A2 control cases. These values are therefore not transformation-operation frequencies. The control assessment concerned the recoverability of the functional relation linking the principal participant, key event, motive or causal relation, plot resolution, and pragmatic closure. Plot resolution and pragmatic closure remained analytically distinct: the former concerns the outcome of the central narrative problem or conflict, whereas the latter concerns the overall communicative effect produced by the completed narrative. The higher number of control cases in A2 corresponds to its greater retention of source-text relations, while A1 more often reduced the same functional chain to a minimally elaborated sequence.

Taken together, the distribution reveals two different profiles of structural-semantic reorganization. A1 relies more heavily on compression, syntactic redistribution, and lexical substitution, with a greater reduction of expressive and genre-related functions. A2 preserves more source-text relations and uses causal explication more frequently to maintain the recoverability of motivational and cause–consequence relations. The six operations therefore constitute not a simple inventory of surface changes but distinct mechanisms whose textual effects emerge in cohesive organization, causal coherence, event sequence, and narrative function.

Segment-level transformation and the reorganization of cohesion, causal coherence, event sequence, and narrative function

The corpus-level patterns were further examined through aligned transformational segments. The purpose of this analysis was not merely to illustrate the six primary operations, but to establish how a locally identifiable structural-semantic change affected referential and lexical cohesion, causal coherence, event sequence, and narrative function. The selected segments represent different narrative configurations in the corpus, including animal narratives, chain-structured folklore, magical tales, and didactic narratives such as “Tapqyr qoian”, “Qoian men tasbaqa”, “Maqta qyz ben mysyq”, and “Altyn saqa”.

In accordance with the two-level coding procedure, Table 4 assigns one primary structural-semantic operation to each illustrated transformation, while its consequences for textual organization are analyzed separately. English translations are provided for all Kazakh-language examples; morpheme-level glosses are added where the interpretation depends on a specific grammatical construction.



Table 4. Segment-level structural-semantic transformations and their textual effects

Source segment	A1 version	A2 version	Primary operation illustrated	Effect on textual organization
Qoian aňdardy basynan aiağyna deñin aralap shyğyp, bärine birdei körsetem degenshe qazandağy su suyp qalady. <i>‘By the time the rabbit has gone around all the animals to show the water to each of them, the water in the pot cools down.’</i>	Qoian sudy aňdarğa körsetti. Su suydı. <i>‘The rabbit showed the water to the animals. The water cooled.’</i>	Qoian sudy barlyq aňğa körsetip shyqqansha, su suyp qaldy. <i>‘By the time the rabbit had shown the water to all the animals, it had cooled.’</i>	Syntactic simplification	A1 redistributes the temporal relation across two independent sentences and maintains continuity through lexical recurrence. A2 retains the temporal dependency within a single construction; <i>-qanşa</i> in <i>shyq-qan-sha</i> marks the terminative relation ‘until/by the time’.
Al qulyğyn asyrğan tapqyr qoian suyp qalğan sudy basyna biraq köterip, sımırıp salady. <i>‘But the resourceful rabbit, having outwitted them, lifts the cooled water and gulps it down.’</i>	Qoian suyp sudy ıstı. <i>‘The rabbit drank the cold water.’</i>	Tapqyr qoian suyp qalğan sudy ıstıp qoıdy. <i>‘The resourceful rabbit drank the cooled water.’</i>	Neutralization of imagery	A1 removes the expressive manner of the action and the characterizing nomination <i>tapqyr</i> ‘resourceful/clever’. A2 also neutralizes the expressive action sequence but retains <i>tapqyr</i> , preserving part of the characterizing function.
Erte, erte ertede, eşki jünü börtede... <i>‘Long, long ago, in ancient times, when the goat’s fleece was mottled ...’</i> [conventional folklore opening]	Bır küni qoian tasbaqany kördi. <i>‘One day the rabbit saw the tortoise.’</i>	Erte kezde bır qoian bolypty. Bır küni ol tasbaqany köripti. <i>‘Long ago there lived a rabbit. One day it saw a tortoise.’</i>	Structural compression	A1 removes the folklore opening completely, eliminating its genre-framing function. A2 retains a reduced folklore frame through <i>Erte kezde</i> ‘long ago’. Genre-framing loss or recasting is recorded as a secondary textual effect rather than as an additional operation.
Qoiannyñ oına bı aram oı kelipti. <i>‘A malicious thought came into the rabbit’s mind.’</i>	Qoian jaman oı oilady. <i>‘The rabbit had a bad thought.’</i>	Qoiannyñ oına jaman oı keldi. <i>‘A bad thought came into the rabbit’s mind.’</i>	Lexical substitution	The evaluatively marked <i>aram</i> ‘malicious/wicked’ is replaced by the more general <i>jaman</i> ‘bad’. Negative evaluation and the referential function remain recoverable; A2 additionally retains the source pattern of internal-state representation.
Maqta qyz üin jinap jürüp, bır meiz taup alady da, mysyqty şaqyrady. <i>‘While cleaning her house, Maqta qyz finds a raisin and calls the cat.’</i>	Maqta qyz meiz taptı. Ol mysyqty şaqyrdy. <i>‘Maqta qyz found a raisin. She called the cat.’</i>	Maqta qyz üi jinap jürüp, meiz taup aldy da, mysyqty şaqyrdy. <i>‘While cleaning the house, Maqta qyz found a raisin and called the cat.’</i>	Syntactic simplification	A1 distributes the sequential actions across separate sentences and reconstructs participant continuity through <i>ol</i> ‘3SG’ (‘she’ in this context). A2 retains clause-internal linkage between the actions.
— Kelmeseñ kelme! — dep, meizdı özi jep qoiady. <i>“‘If you do not come, then do not!’ she says and eats the raisin herself.’</i>	Mysyq kelmedi. Qyz meizdı özi jedi. <i>‘The cat did not come. The girl ate the raisin herself.’</i>	Mysyq kelmegen soñ, Maqta qyz meizdı özi jep qoidy. <i>‘Since the cat did not come, Maqta qyz ate the raisin herself.’</i>	Causal explication	A2 replaces an inferable relation with overt causal marking through <i>kel-me-gen soñ</i> (come-NEG-PTCP since/after) ‘since [the cat] did not come’. The source dialogic sequence is thereby recast as explicitly causally related narration.



