Some Aspects of Lexicalisation of Case Forms with Spatial Meaning in the Udmurt Language

Natalia Kondratieva

December 20, 2023
Abstract. The article is devoted to the study of the issues of lexicalisation of the so-called "somatic lexicon" in the paradigm of spatial case forms in the modern Udmurt language in a typological aspect. The relevance of the research lies in the fact that the study of the above questions of Udmurt linguistics opens new perspectives to some lexico-grammatical problems of Permic linguistics both from a diachronic and synchronistic aspect. As a result of the study it was revealed that the somatic lexicon is represented in different degrees in the case paradigm of the modern Udmurt language, the most common being the use of the illative case. Apparently, this also explains the fact that the process of lexicalisation is most often implemented in the specified grammatical form. This is also confirmed by fixed word combinations, where one element has now lost its independent distribution: "tode vajyny 'to remember', bure vajyny 'to remember', sane ponyny 'to take into account' and others. The phenomenon of lexicalisation, that is the transition of combinations of separate elements into fixed word combinations, in Udmurt is most often affected by linguistic units containing the verb forms "pyryny 'to enter, come in and other words with the given meaning', lyktyny 'to come, come in, etc.', voźyny 'hold', potyny 'go out, go out, etc.', koškyny 'go away, leave, etc.' with examples like "pele pyryny 'remember', jyre lyktyny 'come to mind', ŝulmaz voźyny 'hold in the heart', jyršy potyny 'forget', jyršy koškyny 'forget', etc. The full illative – inessive – elative paradigm can be observed in the word-formation of a limited number of somatic lexical units such as "jyr 'head', pel 'ear', or ŝulem 'heart'. A typologically similar tendency can be observed in other languages, for example in Russian. Evidence is found in written attestations from the turn of the 19th-20th centuries: "vojti v golovu 'to enter the head' – 'derzhat' v golove 'to keep in the head' – vyjti iz golovy 'to leave the head'; priyti v ruki 'to come into the hands' – 'derzhat' v rukah 'to keep in the hands' – uyt' i iz ruk 'to leave the hands'.

Key words: Udmurt language, lexicalisation, somatic lexicon, illative, inessive, elative.
Institute of System Programming of the Russian Academy of Sciences for organising and conducting the advanced training courses, within the framework of which the article was prepared.

1. Introduction

The case paradigm of the Udmurt language has been formed over many centuries and has preserved some unique features of the development of the grammar system. One such example is the case paradigm of a number of words belonging to the so-called somatic lexicon. Unlike other thematic groups of words, somatic lexicon is included in the process of phraseologization in case of functioning in the form of spatial cases. Why does a part of lexemes of the above thematic group feature a full system of nominalisation, while another part is represented by a truncated paradigm of spatial cases? Can all lexical units of the thematic group in question form fixed word combinations with figurative meanings? Why is it possible to say čemone jötyny ‘to hit a nerve’ (lit. to get in the goitre), but it is not grammatical to use čemone potyny by analogy of jyryś potyny ‘to forget (lit. to get out of one's head)’? These and other questions still remain undescribed in Permic linguistics. Meanwhile, they can be of some interest not only within the framework of philological research, but also to compile topical teaching aids for schools and universities.

Taking into account the above issue, the main purpose of this article is to analyse the functions of the somatic lexicon from the point of view of spatial case paradigms and to describe the means of their lexicalisation in the modern Udmurt language. The article also provides a brief insight into these constructions in Russian.

2. Literature Review

The functions of case forms have been investigated in detail in Permic linguistics, but only in recent decades has the Permic linguistic literature been expanded by notable research results relating to this area. The structural and semantic features of each member of the paradigmatic series have been specified and described (by Tarakanov (1997), Kondratyeva (2010; 2011), Nekrasova (2002; 2021), Usacheva (2012; 2012a), Kuznetsov (2012), and others). The interactional characteristics of nominal grammatical categories in the system of substantive word-formation has been revealed and cases of double case marking have been considered by (Kondratyeva (2011) and Nekrasova (2021)). Studies have been published in the field of grammatical synonymy, similarities and differences of morphologically determined case allomorphs and areas of distribution of similar phenomena in Permic languages (by Karakulov, Karakulova (2000), Kel’makov (2014), Nekrasova (2021), Normanskaya, Bezenova (2018), Ponaryadov (2018) and others.). Dialectal peculiarities in the use of case forms have been determined by (Karpova (2007; 2014), Maksimov (2018),
Arkhangelskiy, Usacheva (2018), and others. Despite this, some issues related to the lexicalisation of certain grammatical phenomena, including case forms, still remain unexplored.

