THE CONCEPTION OF AN ELECTRONIC LEXICOGRAPHIC DICTIONARY
(BASED ON FOLK-SPEECH OF THE MIDDLE IRTYSH REGION)

Marina A. Kharlamova & Irina V. Malmberg
Omsk State University * / Saint Petersburg State University *
khr-spb@mail.ru / irina.malmberg@gmail.com

Abstract
The article addresses the possibility and the necessity of compiling a folk-speech dictionary for the multi-ethnic Middle Irtysh region. The dictionary represents the basic constants of contemporary villagers. The article describes the main theoretical guidelines of the new lexicographic project, presents the field structure of the active electronic dictionary, as well as the structure of a dictionary entry, taking into account the purpose to represent fully the functioning of culturally significant notions, which reflect the worldview of a contemporary dialect speaker.

Keywords
mental constants, lexicography, electronic dictionary conception, folk-speech

LA CONCEPCIÓN DE UN DICCIONARIO ELECTRÓNICO
(BASADO EN EL HABLA POPULAR DE LA REGIÓN MEDIA DE IRTYSH)

Resumen
El artículo aborda la posibilidad y la necesidad de compilar un diccionario de habla popular para la región media multiétnica de Irtysh. El diccionario representa las constantes básicas de los aldeanos contemporáneos. El artículo describe las principales pautas teóricas del nuevo proyecto lexicográfico, presenta la estructura de campo del diccionario electrónico activo, así como la estructura de las entradas del diccionario, teniendo en cuenta el propósito de representar plenamente las nociones culturalmente significativas, que reflejan la cosmovisión de un hablante de dialecto contemporáneo.

* Marina A. Kharlamova, Mira av., 55a, Omsk, Russia, 644077.
* Irina V. Malmberg, Universitetskaya emb., 7/9, Saint-Petersburg, Russia, 199034.

©Universitat de Barcelona
1. Introduction

“We do not live in the language as many believe, but the language lives in us. Inside of us it preserves something that could be called intellectually-spiritual genes that are passed from one generation to another” (Kolesov 1999: 137). These words by a prominent Russian scholar Vladimir Kolesov can be considered a blueprint for many linguists, lexicographers who are willing to leave “the folk memory” to the next generations. Studying a national worldview and the way it is reflected in different forms of folk speech is one of the main tasks of contemporary science. Compiling dictionaries of contemporary dialects, which reflect traditional folk culture, contributes to the reconstruction of a national worldview rooted in “the categories and forms of the native language” (Kolesov 1999: 148).

As part of the project Number 12-34-01078 “Semantic and cognitive research and the lexicographic interpretation of the language situation in the multi-ethnic Middle Irtysh region” (2012-2014) by The Russian Humanitarian Scientific Foundation we have compiled the Electronic Folk-Speech Dictionary of the Middle Irtysh Region. The dictionary is available on the website of the Omsk State University.

The object of description in the dictionary is the speech of the Middle Irtysh region villagers, without differentiation in terms of ethnic, social, educational or other characteristics. The culturally significant constants that objectify the worldview of the Omsk Irtysh region villager are the subject of description.

The immediate material for the dictionary is provided by the language facts that have been recorded by the tutors and students of the Omsk State University during the yearly dialectological fieldwork from the year 1976 up to now. We also involve material from the Dictionary of Old-Settlers’ Dialects (СГСрП 1-3; Addition 1, 2), because it includes language phenomena collected by the dialectologists from the Tomsk State
University (1952) and the Omsk State Pedagogical University during the last half of the 20th century, thus, covers the material that has been collected for almost half a century.

2. The dialect clusters of the Middle Irtysh region

The Omsk region is heterogeneous in its dialect structure (Sadredinova 1999: 74). Characteristic aspects of the language situation in the region are determined by the specific conditions of the Middle Irtysh region’s dialect development, such as multi-dialect population, isolation from ”parent” dialects, interaction with non-slavic languages, difference in time when dialects were appearing in the Omsk region.