End of Table 4

Source segment	A1 version	A2 version	Primary operation illustrated	Effect on textual organization
— Endeşe qatyğyňdy tögem! — deidi. “ <i>Then I will spill your qatyq [fermented milk]!</i> ” [the cat] says.’	Mysyq aşulandy. ‘ <i>The cat became angry.</i> ’	Mysyq aşulanyňp, qatyqy tögetiniň aitty. ‘ <i>The cat became angry and said that it would spill the fermented milk.</i> ’	Semantic reduction	A1 retains the emotional state but removes the proposition expressing the threat. A2 recasts direct speech as indirect narration while preserving the threat as a conflict-organizing function.
Baidyň jylqylary osy ökpeden ürküp, su işpei qoiypty. ‘ <i>The bai’s horses were frightened by the lung and would not drink the water.</i> ’	Jylqylar qoryqty. Olar su işpedi. ‘ <i>The horses were frightened. They did not drink the water.</i> ’	Baidyň jylqylary ökpeden qorqyp, su işpei qoidy. ‘ <i>The bai’s horses, frightened by the lung, did not drink the water.</i> ’	Syntactic simplification	A1 separates the initiating condition and its consequence into adjacent sentences while maintaining referential cohesion through the nominal—pronominal chain <i>jylqylar—olar</i> ‘horses—they’. A2 retains the relation within one syntactic construction.
Bai olardy qanşa aidasa da, suğa jaqyndatu mümkin bolmaidy. ‘ <i>However much the bai drove them on, he could not bring them close to the water.</i> ’	Bai jylqylardy suğa apara almany. ‘ <i>The bai could not take the horses to the water.</i> ’	Bai jylqylardy suğa jaqyndatqysy keldi, biraq olar barmady. ‘ <i>The bai wanted to bring the horses closer to the water, but they would not go.</i> ’	Causal explication	A2 overtly separates intention from resistance. The desiderative construction <i>jaqynda-t-qy-sy kel-di</i> (near-CAUS-DESID-3. POSS AUX-PST) encodes the character’s intention, while <i>biraq</i> ‘but’ marks the contrast with the animals’ refusal.
Ökpe suğa batyp ketedi de, köp ūzamai jalmaýyz kempir bolyp shyğady. ‘ <i>The lung sinks into the water and soon emerges as a jalmaýyz kempir [ogress].</i> ’	Ökpe suğa batty. Keiň kempir shyqty. ‘ <i>The lung sank into the water. Then an old woman emerged.</i> ’	Ökpe suğa batyp ketti, sodan keiň jalmaýyz kempir bolyp shyqty. ‘ <i>The lung sank into the water; after that, a jalmaýyz kempir [ogress] emerged.</i> ’	Syntactic simplification	A1 linearizes the event sequence through sentence succession and generalizes the folklore nomination. A2 explicitly marks temporal succession through <i>sodan keiň</i> ‘after that’ and retains the plot-relevant folklore nomination <i>jalmaýyz kempir</i> . The marker is temporal rather than causal and is therefore analyzed as an event-sequence cue, not as causal explication.
Qyz: “Men qüiryğyňdy kesip alamyn!” — deidi. ‘ <i>The girl says, “I will cut off your tail!”</i> ’	Qyz aşulandy. ‘ <i>The girl became angry.</i> ’	Qyz aşulanyp, mysyqy jazalaitynyň aitty. ‘ <i>The girl became angry and said that she would punish the cat.</i> ’	Semantic reduction	A1 removes the specific threatening proposition and retains only the emotional state. A2 generalizes the threat as punishment and recasts direct speech as indirect narration while retaining its conflict-organizing function.

Note. Preservation of the functional narrative core was assessed independently and was not treated as a primary structural-semantic operation. It is therefore neither represented as a separate operation in Table 4 nor included among the six operation types. The control assessment concerns the recoverability of the principal participant, key event, necessary motive or causal relation, plot resolution, and pragmatic closure. The case in which the rabbit avoids danger through ingenuity is considered in the following subsection: both transformed versions preserve the relation between the protagonist’s character and the plot outcome, while A2 retains the more specific characterizing nomination *tapqyr* ‘resourceful/clever’. Abbreviations used in the grammatical glosses: NEG = negation; PTCP = participle; CAUS = causative; DESID = desiderative; 3.POSS = third-person possessive; PST = past; AUX = auxiliary.

The segment-level evidence shows that syntactic simplification changes the distribution of textual relations rather than merely reducing sentence length. In the source narratives, temporal, sequential, and referential relations may be integrated within a single extended construction; A1 frequently redistributes them across adjacent sentences. In *Maqta qyz meiz tapty. Ol mysyqy şaqyrdy* (‘Maqta qyz found a raisin. She called the cat’), the third-person pronoun *ol* reconstructs referential continuity after the source construction has been split. Similarly, *Jylqylar qoryqty. Olar su işpedi* (‘The horses were



frightened. They did not drink the water’) replaces clause-internal participant continuity with the nominal–pronominal chain *jylqylar–olar* (‘horses–they’). Syntactic simplification thus shifts part of the cohesive load from intra-sentential organization to explicit inter-sentential reference.