The concept of "lexicalisation" as a linguistic phenomenon has been the object of special scientific research since the second half of the twentieth century. The term was first recorded in the works of representatives of the Prague Linguistic Circle (TPLK 1967: 27). In the Russian tradition, issues of lexicalisation are most often studied within the framework of lexicology or phraseology (Babkin 1970; Balazs 1972; Itkin 2002; Kubryakova 1995, etc.).

The classic definition of lexicalisation is that by V.V. Lopatin, who understands it as a "transformation of a language element (morpheme, word form) or a combination of elements (word combination) into a separate denominative word or into another equivalent vocabulary unit (for example, into a phraseological unit)" (Lopatin 1990: 208). As special cases of lexicalisation, he singles out: a) transformation of a bound morpheme, word-form, or phrase into a word; b) phraseologisation: the emergence of an idiomatic combination from a free one; c) semantic isolation of one form or part of forms; d) irregular formal phenomena, limited in their manifestation to a certain group (closed list) of words (Lopatin 1990: 208).

The issues of semantic transformations are actualised in the works by E.S. Kubryakova, who defines lexicalisation not only as "an increment of meaning, but also as a peculiar preservation, although in a hidden, implicitly not expressed form, of the information that was associated with the lexical combinability of the motivating word in the original motivating structure" (Kubryakova 1981: 1).

Relying on E. S. Kubryakova's research, E. V. Lukashevich proposes to consider lexicalisation as a cognitive process, which is the result of interaction and intersubjectivity of elements of language-system and language-capability (Lukashevich 2000: 94).

Despite the fact that the issues of lexicalisation at a lexical level have been investigated in relative detail both in theoretical and practical aspects, the study of the linguistic phenomenon in question at a grammatical level (Blinova 1984: 88–89; Pivovarova 2009; Agafonova 2011; Ganenkov 2005; Lyutikova 2017; Miziev 2013; Miziev, Ulakov 2013 and others) requires further scientific research, since the study of lexicalisation issues in relation to grammar will not only allow us to present a
synchronistic description of the language system, but will contribute to the solution of many
diachronic issues as well.

3. Materials and Methods
When writing this article, the main source of the study was the empirical material collected by
continuous sampling from the Lingvodoc 3.0 language database developed by the Institute of
Linguistics of the Russian Academy of Sciences, the V. P. Ivannikov Institute of System
Programming of the Russian Academy of Sciences (supervised by Y. V. Normanskaya, Ph. V. P.
Ivannikov Institute of Linguistics of the Russian Academy of Sciences) (http://lingvodoc.ispras.ru/),
as well as from the Udmurt Language Corpus (supervised by T. Arkhangelsky) (http://udmurt.web-
corpora.net/).

The Lingvodoc 3.0 programme is a system for collaborative multi-user documentation of
endangered languages, creation of multi-layer dictionaries and scientific work with the resulting
sound and text data. The Udmurt language is represented here by materials from early written
records, including manuscripts, dialectal materials, works of fiction, and media texts. The
development of the system architecture and kernel, computational modules, frontend was carried
out by the staff of the V. P. Ivannikov Institute of Programming (O.D. Borisenko, A.N. Tapekhin,
I.V. Bogomolov; I.B. Beloborodov, S.A. Ipatov, A.A. Zharov, etc.). To date, the Udmurt corpus
contains over 300,000 word occurrences. Filling the Udmurt language corpus, processing the parser,
and removing homonymy was carried out by a group of Udmurt scientists consisting of Ph. M. P.
Bezenova, Ph. A. F. Leontieva, Ph. O. V. Titova, Candidate of Philology. A. A. Shibanov. The
scientific supervisors were Doctor of Philology Y. V. Normanskaya, Doctor of Philology A. V.
Dybo. A. V. Dybo.