Apart from Siberian so called old-settlers’ dialects (starozhilcheskiy dialekt – a dialect, which appeared earlier) and new-settlers’ dialects (novoselcheskiy dialekt – a dialect, which appeared later), there are also mixed dialects: Russian-Ukrarian, Russian-Belorussian, Russian-Czech, Russian-Polish, Russian-Lithuanian, Russian-German, Russian-Estonian etc. (Kharlamova 2007: 15-16; Kharlamova 2011b: 4-6).

It has been noted (Kharlamova 2006: 545-551; Kharlamova 2010a: 140-142; Kharlamova 2011a: 229-233), that non-old-settlers’ dialects, which function in the Middle Irtysh region, do not compose a unity in terms of phonetics, grammar, vocabulary and phraseology, therefore they can not be defined as a single dialect, which makes it difficult to present them in a differential dialect dictionary. As a result, we had to look for new approaches for presenting lexical material, which is being continuously collected in the multi-ethnic environment by students, post-graduate students and tutors of The Omsk State University.

3. Theoretical outline of the main principles underlying the electronic folk-speech dictionary

The necessity of compiling the dictionaries which would present the worldview of a contemporary dialect speaker has been addressed by many folk-speech researchers
(see G. Rakov, V. Vasiliev, E. Vasilieva, T. Demeshkina etc.): “...describing the sphere of concepts challenges the researchers with the task of making dialect dictionaries of a new type, which implies turning not to the word semantics, but to the main concepts of the folk-speech culture” (Demeshkina 2002: 59).

Some of the theoretical principles of the electronic issue of the dictionary have already been addressed before (Kharlamova 2010a and others). Thus, one of the basic general-purpose principles of the new folk-speech dictionary is the focus on reconstructing the dialect view of the world. That brings us to the next principle: founding on everyday mundane metatext consciousness, which captures and records the system of dialect speakers’ knowledge and ideas about the world. At the same time, the everyday mundane consciousness provides the factual base for the research.

Another theoretical guideline of the dictionary is commitment to the meaning, which should enable us to take into account all the shades of meanings objectified in the words, their variant forms, connotative lexical units, idiomatic expressions, general dialect utterances and patterns. Taking into account such an important theoretical principle as the focus on semantics allows us to objectify the results of describing culturally significant constants, which are traditional for folk mentality.

We have come to the conclusion that, since we are going to found on the original “live” material, neither chronological framework nor the structure of the dictionary will stay unchanged: dialect facts determine the making of the dictionary entry, examples, grammar and status labels, linguoculturological comments.

It is worth noting, that we do not always maintain the same structure of a dictionary entry, because we believe, that in each case both the semantics and the form of the entry are determined by the material. The language facts “lead” us and “impose” a certain making of the entry. Thus, following the language material is essential, because it enables to make a word description “dimensional” and “distinctive”. The implementation of this principle appears possible in the dictionary entries which present word families, and is associated with the commitment to reconstruct the dialect picture of the world.

The dictionary principles also stipulate the maximum broad interpretation of the semantic word-building, when a change in meaning is always considered a word-building
instrument and leads to homonymy. However, this concept is associated with a number of challenges described in (Kharlamova 2012: 108-113).

Including the dialect words, dialect variants of the countrywide words as well as the standard and colloquial equivalents encountered in the dialects is essential, since it is the only way to give the full picture of how a notion “lives” in the Middle Irtysh region.

After the preliminary research of the main mental categories existing in the folk-speech of the Middle Irtysh region we are going to present the complete view of the world, nature, society, the real world transformed in the word. We suppose, that the world has expanded for a folk culture-bearer, which shows in the inner dynamics of all elements of the dialect system.

Needless to say, we are facing a complicated task, which can be solved only if new criteria of choosing and delivering language facts are developed.

It is believed, that the electronic format can be the most adequate and up-to-date form of implementing the new lexicographic project. Since the dictionary presents the most significant constants of the folk world perception (House, Life, Bread, Field, Crop, etc.) verbalized in the villagers’ speech, the dictionary structure does not suggest, that key lexical items should be situated in the alphabetical order. Nevertheless, the electronic version allows us to put the language facts into the alphabetical order if necessary.