The lexical evidence confirms the boundary between lexical substitution and neutralization of imagery established in the coding matrix. In *aram oı → jaman oı* (‘malicious thought’ → ‘bad thought’), the lexical choice becomes less specific, but negative evaluation and the relevant referential function remain recoverable; the change is therefore classified as lexical substitution. By contrast, *Qoian suyq sudy ıstı* (‘The rabbit drank the cold water’) removes both the expressive manner of drinking and the explicit characterizing nomination *tapqyr* ‘resourceful/clever’ present in the source segment. The loss concerns expressive and characterizing function rather than denotative reference and is therefore coded as neutralization of imagery. Structural compression of the folklore formula *Erte, erte ertede, eşki jüm börtede...* (‘Long, long ago...’) produces a different secondary effect: the event sequence remains recoverable, but the conventional genre frame is removed.

Causal explication is distinguished from the mere explicit ordering of events. In *Mysyq kelmegen soñ, Maqta qyz meıızdı özi jep qoidy* (‘Since the cat did not come, Maqta qyz ate the raisin herself’), the construction *kel-me-gen soñ* explicitly encodes the relation between the cat’s absence and Maqta qyz’s subsequent action. Likewise, *Bai jylqylardy suğa jaqyndatqysy keldi, biraq olar barmady* (‘The bai wanted to bring the horses closer to the water, but they would not go’) overtly represents intention and the event that prevents its realization. In both cases, causal or motivational information that had depended more strongly on contextual inference becomes linguistically recoverable. By contrast, *sodan keıın* ‘after that’ in the “Altyn saqa” segment marks succession in the event sequence but does not itself encode a causal relation. This contrast maintains the analytical distinction between causal coherence and chronological or sequential continuity.

Semantic reduction in dialogic material further demonstrates that propositional preservation and functional preservation are not equivalent. The direct threat *Men qııryğyñdy kesip alamyn!* (‘I will cut off your tail!’) simultaneously contributes propositional content, emotional intensity, and conflict progression. A1 reduces the segment to *Qyz aşulandy* (‘The girl became angry’), preserving affect but removing the threatening proposition and its speech-act function. A2 generalizes the threat as *mysyqty jazalaitının aıttı* (‘said that she would punish the cat’), thereby reducing lexical and expressive specificity while retaining the conflict-organizing function. The corresponding *qatyq* example follows the same pattern: the A1 version retains anger alone, whereas A2 preserves the threat through indirect narration.

Taken together, the segment-level evidence identifies three corpus-wide patterns. First, syntactic splitting redistributes cohesive relations from clause-internal organization to inter-sentential reference. Second, causal explication can restore the recoverability of motivational or cause–consequence relations without being conflated with explicit event sequencing. Third, lexical, semantic, and expressive reduction may leave the event sequence intact while reducing or recasting characterization, genre framing, conflict presentation, or another narrative function. A1 relies more strongly on repeated reference and linear sequencing after local reduction, whereas A2 retains a denser network of source-text relations and more often preserves or explicitly reconstructs causal links. Structural-semantic transformation therefore operates as a reorganization of the linguistic resources supporting cohesion, causal coherence, event sequence, and narrative function rather than as a uniform movement from formally complex to formally simple text.

Preservation of the functional narrative core and the limits of structural-semantic transformation

The final stage of the analysis examined which textual relations remained functionally stable across the original–A1–A2 sequences and at what point local transformation began to compromise narrative integrity. The components differed in their tolerance to reduction or recasting. Description, imagery, dialogue, and culturally marked detail could undergo substantial modification, whereas coherent



transformed versions retained a recoverable functional chain linking the principal participant, key event, necessary motive or causal relation, plot resolution, and pragmatic closure. The functional narrative core is therefore understood not as an abstract residue of “main meaning”, but as a relational invariant whose recoverability supports macro-level coherence.

Participant continuity formed the first component of this invariant. Plot-determining characters remained identifiable even when their descriptive, evaluative, or lexical representation changed. The rabbit (*qoian*) retained its role as the protagonist whose actions determine the outcome; the tortoise (*tasbaqa*) remained a contrastive participant; Maqta qyz and the cat (*mysyq*) continued to organize the chain conflict; and the *jalmaýyz kempir* (‘ogress’) remained the source of threat in “Altyn saqa”. Individual nominations could be repeated, simplified, or semantically neutralized without altering these narrative roles. Removing a plot-determining participant, by contrast, would disrupt referential continuity and the functional organization of the event sequence. Character representation was therefore more susceptible to transformation than character function.

The key event showed comparable stability. None of the coherent A1 or A2 versions eliminated an action on which the subsequent event sequence depended. A1 frequently reduced such events to shorter propositions and redistributed associated information across adjacent sentences, whereas A2 retained more circumstantial, dialogic, and sequential material. The contrast concerned the linguistic elaboration of the event rather than its functional presence within the narrative. Event preservation thus operated at the level of narrative organization, while sentence length, lexical density, and the amount of descriptive material remained variable.

Causal coherence constituted a more sensitive boundary. Source narratives frequently made motivation or cause–consequence relations recoverable through event succession, dialogue, or contextual inference rather than through overt causal marking. Transformation could alter the linguistic realization of these relations without eliminating them. A1 more often distributed related events across short propositions and relied on linear event sequencing and contextual adjacency, whereas A2 more often retained the contextual basis of the relation or made it explicit through causal explication. The distinction between causal coherence and event sequence is crucial here: preservation of chronological succession did not by itself guarantee that the motive or cause connecting two events remained recoverable. Once a functionally necessary motivational or cause–consequence relation was lost, the propositions could remain grammatically connected and chronologically ordered while no longer preserving equivalent causal coherence.