The Udmurt language corpus was developed by M. Medvedeva and T. Arkhangelsky as part of a
project at the School of Linguistics of the National Research University Higher School of
Economics. Three corpora are currently available: the corpus of modern literary Udmurt (the «main
corpus»), the corpus of Udmurt social networks, and the sound corpus of Udmurt dialects. The main
corpus has 9.57 million word occurrences, the social media corpus has 2.66 million word
occurrences and the sound corpus has 11,000 word occurrences.

The material for the study was collected using the Udmurt–Russian Dictionary (URS 2008) and the
Means of Figurative Expressions in the Udmurt Language (SOVUYA 1996), as well as materials
collected by Udmurt linguists G. N. Lesnikova (Lesnikova 2003) and A. V. Egorov (Egorov 2010). The National Corpus of the Russian Language (https://ruscorpora.ru) was used to identify examples from the Russian language.

To solve the scientific task in question, we selected all cases of word usage of the somatic lexicon marked with suffixes of spatial cases by continuous sampling. Furthermore, on the basis of component and contextual-situational analyses, we determined the direct and figurative meanings of each case. The descriptive method was used to present the results of the study.

4. Discussion

4.1 The system of spatial cases in literary Udmurt consists of seven members of the paradigmatic series: inessive (static location), illative (movement towards), elative (movement from), egressive (origin of movement), prolative (transitive movement), terminative (end of movement) and approximative (direction of movement). All of them can be used to express both spatial and temporal meanings (Kondratyeva 2011a; Nekrasova 2021).

In this paradigm, a special place is occupied by the use of somatic lexis, which can acquire a figurative meaning in verbal word combinations. As the examples extracted from linguistic corpora show, somatic lexis in literary Udmurt can be used in six out of seven case forms:

- **inessive**: Palʹpoton noš ymnyryn. ‘A joy on the face’;
- **illative**: Stas piez no jöspöriem dišaškemyń, pel'az kuin' poles ugy ošemyn. ‘His son, Stas, is dressed in a special way, three pairs of earrings are hung into his ears’;
- **elative**: Vot i umoj, vot i šon’er; ‒ jyryśtyz kartuzze pukon vyle yrgaltiz kużo. ‘That’s good, that’s right, – the host tossed the hat from his head onto the chair’;
- **egressive**: ’Hm, zootehn’ik,’ ‒ nylze jyryśenyz pydožaz mertaz Volod’a. ‘Hm, zootechnician, – Volodya looked at the girl from head to toe’;
- **terminative**: Pumiśkem murte jyryśenyz pydožaz učko, dišeža dunjalo. ‘The person they meet they look at them from head to toe’;
- **prolative**: Śekyteś siźyl zor šapykjos mazen, ymnyrtiz kuryt šinkylios pityrskizi. ‘Tears rolled down her face like heavy drops of autumn rain’.

It is important to emphasise that we did not find any cases of the case approximative in the studied corpora.
In order to determine the peculiarities of the representation of each case form within the somatic lexicon, we have selected cases of verb phrases acceptable both in the conditions of direct and figurative meanings, which were found in the corpus of the texts mentioned above. On the basis of analysing the examples found, we can conclude that in verb phrases including lexemes of the somatic lexicon, the most frequent combination is \textit{verb + illative form}. At the same time, different verbs can combine with the lexemes in question to varying degrees. In particular, the verb \textit{nypjany} ‘to put on (put on) the back’ can only attach the lexemes \textit{tybyr} ‘spina’ or \textit{pel′pum} ‘shoulders’: \textit{Van′ze soe tybyramy nypjasa vaj ylimy}. ‘All these things we carried with them thrown onto our backs’. Or the verb \textit{śipyrtyny} ‘to whisper’ can only be used with the word-form \textit{pel'e} ‘in the ear’: \textit{Pelyaz luškem śipyrtm}. ‘In his ear he whispered quietly’. In quantitative terms, we found 139 cases (60.43\%) of illative, 35 cases (15.22\%) of inessive, 30 cases (13.04\%) of elative, 3 cases (1.31\%) of egressive, 15 cases (6.52\%) of prolative, and 8 cases (3.48\%) of terminative.