4. The structure of a dictionary entry

Thus, each dictionary entry is based on the family of words. As for placing the entries in the dictionary, here we proceed from the language (or, to be more precise, speech) material. We determine the significance of the constants in the folk mentality of the Omsk Irtysh region. The “word-family” principle allows to include into the same dictionary entry semantically and derivationally close representations of the basic notions, that have always been significant for the Middle Irtysh region speakers, i.e. both now and in the past. It also enables us to trace the “life” of a word in history and in present time.
4.1 "Zoning" principle

However, the absence of the same dictionary entry structure does not mean, that the material is not structured and “formalized” for the electronic version. Each dictionary entry is headed by the word, which represents a culturally significant notion / constant (often it is the name of a concept): *House*, *Yard*, *Family*, *Rib Grass*, etc. By means of semantic and cognitive analysis we have defined the *main* meanings for each such word. As a rule, there are maximum 4 meanings (if there are more meanings in the dialect, they are always presented in the dictionary). The electronic version stipulates zones: (A) – synonyms, (Б) – dialect variants and (B) – derivational forms.

The words listed in these zones and used metaphorically are labeled with (*) and placed into zone (Г) – metaphorical meaning – with the description of the new semantics. This zone includes semantic derivational forms. For example, in the entry for *Дом* (*House*) zone (Г) includes the diminutive form *домик* (*little house*) which means ‘beehive’ and *скворечник* (*birdhouse*) which means ‘a house, a city apartment in a many-storeyed house’.

In accordance with listed zones, the contexts are placed into the example zone (Е), the zone and the actualized semantics of the correspondent lexical unit are labeled. Idiomatic expressions, phraseological units and collocations (Ж) are located in zone (Д). They often objectify the image and symbol as constituents of a concept (Kolesov 1999: 81), which is important when reconstructing the components of the sphere of folk concepts. It seems expedient to place precedent utterances, which clarify the specific features of folk world perception and the dialect picture of the world, into a separate zone (З).

The specific of folk mentality of the Middle Irtysh region is reflected in zone (И), where one can find linguoculturological comments for the basic constants’ representatives. Thus, a dictionary entry about plants contains a great number of synonyms and dialect variants of the naming units. The difference between them in terms of use and circulation, located in zone (И), allows comprehending the complex processes of naming, determining the synonymic and word-building links of the lexical
units, which eventually provides for the adequate presentation of the units’ semantics, and enables us to identify the plant naming in the worldview of a contemporary villager.

Besides, the zone *in Latin* (L.) is very active for plants. It allows narrowing down the numerous folk plant names to one naming unit by means of identification. For instance, in the dialects in question the same naming unit can refer to different plants, coming from different botanic families: жабре́й ‘bee nettle – *Galeopsis tetrahit* (L.)’ – жабре́й ‘amber / touch-and-heal – *Hypericum perforatum* (L.).’ Therefore, there is a separate dictionary entry with the Latin name for each of these plants.

Grammatical and stylistic features appear next to the described units when they are relevant. They can be placed to zones (A), (B), (В), (Г).

4.2 *The content of the zones*

Now we turn to the content of some of the dictionary entry zones, which show the specific features of the new dictionary and of the Middle Irtysh region villagers’ worldview.

We consider the examples to be inherently valued, since they are essential for rendering and describing the whole scope of meanings embodied in the folk consciousness, and for proving the existence of a constant in the folk-speech.

It is of crucial importance to highlight, that metatext utterances of dialect speakers are relevant as a source of material for the dictionary. On the one hand, they reflect so called "naive" view of the world, on the other hand, they explain the meaning of a dialect word, which is necessary for a lexicographer:

1. Зарница каг-зарева такие/ асвещает асвещает как-молния// Зарница светим/ эта уже фсё/ хлеб паспел называли уборач’ная// *(Distant lightning is like a kind of glow/ it lights, it lights like-a lightning// Distant lightning glows/that's it/ the bread wheat is ripe it was called harvest time//* (Zelinnaya village, Cherlak region, I. Domracheva, female, 65 years old, literate, old settler, 2004).