Plot resolution and pragmatic closure formed the final two components of the functional narrative core but were assessed separately. Plot resolution concerned whether the outcome of the central problem or conflict remained recoverable after transformation. Pragmatic closure, by contrast, concerned the overall communicative effect of the completed narrative, including an evaluative or didactic implication where such an implication followed from the preceding event structure. A concluding moral or evaluative statement was therefore insufficient on its own to establish pragmatic closure. If the relation linking conflict, action, consequence, and plot resolution had been disrupted, a final evaluative sentence would function as an externally appended comment rather than as a communicative effect generated by the narrative as a whole. The relevant criterion was consequently not the formal presence of a didactic conclusion, but its functional derivability from the preceding narrative structure.

Imagery, expressive detail, dialogue, and folklore formulae were comparatively more variable. Their reduction did not necessarily compromise macro-level coherence, although it could reduce or recast genre framing, characterization, conflict intensity, or expressive function. Removal of an opening folklore formula, for example, left the event sequence recoverable while weakening the conventional folklore frame. Similarly, recasting a direct threat as indirect narration could preserve its role in conflict progression while reducing its illocutionary and expressive force. The same distinction applies to folklore formulae more generally: the presence or absence of a conventional opening or closing formula

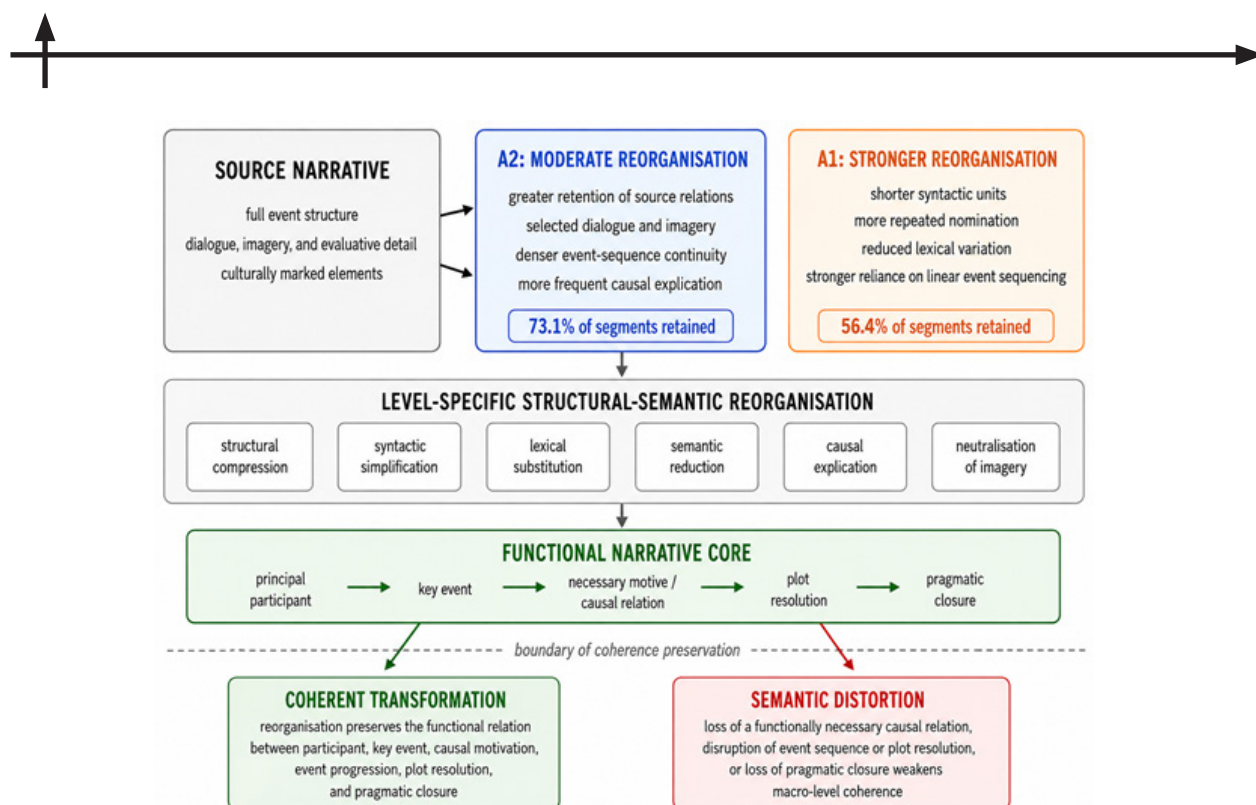


Fig. 3. Functional narrative core and the boundary of structural-semantic transformation in Kazakh A1/A2 narratives

was not equated with pragmatic closure, which depended on the communicative interpretation supported by the narrative as a whole. A2 retained or recast these secondary functions more often, whereas A1 more frequently preserved only the relations necessary for narrative continuity.

These patterns establish the boundary between structural-semantic transformation and semantic distortion. Reduction remained compatible with narrative integrity when participant continuity, the key event, the functionally necessary causal relation, the event sequence, plot resolution, and pragmatic closure remained recoverable. Transformation crossed this boundary when it removed a plot-determining participant, made the motive or cause of a decisive action unrecoverable, disrupted a necessary event connection, altered the plot outcome or its evaluative orientation, or detached the overall communicative effect from the event structure from which it should arise.

Fig. 3 synthesizes this corpus-based model by showing how the six structural-semantic operations may reorganize source-text material to different degrees while preservation of the functional narrative core remains a separate constraint on acceptable transformation.