Another group of verbal units, e.g. \textit{ponyny} ‘to put’ or \textit{jötyny} \textit{~ jötskyny} ‘to hit, touch’ can potentially be used with most lexical units: \textit{śulme jötyny} ‘to hit a nerve’ (URS 2008: 628), \textit{jure ponyny} ‘to remember’, \textit{pel'e ponyny} ‘to remember’, etc.: \textit{śinme jôte na soiz no: trosaz suredjosaz voźmatemyn šundy pukšon vakyt}. ‘It is also striking that many paintings depict the sunset’; \textit{Malpaško: kyk poles umoj so, myneštym veramjosme kot'ku pelaz ponoz}. ‘I think this is good in two ways: my words will be firmly remembered (lit. put in the ear)’.

A whole group of verb forms that actively participate in the process of lexicalisation is of particular interest. These include such verbs as: \textit{ponyny} ‘to put’, \textit{pyryny} ‘to enter’, \textit{voźyny} ‘to hold’, \textit{koškyny} ‘to leave’, \textit{potyny} ‘to go out’. It is in combination with these verbs that idiomatic expressions most often occur when undergoing the process of lexicalisation. Let us consider them in more detail in the next subsection.

4.2 When studying the peculiarities of the lexicalisation process, O. I. Blinova proposes to distinguish the following characteristics (Blinova 1984):

- Linguistic units
  1) are connected with motivated words;
  2) reflect the tendency towards the arbitrariness of the language sign;
  3) are related to isomorphism/non-isomorphism;
  4) are not accompanied by phonological reduction.
These features are also typical for cases of lexicalisation of verb constructions containing somatic lexicon in their structure.

4.2.1 The most common verb form involved in the creation of fixed word combinations with somatic vocabulary is *ponyny*. The lexicographical works mention the following semantic equivalents of *ponyny*: 1. to put; 2. to pour; 3. to drop; 4. to name; 5. to lay; 6. to put on: to shoe; 7. to attach (URS 2008: 538). The following meaning is also specified here: *pele (sane, čote) ponyny* ‘to listen to, pay attention to, take into account’ (URS 2008: 538).

As can be seen from the lexicographic description, no other word combinations in figurative meaning are recorded. Meanwhile, a whole range of somatic lexicon is used in connotation with this verb:

a) *jyre ponyny* ‘to remember (lit. to put in the head)’ (Lesnikova 2003: 39), (URS 2008: 266);

b) *pele ponyny* ‘to take into consideration, to listen to; to mottle on the ear (lit. to put in the ear)’ (URS 2008: 519), (Lesnikova 2003: 39);

c) *pele ug pon(y)* ‘and does not lead with his ear’ (URS 2008: 519), (Egorov 2010: 66);

d) *śulme ponyny* ‘to hitch a nerve (lit. to enter the heart)’ (Lesnikova 2003: 103);

e) *yme ponyny* ‘try’.

4.2.2 Similar meanings are also characteristic of fixed combinations with the verb form *pyrny*. The lexicographical works mention the following semantic equivalents of the verb *pyrny*: 1. to go in; 2. to enter; 3. to get in; 4. to interfere; to invade; 5. to steam (in a bathhouse) (URS 2008: 572).

As evidenced from the lexicographical description, no records are found for expressions with figurative meanings, where somatic lexicon would be used inflected with spatial cases. Meanwhile, these forms are widespread in fiction: *Veramed kylzišlen jyraz no, śulmaz no med pyroz*. ‘It is important that your words stay in the listener’s memory and their heart’.

On the basis of the linguistic material, we can distinguish the following constructions with the verbal unit in question:

a) *jyre pyrny* ‘to appear, to be remembered’ (lit. to get into the head);

b) *yire ug pyr* ‘1. impossible to remember (lit. not to go into the head)” (Dzyuina 1996: 40); 2. impossible to imagine ”;
c) pele pyryny: ‘1. to hear, to hear; 2. to be unpleasant to hear, to irritate’ (URS 2008: 519), (Lesnikova 2003: 81);

d) śinme pyre ‘1. to catch the eye; 2. to annoy’.

e) śulme pyryny ‘to come to one's heart (lit. to enter one's heart)’ (URS 2008: 628);

f) śulmaz ud pyry ‘you cannot look into the heart (lit. you cannot enter the heart)’ (Dzyuina 1996: 112);

g) sil'vire pyryny ‘to get better (lit. to enter the body)’ (Egorov 2010: 100);

h) lule-puške pyryny ‘to enter the soul’.

i) ym ug pyry ‘not to dare, not to open (lit. the mouth does not enter’).

