Everita Andronova
Institute of Mathematics and Computer Science, University of Latvia

Renāte Siliņa-Piņķe
Latvian Language Institute of the University of Latvia

Anta Trumpa
Latvian Language Institute of the University of Latvia

Pēteris Vanags
Latvian Language Institute of the University of Latvia

Riga

The Electronic Historical Latvian Dictionary Based on the Corpus of Early Written Latvian Texts[1]

1. Introduction

The Corpus-based Electronic Historical Dictionary of Latvian (*Korpusā balstīta elektroniska latviešu valodas vēsturiskā vārdnīca*; henceforth – LVVV) is the title of a four-year-long research project initiated in 2013 and financed by the Latvian Council of Science (Latvijas Zinātnes padome). The aim of the project is to compile a representative pilot of a modern electronic dictionary of 16th–17th century Latvian. This dictionary will contain approximately 1200 entries and is based on data from the Corpus of Early Written Latvian Texts (Latviešu valodas seno tekstu korpuss,

available at http://www.korpuss.lv/senie/). During the project the corpus has also been supplemented with formerly unincluded texts, namely the earliest printed or handwritten Latvian texts.

The project was made possible through several years' work on the Corpus of Early Written Latvian Texts as well as work on defining a dictionary structure and producing several hundreds of sample entries during the period 2002–2009[2]. During that period: 1) the early text corpus was compiled, with more than 40 publications from 16th–18th century,[3] 2) the editorial guidelines for the Historical Dictionary of Latvian were written on the basis of studies of historical dictionaries of other languages and 3) approximately 500 sample entries were written (300 common nouns and 200 proper names[4]); about one third of them are published in Word format on the project website (http://www.tezaurs.lv/lvvv/#vardnicas skirkli).[5]

The objective of this article is to describe the elaboration process of LVVV from the development of the corpus to the moment when the dictionary entries were uploaded to the internet, and to highlight some problems related to the compilation of the dictionary.

2. Sources

The history of the Latvian written language dates back to the 16th century and is largely linked to the Reformation of the Church. Only a few individual records of words have been preserved from before the 16th century. The earliest texts from the 16th century are various versions of translations of the Lord's Prayer, as well as separate short records in the books of Riga trade associations. The first longer texts are translations

---

[2] The first preliminary work on the dictionary was done between 2004 and 2007 within the framework of project No. 04.1307 of the Latvian Council of Science. From 2006 to 2009, the work continued within the cooperation project 'Latvian Cultural Space in European Context' and its subproject 'Main Lexical Units of Latvian Language Subsystems – Exploration, Publishing and Study'. These projects involved experts from various institutions – the Artificial Intelligence Laboratory at the Institute of Mathematics and Computer Science of the University of Latvia; the Chair of Baltic Linguistics at the Faculty of Philology (later: Faculty of Humanities) of the University of Latvia, as well as the Latvian Language institute.

[3] The chronological scope of the sources in the corpus is larger than that of the dictionary currently in making. The most recent text in the corpus is from the year 1800.

[4] Unlike several other historical dictionaries the Historical Dictionary of Latvian, along with a Polish historical dictionary of the 17th–18th century (SXVII), includes not only common nouns but also proper names.

[5] The principles of the Corpus of Early Latvian Texts are described in more detail by E. Andronova (Milčonoka, 2003; Andronova, 2007). The earlier stages of the work on the Historical Dictionary of Latvian are also described in Andronova, Trumpa, and Vanags (2006); Andronova, Siliņa-Piņķe, Trumpa, and Vanags (2008); Andronova, Trumpa, and Vanags (2012); Siliņa-Piņķe and Trumpa (2014); Vanags (2014).

for the needs of the Lutheran church — catechism, pericope, and hymn translations. For a long period of time they existed as manuscripts, but starting from 1586/1587 they appeared in printed versions. The first Catholic catechism, which is linked to the Counter Reformation efforts in the Polish-Lithuanian Commonwealth, was also printed around the same time (1585).

The quantity and volume of Latvian texts gradually increased in the 17th century. The above-mentioned texts were reissued and improved together with the first Book of Sermons (1654), a complete Bible translation (1685–1694), and other texts of a religious content. The first dictionaries and grammar books were also compiled at this time.

There are two significant aspects of these early Latvian texts. The first is that most of the texts were translations from German, Latin and Polish, and there were very few original texts. One of the few exceptions is G. Mancelius's Book of Sermons. The second aspect is that the majority of the translators were not native speakers of Latvian, and some translators had a rather limited command of Latvian. Hence, it is assumed that the written Latvian language in the texts from the 16th–17th century is a far cry from the Latvian spoken at that time (further readings see Rūķe-Draviņa, 1977; Ross & Vanags, 2008; Vanags, 2009).

The Historical Dictionary of Latvian is based on the Corpus of Early Written Latvian Texts, i.e. on texts dating from the 16th and 17th centuries. Thus, before writing the dictionary entries it was important to expand the corpus in order for it to be sufficiently representative. Additions to the corpus indluded the main Latvian lexicographical sources of the 17th century, as well as religious and secular texts. Unlike previously, the corpus currently encompasses not only printed but also handwritten texts.

Currently, the Corpus of Early Written Latvian Texts contains:

1) ten 16th century sources – all the oldest known books in Latvian,[6] as well as the earliest Latvian versions of the Lord's Prayer (both printed and handwritten);

2) 46 sources from the 17th century, including three dictionaries (both manuscripts of Christopher Fürecker's dictionary (2nd half of the 17th century), the dictionary part of the work Lettus and the so-called 'Ten conversations' (1638) by Georg Mancelius); 3) 17 sources from the 18th century (including one lexicographic source – Liborius Depkin's dictionary). The sources added to the corpus during the last three years were significant contributions from a chronological, thematic and also a stylistic point of view. Each of these sources is also a valuable complement to the LVVV.