The texts of dialect speech, which contain the definitions of the notions unknown for a standard Russian speaker, are particularly valuable. For instance, the words
надышбица ‘loft’ and сестренца ‘female cousin’ do not exist in the standard Russian language, they are defined by the respondents in the following contexts:

2. В-выбе между крышей и потолком надышбица/ на-дому чирдак/ (In an izba (Russian log hut) between the roof and the ceiling there is a loft/ on a house there’s a loft/ (area centre Nizhnjaya Omka, Nizhneomsk region, an, 77 years old, new settler, 2007).

In case of semantic or other similarity of the synonyms, the metatext examples provide all the semantic shades, leaving the “right of vote” with the “masters” of the folk-speech:

3. Хто сарай назовёт хто двор/ мы двором зовём/ Some call it saraj ‘shed’, some call it dvor ‘yard’/ we call it dvor ‘yard’/ (Kachesovo village, Muromtsevo region, V. Feoktistova, 72 years old, illiterate, old settler, 2005).

The main difference, if compared with a traditional explanatory dictionary, consists in the fact, that the authors abandon the idea of polysemy (as it was already done in the Dictionary of the Old Settlers’ Dialect. Addition 1-2) and implement the ideas of semantic word-building (V. Markov, I. Toroptsev, B. Osipov and others) (Markov (1981); Osipov (1984); Toroptsev (1979)): each new meaning is a new word. According to that stance, the word skvorechnik ‘birdhouse’, in the dialect meaning ‘city flat’ (e.g. 4. Я не-пойду в скворечник жить – ‘I will not go to live in a birdhouse’ (‘city flat’)) will be placed into the dictionary entry as a connotative synonym of the word дом ‘house’.

It is essential for a dialectologist to describe the archaic level of a dialect, because it forms the basis of the dialect system. The hectic pace of modern lifestyle, active interference processes in dialects, etc. make heavy influence on all dialect speakers. The changes that take place in contemporary dialects are reflected in the speech of all generations. We consider it necessary to use the speech of the representatives of all dialect levels in order to show the wealth, diversity and dynamics of the dialect speech.

4.3 Informant’s passport

Another significant part of a dictionary entry is the informant’s passport. It is placed after each example and includes the place of recording, age at the time of
recording, education, gender (male/female), dialect, year of recording. We tend to believe, that it is more expedient to indicate the dialect rather than the nationality of an informant, because there are not only dialects of different ethnic groups, but also dialects with different dialect basis, which are spread across the Middle Irtysh region. We presume, that the informant’s passport is important for further sociolinguistic, gender-based, culturological, etc. research of the region.

4.4 Variations of the naming units

While working at the experimental dictionary entries, we came to the conclusion, that it was expedient to introduce grammar and stylistic features of the words which represent certain notions. The relevance of grammar labels is explained by the fact, that dialects can contain lexical-phonetic and lexical-grammatical variations of the nationwide words; a dialect word can have lexical and syntax combinatorial characteristics, expressive meaning, that are different from the standard language. All that is certainly essential for comprehending specific dialect features of an ethnos.

Thus, in the old settlers’ dialects there are synonymic lexical units, which are illustrative of various types of peasant houses: одностопный дом, одностопная изба; одностенка, одностопка, одноколка, одностопочка, feminine gender, diminutive ‘a house, which consists of one room, Russian log house’; пятистенный дом (literally – ‘five-wall house’), пятистенник (literally – ‘five-waller’), четырёхжильный (literally – ‘four-dweller’) – in the old settlers’ dialects, пятистенок (literally – ‘five-waller’), пятистенный – in the new settlers’ dialects; here they are presented both as lexical-grammatical variants of the phrase одностопный дом ‘a house consisting of one room’, and as lexical units with connotative labels. Besides, the variations of the naming units in the dialects with different dialect base are also provided.