4.2.3. The above group of somatic vocabulary can also be used with the verb voźyny ‘hold’. Lexicographic works mention the following semantic equivalents of the verb voźyny: 1. to hold; to retain; 2. to keep; 3. to observe, to establish, to found; 4. todyn voźyny ‘to keep in memory, to remember’ (URS 2008: 131).

Based on the linguistic material, we can conclude that the same group of somatic lexicon considered above can be actively used with this verb:

a) jyryn voźyny ‘to keep in the head (in the mind), to remember’ (URS 2008: 266);

b) śulmyn voźyny ‘to come to the heart, to catch’ (URS 2008: 628);

c) kiyn voźyny ‘to control, literally, to hold in one's hands’;

d) todaz voźyny ‘to remember’.

With the lexical units śin ‘eye’, pel’ ‘ear’, pyd ‘leg’, pin′ ‘tooth’ the verb voźyny is possible with postpositional forms: śin aźyn voźyny ‘to keep in view’ (Dzyuina 1996: 113); pel’ śöryn voźyny ‘to remember’, pyd ulyn voźyny ‘to oppress, to keep under the heel (lit. to keep underfoot’, kylze pin′ šәryn voźyny ‘to keep silent, lit. to keep one's mouth shut’.

4.2.4 The somatic lexicon can also be used in a figurative sense with the verb potyny ‘to go out’. Lexicographical works mention the following semantic equivalents of the verb potyny: 1. to go out; to leave, 2. to grow, to rise, to sprout; to graft; 3. go, go out 4. cross over; 5. hatch; 6. Appear; 7. want; wish; 8. acts as the second component in compound verbs (URS 2008: 542). As the examples show, the following cases of lexicalisation are possible:

a) jyryś potyny ‘to get out of one's head, to forget’ (URS 2008: 266);

b) śulmyś potyny ‘to let go (lit. to go out of the heart)’.
c) jyrpydesti potyny: pydpydesti pyrysa, jyrpydesti potoz ‘insistently (lit. having entered through the heels, will leave through the top)’ (Egorov 2010: 88);

In this context, the verb form koškyny ‘to go away’ or verb forms close to this meaning can also be used, e.g. tölzdźyny ‘to go away’: jyryś tölzdźyny ‘to go out of one's head, to forget’ (URS 2008: 266). The lexeme pel’ ‘ear’ can be used in constructions with postpositional forms or with the prolative case: pel’ pyrti potyny ‘to pass by the ears’ (URS 2008: 519) ~ pel’ti potyny ‘to forget (lit. to go out through the ears)’.

Thus, the group of words jyr ‘head’, śulem ‘heart’, pel’ ‘ear’, šin ‘eye’ can create whole paradigmatic series with case forms of illative – locative – elative meanings.

Based on the above examples, we can assume that cases of lexicalisation are relict phenomena of the development in a certain stage of life of a language. This is confirmed by examples from other languages. We can identify similar tendencies based on materials of the Russian language, in particular. The majority of the examples reflect the development of the language at the turn of the 19th–20th centuries:

a) Ne daj vojit̆i v serdce iskusheniyu mirom Božiim pod predlogom grekhovnogo ego sostoyaniya (Buldakov N. S., 1924). Do not let temptation by the world of God enter the heart under the pretext of its sinful state.

b) Pomogi mne, Gospodi, derzhat’ v serdce svoe yasnost’, tishinu, ponimanie (Prishvin M. M., 1914). ‘Help me, Lord, to keep clarity, silence, understanding in my heart’.

c) Priznayus’, moe obrashchenie ne bylo sledstviem kakogo-nibud’ velikogo sverh estestvennogo otkroveniya; ubezhdenie moe vyshlo iz serdca; ya plakal, i ya uveroal (Vinogradov A.K. 1932). ‘I confess that my conversion was not the result of some great supernatural revelation; my conviction came out of my heart; I cried, and I believed’.