The model emphasizes that narrative integrity is not proportional to the quantity of verbal material retained. Extensive compression can preserve macro-level coherence when the relational chain linking participant, key event, necessary motive or causal relation, event progression, plot resolution, and pragmatic closure remains reconstructable. Conversely, a formally fuller version may still weaken narrative integrity if a functionally necessary relation is disrupted. A1 and A2 therefore represent different configurations of textual reorganization rather than merely different degrees of shortening: A1 relies more heavily on reconstructed referential continuity and a more linear event sequence after extensive compression, whereas A2 preserves a denser network of source-text relations and more frequently makes causal relations explicit while retaining additional genre-expressive and narrative functions.



Discussion

The findings show that structural-semantic transformation of narrative text is not a linear reduction of linguistic complexity but a reorganization of textual relations. Across the aligned original–A1–A2 sequences, changes in text length and syntax were accompanied by redistribution of cohesive resources, changes in causal recoverability, modification of the event sequence, and reduction or recasting of narrative functions. This supports operation-based approaches to simplification [1, 11] and document-level research showing that locally successful sentence simplification does not necessarily preserve discourse organization [12, 27–29].

The first major result concerns cohesive reconfiguration. In A1, syntactic splitting frequently transferred relations previously encoded within extended constructions to adjacent independent sentences, increasing the role of repeated nominations and pronominal reference in participant tracking. This is consistent with [17], where it is shown that referential cohesion depends on the distribution of referring expressions across discourse, and with [13], where coherence risks in split-and-rephrase procedures are identified. A1 retained 56.4% of source segments and relied more strongly on repeated reference and linear adjacency, whereas A2 retained 73.1% and preserved a denser network of source-text relations. These proportions reflect different structural conditions of coherence preservation [28, 29].

The second major result concerns the distinction between causal coherence and event sequence. Causal explication occurred in 37 A2 and 33 A1 cases despite the smaller number of primary-operation cases in A2. Its function was not simply to make chronology explicit, but to encode a motive or cause–consequence relation whose inferability had been weakened by compression, dialogue reduction, or restructuring. This accords with work treating causal relations as a distinct class of discourse coherence relations [4] and with evidence that causal connections are central to narrative coherence [6]. Zwaan’s multidimensional account of narrative events [23] likewise supports separating chronological succession from causal interpretation. In the present corpus, causal explication functions as a compensatory coherence mechanism: descriptive or episodic material may be reduced while a functionally necessary relation is made more explicit. Event sequence can therefore remain intact even when causal coherence is weakened.

The third result is the non-equivalence of propositional and functional preservation. Folklore formulae, expressive nominations, figurative descriptions, and direct speech contribute to genre framing, characterization, conflict presentation, and evaluative orientation beyond factual content. This corresponds to [21] on lexical and phraseological organization in children’s folklore and [22] on the multiple functions of elements in Kazakhstani narrative. In the present corpus, removal of a folklore opening could preserve the event sequence while reducing genre framing; neutralization of an expressive action could preserve the event while weakening characterization; and reduction of a direct threat to an emotional-state statement could retain affect while diminishing its conflict-organizing function. Narrative transformation is therefore better described in terms of whether functions are retained, reduced, or recast than through a binary opposition between “meaning preserved” and “meaning lost”.

The treatment of formulaic material further supports this distinction. In [8], folklore formulae are identified as structurally and pragmatically relevant elements of fairy-tale discourse, while in [26], it is demonstrated that closing formulae form a differentiated class of conventional narrative units. A folklore formula is an identifiable linguistic unit contributing to genre organization; pragmatic closure is the overall communicative effect of the completed narrative. A closing formula may therefore be removed without eliminating pragmatic closure, while its preservation cannot guarantee pragmatic closure if the preceding conflict–action–outcome relation has been disrupted.

The lexical results similarly confirm the boundary between lexical substitution and neutralization of imagery. Research on lexical simplification emphasizes preservation of source meaning and contextual appropriateness [7, 18–20]. The present data add a narrative-functional criterion: replacement remains lexical substitution when denotative reference and relevant textual function remain



substantially stable; when figurative, evaluative, emotional, characterizing, or genre-marking meaning is weakened, the operation is classified as neutralization of imagery.

The most general result concerns the functional narrative core, treated as a control category rather than a seventh transformation operation. Across coherent transformed versions, the relation linking the principal participant, key event, necessary motive or causal relation, plot resolution, and pragmatic closure remained recoverable despite substantial formal variation. These components form a relational configuration: participant continuity supports reference; the key event anchors progression; motive or causal relation connects action and consequence; plot resolution provides the outcome of the central problem or conflict; and pragmatic closure concerns the overall communicative effect of the completed narrative. Plot resolution and pragmatic closure are therefore related but analytically distinct.

This relational account explains why retained verbal quantity was not proportional to narrative integrity. Extensive reduction could preserve macro-level coherence when the functional narrative core remained reconstructable, whereas a smaller local transformation could produce semantic distortion by removing a necessary causal relation, disrupting the event sequence, altering the plot resolution, or detaching pragmatic closure from the preceding event structure. Recent document-level studies similarly show that meaning preservation must be evaluated independently from surface simplification [28, 29].

The findings thus support the proposed two-level model: the six primary operations identify what changes formally, while the second level records consequences for cohesion, causal coherence, event sequence, and narrative function. This explains why formally similar operations may have different textual outcomes. The Kazakh material is significant because previous research has addressed A2-level lexical-grammatical selection [10], children's folklore [21], narrative semantics [22], and the under-resourced status of Kazakh in language-processing research [25], but not aligned segment-level transformations of the same narratives.