4.5 Phraseological units in the dictionary

Apart from being manifested in lexical units, certain constants can show by means of collocations or fixed phrases; such examples are located in the dictionary in zone (Ж).
But which dictionary entry/entries they are to be included in? Linguists have often highlighted the significance of the language phraseological corpora, which is “the mirror, where a linguoculturological community identifies its national self-comprehension” (Teliya 1986: 9). The presence of cultural traces in the meaning of the widely used phraseological locutions, as well as turning to the etymology of the phraseological units’ “literal” meaning, allow us to place them to the main dictionary entries of the new dictionary: 5. Муromtsevo была хорошая/ весёлая/ а сичас как е-яме// // ‘Muromtsevo [village] used to be good/ cheerful/ but now it’s like in a pit// (Kam-Kurskoye village, Muromtsevo region, A. Krivenko (Voronova), 82 years old, female, 3 years of primary school, old settler, 2005). We presume, that image-bearing comparisons can be located in zone (Д) in accordance with their semantics: comparison как е-яме ‘like in a pit’ has the meaning ‘bad’, so it is placed to the corresponding dictionary entry.

In the active dictionary the indication of typical combinatorial characteristics of a word appear to be the important part of a dictionary entry, e.g.:

克莱стовый дом ‘a peasant’s house consisting of four rooms’: 6. А крестовые крест-наперекрест/ там четыре комнаты// And as for cross house/ there four rooms in it// (Panovo village, Ust-Ishim region, old settler, 1979);

односторонний дом ‘a house consisting of one room, Russian log house’: 7. Изба без-горницы изделана// Izba viz-gornitsy izdelana// ‘A log house made with no ben’ (Orlovo village, Tara region, M.Polonnikova, female, 78 years old, illiteral, old settler, 1979);

кулацкий дом ‘a house, which used to belong to wealthier peasants’: 8. И вот рядом тоже кулацки и вот идя садик был у-нас зделан ище кулацкий дом один стаит// And nearby there are also wealthy peasants/ and where a kindergarten was made there was also a wealthier peasants’ house// (Kopejknno village, Tavrichesky region, Neklesova, female, 88 years old, new-settler, 2004).
4.6 Origins of the words

The semantic space of Omsk region dialects includes lexical units of various origins, which is determined by the specific of peopling the region and forming of the dialects as well as by intensive inter-dialect and inter-language contacts. Therefore we have decided to include labels Russian-Ukrainian, Russian-Belorus, Russain-Czech, Russian-Polish into the dictionary entries, i.e. to indicate the informant’s dialect in the informant’s passport.

Thus, while working at the dictionary entry of the word Хлеб ‘bread’ ‘food product’, we found nationwide lexical units: под ‘hearth’ ‘lower surface of the oven’ (Ozhegov 1984: 458), ‘a place in the oven for bread baking’ (dialects) and помело ‘broom’ marked as rare (Ozhegov 1984: 483). But in Siberia and in the northern Russian dialects the word помел ‘broom’ (Onega, Pechora, Karel Russian dialects etc.) (The Dictionary of Russian Folk Dialects 1992: 320-321) is spread everywhere (in our dictionary it is marked повсем. ‘widespread’). The word крыло, крылышко ‘usually a goose wing for sweeping litter, dust’ came from the northern dialects. According to the prominent lexicographer I. Dal, the word is widespread in Siberia and across Russia, but we recorded it in the speech of Ukranian-born woman, which is the result of inter-dialect contacts.

The dictionary user can find such information in the informant’s passport, but we place it to the field Linguoculturological comment. Presumably, the interpreting of various forms of speech reflects the specific of the language situation in Omsk Irtysh region, which we strive to represent in the dictionary.

In the field Linguoculturological comment we locate, for example, the analysis of etymologically obscure words. The etymology of the word чиринь ‘hearth, lower surface of the oven’ is not clear. The word was recorded in the speech of a Ukranian-born woman: Чиринь разметешь и хлеб посадишь ‘You sweep the hearth and put the bread’ (Pristavnoe village, Cherlak region, T. Skripnik, female, 67 years old, Russian-Ukrainian, 2004).