A similar paradigmatic series is typical for the lexeme golovo ‘head’:

a) Kak eto voshlo v golovu staruše — ne znayu: u nee iz rodni edva li kto byl vzyat (Ogarev N. P. 1860–1862). ‘How it entered the old woman's head - I do not know: hardly anyone was taken from her kin’ (Ogarev N. P. 1860-1862).

b) Ya vsperinaya to muchitel’noe dlya nee vremya. chasto derzhu v golove slova Avvakuma (Prishvin M. M., 1927). ‘When I remember that painful time for her, I often keep in my mind the words of Avvakum’.
c) *U menya vse ne mogli vyjti iz golovy ispanske del* (Gogol’ N. V., 1832). ‘I could not get Spanish affairs out of my head’.

5. Conclusion

This article presents only a small part of the work that can be devoted to the study of the peculiarities of somatic lexicon found in the Udmurt language.

As a result of the study it was revealed that somatic lexicon is represented to a different extent in the case paradigm of the modern Udmurt language. The most common instance is the use of illative forms. This is confirmed by such fixed word combinations, as *tode vajyny* ‘to remember’, *bure vajyny* ‘to remember’, *sane ponyny* ‘to take into account’, etc. Perhaps, this tendency is connected with the fact that illative is the most frequent case among the spatial cases of the Udmurt language.

The phenomenon of lexicalisation, that is the transition of combinations of separate elements into fixed word combinations, in Udmurt is most often affected by linguistic units containing the verb forms *pyryny* ‘to enter, come in and other words with the given meaning’, *lyktyny* ‘to come, come in, etc.’, *voźyny* ‘hold’, *potyny* ‘go out, go out, etc.’, *koškyny* ‘go away, leave, etc.’ with examples like *pele pyryny* ‘remember’, *jyre lyktyny* ‘come to mind’, *śulmaz voźyny* ‘hold in the heart’, *jyryś potyny* ‘forget’, *jyryś koškyny* ‘forget’, etc.

As the results of the study have shown, the full paradigm illative – inessive – elative can be attested in the word-formation of a limited number of instances from the somatic lexicon such as *jyr* ‘head’, *pel’* ‘ear’, *śulem* ‘heart’. It is important to emphasize that the above word combinations function both in the form of the literary Udmurt language and in its dialectal forms. This allows us to conclude that the phenomenon under study reflects the state of the Udmurt language in the early period of its development.

The study of the somatic lexicon within the framework of lexical cases can open new perspectives for studying the main stages of development of the Udmurt lexico-grammatical system in a diachronic aspect. The study will also be continued in terms of typological parallels with the contacting Russian and Tatar languages.

This article presents only the most striking cases of lexicalisation of case forms presented in modern Udmurt. Undoubtedly, there is still a whole stratum of vocabulary that requires detailed study.
particular, the study of fixed combinations with the components ĺin ‘eye’ and pe/l ‘ear’, which are actively used with serial afterwords, requires special attention.

**Literature**


Balazs 1972 – Balazs J. Syntagmatisation and lexicalisation (to theoretical issues of lexicology and lexicography). In: Lexicology and lexicography. Moscow, 1972. (in Russian)


Kondratyeva 2010 – Kondratyeva N. V. Intercategorical relations in the grammar of the Udmurt language (on the material of the case of the direct object). Izhevsk, 2010. (in Russian)


Kubryakova 1981 – Kubryakova E. S. Types of linguistic meanings Semantics of the derived word. Moscow, 1981. (in Russian)


Usacheva 2012a – Usacheva M. N. Locative cases as part of groups with spatial meaning in Permian languages: dissertation. ... kand. philol. n.. Moscow, 2012a. (in Russian)


Doctor of Philological Sciences; Leading Researcher, (Izhevsk); Leading Researcher, V.P. Ivannikov Institute of System Programming, Russian Academy of Sciences (Moscow). V. P. Ivannikov Institute of System Programming of the Russian Academy of Sciences (Moscow); nataljakondratjeva@yandex.ru