The study is limited to eight narratives with relatively stable event structures and to versions produced under a unified research protocol. Its conclusions concern textual organization rather than reader comprehension. Further research should test the model on independently transformed corpora, broader narrative genres, and naturally occurring level-modified texts, particularly narratives with multiple plot lines, temporal discontinuity, or more complex participant systems.

Conclusion

The study demonstrates that structural-semantic transformation of Kazakh folklore-didactic and children's narratives involves an asymmetrical reorganization of textual relations rather than a proportional reduction of linguistic material. Structural compression and syntactic simplification may reduce or redistribute cohesive resources, while repeated reference and inter-sentential linkage reconstruct referential continuity after syntactic restructuring. Causal explication performs a different function: it can compensate for contextual information weakened by compression by making a necessary motive or cause–consequence relation linguistically recoverable. The findings therefore confirm the analytical importance of distinguishing causal coherence from event sequence, since chronological continuity may remain intact even when the causal relation between events has been weakened.

The analysis also shows that preservation of propositional content does not necessarily entail preservation of narrative function. Genre framing, characterization, conflict presentation, imagery, and expressive force may be reduced or recast while the central event sequence remains recoverable. Macro-level coherence was maintained where the functional narrative core remained reconstructable as a relation linking the principal participant, key event, necessary motive or causal relation, plot resolution, and pragmatic closure. Within this configuration, plot resolution denotes the outcome of the central narrative problem or conflict, whereas pragmatic closure denotes the overall communicative effect of the completed narrative; the two are therefore functionally related but analytically distinct.



A1 and A2 revealed different profiles of textual reorganization rather than two proportional degrees of shortening. A1 relied more strongly on compression, syntactic redistribution, repeated reference, and linear event sequencing, whereas A2 retained a denser network of source-text relations and more frequently made causal relations explicit while preserving additional narrative functions.

The proposed two-level model consequently distinguishes six primary transformation operations from their effects on cohesion, causal coherence, event sequence, and narrative function, while treating preservation of the functional narrative core as a separate control category. This distinction provides a linguistically grounded framework for analyzing how transformed narratives preserve coherence and where structural-semantic modification begins to develop into semantic distortion.

REFERENCES

- [1] **Alva-Manchego F., Scarton C., Specia L.**, Data-Driven Sentence Simplification: Survey and Benchmark, *Computational Linguistics*, 46 (1) (2020) 135–187. DOI: 10.1162/coli_a_00370
- [2] **Cripwell L., Legrand J., Gardent C.**, Document-Level Planning for Text Simplification, *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023)*, (2023) 993–1006. DOI: 10.18653/v1/2023.eacl-main.70
- [3] **Agrawal S., Carpuat M.**, Do Text Simplification Systems Preserve Meaning? A Human Evaluation via Reading Comprehension, *Transactions of the Association for Computational Linguistics*, 12 (2024) 432–448. DOI: 10.1162/tacl_a_00653
- [4] **Scholman M.C.J., Demberg V., Sanders T.J.M.**, Descriptively Adequate and Cognitively Plausible? Validating Distinctions between Types of Coherence Relations, *Discours*, 30 (2022) 1–33. DOI: 10.4000/discours.12075
- [5] **Mulder G., Sanders T.J.M.**, Causal coherence relations and levels of discourse representation, *Discourse Processes*, 49 (6) (2012) 501–522. DOI: 10.1080/0163853X.2012.692655
- [6] **Cevasco J., Acosta Buralli K.**, Construction of coherence in the comprehension of narratives: Studies on the importance of the establishment of causal connections, gaps in current research and future directions, *Psychologist Papers*, 44 (1) (2023) 45–54. DOI: 10.23923/pap.psicol.3010
- [7] **North K., Ranasinghe T., Shardlow M., Zampieri M.**, Deep learning approaches to lexical simplification: A survey, *Journal of Intelligent Information Systems*, 63 (1) (2025) 111–134. DOI: 10.1007/s10844-024-00882-9
- [8] **Tyshchenko O., Fígedyová M.**, The image base of fairy-tale formula in the structure of folklore text (based on Russian, Slovak, and Czech fairy tales and their translations), *Slavistica Vilnensis*, 68 (2) (2023) 115–127. DOI: 10.15388/SlavViln.2023.68(2).9
- [9] Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*, Strasbourg: Council of Europe Publishing, 2020. 280 p.
- [10] **Iskakova Z., Demessinova L., Zinedina A., Yesbossyn L., Ospanova D., Zhumabay M., Yerbaty-rova Z.**, The significance of the lexical-grammatical minimum in teaching Kazakh at the A2 level, *Theory and Practice in Language Studies*, 15 (10) (2025) 3477–3488. DOI: 10.17507/tpls.1510.36
- [11] **Cardon R., Bibal A.**, On operations in automatic text simplification, *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, (2023) 116–130.
- [12] **Vásquez-Rodríguez L., Shardlow M., Przybyła P., Ananiadou S.**, Document-level text simplification with coherence evaluation, *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, (2023) 85–101.
- [13] **Grubišić A., Žitko B., Gašpar A., Vasić D., Dodaj A.**, Evaluation of split-and-rephrase output of the knowledge extraction tool in the intelligent tutoring system, *Expert Systems with Applications*, 187 (2022) 115900. DOI: 10.1016/j.eswa.2021.115900
- [14] **Schomacker T., Dönicke T., Tropmann-Frick M.**, Exploring automatic text simplification of German narrative documents, *Proceedings of the 19th Conference on Natural Language Processing (KONVENS 2023)*, (2023) 139–148. DOI: 10.48550/arXiv.2312.09907
- [15] **Farajidizaji A., Raina V., Gales M.**, Is it possible to modify text to a target readability level? An initial investigation using zero-shot large language models, *Proceedings of the 2024 Joint International*



Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), (2024) 9325–9339. DOI: 10.48550/arXiv.2309.12551

[16] **Burmakina N.G.**, Automated text simplification in the field of inclusive social addressing, *Terra Linguistica*, 15 (3) (2024) 7–14. DOI: 10.18721/JHSS.15301

[17] **Díaz-Negrillo A., Espinola Rosillo M.C.**, Mode of production and (referential) cohesion: An L2 English corpus-based study of syntactic coordination, *System*, 121 (2024) 103247. DOI: 10.1016/j.system.2024.103247

[18] **Truică C.-O., Stan A.-I., Apostol E.-S.**, SimpLex: a lexical text simplification architecture, *Neural Computing and Applications*, 35 (8) (2023) 6265–6280. DOI: 10.1007/s00521-022-07905-y

[19] **Graichen E., Jönsson A.**, Context-aware Swedish lexical simplification, *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, (2023) 11–20.

[20] **Ortiz-Zambrano J.A., Espín-Riofrío C.H., Montejo-Ráez A.**, Deep encodings vs. linguistic features in lexical complexity prediction, *Neural Computing and Applications*, 37 (2025) 1171–1187. DOI: 10.1007/s00521-024-10662-9

[21] **Tokpayeva L.S., Kuzembayeva G.A., Spulber D.P.**, A comparative lexical analysis of Kazakh, Russian and English children's folklore, *RUDN Journal of Language Studies, Semiotics and Semantics*, 15 (1) (2024) 107–125. DOI: 10.22363/2313-2299-2024-15-1-107-125

[22] **Tileubay G.K., Joldasbekova B.U., Tattimbetova K.O.**, Semiotics of objects in a modern Kazakhstani story // *RUDN Journal of Language Studies, Semiotics and Semantics*, 15 (4) (2024) 1390–1402. DOI: 10.22363/2313-2299-2024-15-4-1390-1402

[23] **Zwaan R.A.**, Comprehension: from clause to conspiracy narrative, *Discourse Processes*, 61 (4–5) (2024) 166–179. DOI: 10.1080/0163853X.2024.2327271

[24] **Espinosa-Zaragoza I., Abreu-Salas J., Lloret E., Moreda P., Palomar M.**, A review of research-based automatic text simplification tools, *Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing*, (2023) 321–330.

[25] **Veitsman Y., Hartmann M.**, Recent advancements and challenges of Turkic Central Asian language processing, *Proceedings of the First Workshop on Language Models for Low-Resource Languages*, (2025) 309–324.

[26] **Karagöz E.**, Closing formulae in tales: The case of Idel-Ural (Tatar) tales, *Motif Academy Journal of Folklore*, 14 (36) (2021) 1207–1226. DOI: 10.12981/mahder.1016619

[27] **Bakker J., Kamps J.**, Beyond sentence-level text simplification: Reproducibility study of context-aware document simplification, *Proceedings of the Workshop on DeTermIt! Evaluating Text Difficulty in a Multilingual Context @ LREC-COLING 2024*, (2024) 27–38.

[28] **Cripwell L., Legrand J., Gardent C.**, Evaluating document simplification: On the importance of separately assessing simplicity and meaning preservation, *Proceedings of the 3rd Workshop on Tools and Resources for People with READING Difficulties (READI) @ LREC-COLING 2024*, (2024) 1–14.

[29] **Maddela M., Alva-Manchego F.**, Adapting sentence-level automatic metrics for document-level simplification evaluation, *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1 (2025) 6444–6459. DOI: 10.18653/v1/2025.naacl-long.327

СПИСОК ИСТОЧНИКОВ

1. **Alva-Manchego F., Scarton C., Specia L.** Data-Driven Sentence Simplification: Survey and Benchmark, *Computational Linguistics*, 46 (1) (2020) 135–187. DOI: 10.1162/coli_a_00370

2. **Cripwell L., Legrand J., Gardent C.** Document-Level Planning for Text Simplification, *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023)*, (2023) 993–1006. DOI: 10.18653/v1/2023.eacl-main.70

3. **Agrawal S., Carpuat M.** Do Text Simplification Systems Preserve Meaning? A Human Evaluation via Reading Comprehension, *Transactions of the Association for Computational Linguistics*, 12 (2024) 432–448. DOI: 10.1162/tacl_a_00653

4. **Scholman M.C.J., Demberg V., Sanders T.J.M.** Descriptively Adequate and Cognitively Plausible? Validating Distinctions between Types of Coherence Relations, *Discours*, 30 (2022) 1–33. DOI: 10.4000/discours.12075