According to the “Etymological Dictionary of Russian Language” by A. Preobrazhensky, in Ukrainian language there is the word черинь ‘hearth’, the author also
refers to the Church Slavonic чернь ‘three-legged frying pan’ (Preobrazhensky 1939: 69). M. Fasmer gives Russian Church Slavonic form чрнъ and Old Russian черень ‘a pan for salt boiling off’, Serbian Church Slavonic чернь ‘a frying pan on an iron tripod stand’ and Ukrainian черень ‘fireplace, oven, hearth’. M. Fasmer also points out, that Lettish cēri ‘red-hot stones in the oven, drying house or bath-house’ belongs to the same word family and is traced back to *чертъ (Fasmer 1986: 340). Consequently, the meaning of чирень is traced to Ukrainian черень ‘fireplace, oven, hearth’. The pronunciation of the analysed word can be interpreted basing on the regular transformations of the Church Slavonic form чернь in Ukrainian language and subsequent common Russian phonetic changes: the second vowel у appeared after the so called “fall of reduced vowels” and as a result of the regular development from “new” һ; while the first у might be the result of the common Russian process – the reduction of unstressed vowels – retention of unstressed ”и”.

In the analysed dialects there is a widely spread lexical unit курник and its cognates куринка, куренёк, акурник, курень with the meaning ‘stuffed pie’. In the north of Omsk region the stuffing is usually сало ‘lard’, мясо ‘meat’, капуста ‘cabbage’, in the south – курица ‘chicken’, so it turns out, that the lexical motivation of the word exists in the new settlers’ and mixed dialects. While in the old settlers’ dialects the word курник preserves its ceremonial significance: 9. Ежели свадьба курник стапают ‘If there’s a wedding, a stuffed pie is baked’ (Bolsherechje village, Bolsherechje region, M. Kalinina, female, 73 years old, 7 years of school, old settler, 1994). The naming unit курень ‘stuffed pie’, recorded in the new settlers’ dialects, is of interest for a lexocologist and for a researcher of the dialect speakers’ linguistic consciousness. It is known, that курень means ‘house’ in the dialects spoken in Southern-European part of Russia. Probably, the new settler dialect speakers simply adjusted the familiar and similarly sounding word курень to naming the popular across this territory pie курник.

We assume, it is necessary to clarify the questions of certain words’ origins, the source of borrowing, because in the heterogeneous (in terms of dialect base) dialects of the Middle Irtysh region the processes of interference and inter-language contacts are very active. These processes create the specific “fusion” of folk-speech, which we strive to render in the new dictionary.
5. Conclusion

The Middle Irtysh region is a multi-ethnic and multi-dialect heterogeneous environment, which contains features of both Slavic and non-Slavic languages and dialects. The electronic dictionary of a new type, containing unique language material collected for more than fifty years, allows preserving the inner dynamic and multi-dimensional mental constants of the contemporary folk speech in a most comprehensible form. The main principles of the dictionary are determined by the commitment to reconstruct and present the worldview of the dialect speakers actualized and objectified in the word. This principle preconditions the active type and the field structure of the dictionary, with focus on semantic word-building and word-family principle, as well as ‘zoning’ of a dictionary entry, selection and classification of language material, examples, comments and labels included in the dictionary, research conducted, etc.

Obviously, the work at the electronic dictionary will not be over when the grant project ends, because, as it is known, compiling a dictionary is an effortful and time-consuming task. The lexicographic project requires further developing, solving the problems which appear during the analysis of specific speech, as well as other tasks both linguistic and related to the dictionary software.

References


Collection of articles devoted to the 70th anniversary of prof. V. V. Kolesov], St. Petersburg: Philology Department of St. Petersburg State University, 545-551.


Kolesov, V. (1999) *Zhizn proishodit ot slova... [Life stems from word...]*, Saint Petersburg: Zlatoust, 368 c.


PREOBRAZHENSKII, A. (1939) *Etimologicheskiy slovar russkogo yazika* [Etymological dictionary of Russian language], Leningrad: Academy of Science Press.