5. **Mulder G., Sanders T.J.M.** Causal coherence relations and levels of discourse representation, *Discourse Processes*, 49 (6) (2012) 501–522. DOI: 10.1080/0163853X.2012.692655
6. **Cevasco J., Acosta Buralli K.** Construction of coherence in the comprehension of narratives: Studies on the importance of the establishment of causal connections, gaps in current research and future directions, *Psychologist Papers*, 44 (1) (2023) 45–54. DOI: 10.23923/pap.psicol.3010
7. **North K., Ranasinghe T., Shardlow M., Zampieri M.** Deep learning approaches to lexical simplification: A survey, *Journal of Intelligent Information Systems*, 63 (1) (2025) 111–134. DOI: 10.1007/s10844-024-00882-9
8. **Tyshchenko O., Fignedyová M.** The image base of fairy-tale formula in the structure of folklore text (based on Russian, Slovak, and Czech fairy tales and their translations), *Slavistica Vilnensis*, 68 (2) (2023) 115–127. DOI: 10.15388/SlavViln.2023.68(2).9
9. Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*, Strasbourg: Council of Europe Publishing, 2020. 280 p.
10. **Iskakov Z., Demessinova L., Zinedina A., Yesbossyn L., Ospanova D., Zhumabay M., Yerbattyrova Z.** The significance of the lexical-grammatical minimum in teaching Kazakh at the A2 level, *Theory and Practice in Language Studies*, 15 (10) (2025) 3477–3488. DOI: 10.17507/tpsl.1510.36
11. **Cardon R., Bibal A.** On operations in automatic text simplification, *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, (2023) 116–130.
12. **Vásquez-Rodríguez L., Shardlow M., Przybyła P., Ananiadou S.** Document-level text simplification with coherence evaluation, *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, (2023) 85–101.
13. **Grubišić A., Žitko B., Gašpar A., Vasić D., Dodaj A.** Evaluation of split-and-rephrase output of the knowledge extraction tool in the intelligent tutoring system, *Expert Systems with Applications*, 187 (2022) 115900. DOI: 10.1016/j.eswa.2021.115900
14. **Schomacker T., Dönicke T., Tropmann-Frick M.** Exploring automatic text simplification of German narrative documents, *Proceedings of the 19th Conference on Natural Language Processing (KONVENS 2023)*, (2023) 139–148. DOI: 10.48550/arXiv.2312.09907
15. **Farajidizaji A., Raina V., Gales M.** Is it possible to modify text to a target readability level? An initial investigation using zero-shot large language models, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, (2024) 9325–9339. DOI: 10.48550/arXiv.2309.12551
16. **Бурмакина Н.Г.** Дигитальная симплификация текстов в сфере инклюзивной социальной адресации // *Terra Linguistica*. 2024. Т. 15. № 3. С. 7–14. DOI: 10.18721/JHSS.15301
17. **Díaz-Negrillo A., Espinola Rosillo M.C.** Mode of production and (referential) cohesion: An L2 English corpus-based study of syntactic coordination, *System*, 121 (2024) 103247. DOI: 10.1016/j.system.2024.103247
18. **Truică C.-O., Stan A.-I., Apostol E.-S.** SimpLex: a lexical text simplification architecture, *Neural Computing and Applications*, 35 (8) (2023) 6265–6280. DOI: 10.1007/s00521-022-07905-y
19. **Graichen E., Jönsson A.** Context-aware Swedish lexical simplification, *Proceedings of the Second Workshop on Text Simplification, Accessibility and Readability*, (2023) 11–20.
20. **Ortiz-Zambrano J.A., Espín-Riofrío C.H., Montejo-Ráez A.** Deep encodings vs. linguistic features in lexical complexity prediction, *Neural Computing and Applications*, 37 (2025) 1171–1187. DOI: 10.1007/s00521-024-10662-9
21. **Токпаева Л.С., Кузембаева Г.А., Спубер Д.П.** Сравнительный анализ лексических средств казахского, русского и английского детского фольклора // *Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика*. 2024. Т. 15. № 1. С. 107–125. DOI: 10.22363/2313-2299-2024-15-1-107-125
22. **Тилеубай Г.К., Джолдасбекова Б.У., Таттимбетова К.О.** Семиотика предметов в современном казахстанском рассказе // *Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика*. 2024. Т. 15. № 4. С. 1389–1402. DOI: 10.22363/2313-2299-2024-15-4-1389-1402
23. **Zwaan R.A.** Comprehension: from clause to conspiracy narrative, *Discourse Processes*, 61 (4–5) (2024) 166–179. DOI: 10.1080/0163853X.2024.2327271
24. **Espinosa-Zaragoza I., Abreu-Salas J., Lloret E., Moreda P., Palomar M.** A review of research-based automatic text simplification tools, *Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing*, (2023) 321–330.



25. **Veitsman Y., Hartmann M.** Recent advancements and challenges of Turkic Central Asian language processing, *Proceedings of the First Workshop on Language Models for Low-Resource Languages*, (2025) 309–324.

26. **Karagöz E.** Closing formulae in tales: The case of Idel-Ural (Tatar) tales, *Motif Academy Journal of Folklore*, 14 (36) (2021) 1207–1226. DOI: 10.12981/mahder.1016619

27. **Bakker J., Kamps J.** Beyond sentence-level text simplification: Reproducibility study of context-aware document simplification, *Proceedings of the Workshop on DeTermIt! Evaluating Text Difficulty in a Multilingual Context @ LREC-COLING 2024*, (2024) 27–38.

28. **Cripwell L., Legrand J., Gardent C.** Evaluating document simplification: On the importance of separately assessing simplicity and meaning preservation, *Proceedings of the 3rd Workshop on Tools and Resources for People with READING Difficulties (READI) @ LREC-COLING 2024*, (2024) 1–14.

29. **Maddela M., Alva-Manchego F.** Adapting sentence-level automatic metrics for document-level simplification evaluation, *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1 (2025) 6444–6459. DOI: 10.18653/v1/2025.naacl-long.327

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Bibikhan Myrzabek

Мырзабек Бибихан

E-mail: bibikhanmm@mail.ru

ORCID: <https://orcid.org/0000-0002-6846-1132>

Manshuk Mambetova

Мамбетова Маншук

E-mail: manshyk.mambetova@kaznu.edu.kz

Submitted: 03.06.2026; Approved: 26.08.2026; Accepted: 13.09.2026.

Поступила: 03.06.2026; Одобрена: 26.08.2026; Принята: 13.09.2026.