We will now describe the main groups of sources in more detail.

---

[6] *Catechismvs Catholicorum* (1585), the Lutheran handbook Enchiridion (1586), *Euangelia vnd Episteln* (1587) and *Vndeudsche Psalmen* (1587).

2.1. The Lord's Prayers of the 16th century

The earliest known Latvian texts are several versions of the Lord's Prayer in various 16th century sources. Within the framework of this project, the Corpus of Early Written Latvian Texts was supplemented with the following: 1) the Lord's Prayer by Ghisbertus (Gis1507_PN[7]), written by hand in a Catholic Agenda printed in 1507 in Leipzig; further-more, two more handwritten versions of the Lord's Prayer, dated to approximately 1520, namely: 2) the Lord's Prayer by Bruno (Br1520_PN) and 3) the Lord's Prayer by Grunau (Gr1520_PN), 4) the so-called Lord's Prayer of Hasentöter printed in 1550 (Has1550_PN), which is the oldest known printed Latvian text; 5) the Lord's Prayer by Lazius printed in 1557 (Laz1557_PN), and 6) the Lord's Prayer by Megiser, approx. 1593 (Meg1593_PN).

The earliest Latvian versions of the Lord's Prayer are denoted by the names of the individuals who either wrote them or owned the manuscripts. While most of these prayers exist in several copies, two of them – those by Ghisbertus and Bruno – are only extant in one single copy. The Lord's Prayer by Grunau, along with the chronicle of Prussia by the same author, exists in several slightly different copies. The version by Hasentöter was first included in the 1550 German edition of Sebastian Münster's 'Cosmographia', in the Latin edition of the same year, and later in other editions in other languages as well. S. Münster's books is the source from which the Latvian text of the Lord's Prayer was later copied to be included in the publications by Wolfgang Lazius, Hieronimus Megiser and others; therefore some authors (e.g. Breidaks, 1994) also speak of the Lord's Prayers by Megiser, etc.[8]

Due to the numerous versions, we had to choose which texts of the Lord's Prayer to include in the corpus. With the historical dictionary in mind, we decided that only the earliest version from each set of lexically and grammatically identical texts should be included. However, to meet the needs of researchers wishing to use the corpus to analyse spelling other details, the other versions should also be added in the future.

The first Latvian translations of the Lord's Prayer are of a varying quality. Nevertheless, these texts are very important because they are the first written attestations of the approximately 40 Latvian lexemes used in this prayer.

2.2. Dictionaries

In addition to C. Fürecker's dictionary manuscripts, which had been added at an earlier stage, the corpus is now supplemented with G. Mancelius's German-Latvian dictionary Lettus (1638), which is the oldest lexicographic source of Latvian. Lettus consists of three parts that together comprise the first textbook for learning Latvian: a dictionary (Manc1638_L), the part *Phraseologia Lettica* (Manc1638_PhL) comprising 51 sections of lexis on different

---

[7] The abbreviations used in the corpus are used in this paper as well.
[8] On the oldest Latvian translations of the Lord's Prayer see also: Augstkalns, 2009, pp. 71–72; Pokrotniece, 2012.

topics, and 10 conversations about everyday issues (travelling, agriculture, etc); these are considered the first Latvian original short stories (Manc1638_Run). *Lettus* is an extremely valuable source for the compilation of a historical Latvian dictionary because 1) the beginning of each LVVV entry contains data on 17th-century Latvian lexicographic sources, and 2) all three parts of Lettus contain a large amount of secular vocabulary. Moreover, for many words this was the first time when they were fixed in printed form, and one of the tasks of the LVVV is to give information about this as well.

During the final year of the project, we will prepare Georg Elger's Polish-Latin-Latvian dictionary *Polono-Latino-Lottauicum* (1683) for inclusion in the corpus. This is the second printed Latvian dictionary, after G. Mancelius's *Lettus*.

2.3. Larger religious texts

The most voluminous source of 17th-century written Latvian, without which the compilation of LVVV would not be possible, is Ernst Glück's Latvian translation of the Bible. Until now the early text corpus contained only the New Testament part of this translation (JT1685). Due to the large size of this source, many LVVV entries take their usage examples chiefly from this text. Also, since the Bible was translated into Latvian in the late 17th century, the quotes from this New Testament text are often the final examples in the LVVV entries.

The quality and significance of E. Glück's Bible translation in the development of written Latvian is undeniable. A comparison with earlier translations of some parts of the Old Testament, e.g., with G. Mancelius's translation of Ecclesiastes (Manc1637_Sal) published in 1637, shows noticeable lexical differences. The 1689 translation of the Old Testament is richer and more nuanced.

Currently,[9] the corpus contains the first three books of the Genesis, but intense work on the rest of E. Glück's translation of the Old Testament (1689–1694) is under way at the moment. The digital version of the text is compared with the original and text structural mark-up is provided for inclusion in the corpus, and we plan to add the text to the corpus during 2016.

The corpus is gradually supplemented with other religious texts – namely with the Lutheran handbook published in the region in 1685. Its first part, the hymnal (LGL1685_K1), was one of the first sources included in the Corpus of Early Written Latvian Texts. Later it was followed by the pericopes translated and edited by C. Fürecker (VLH1685), Ecclesiastes (VLH1685_Sal), the book of Sirach (VLH1685_Syr), and a catechism (VLH1685_Cat).

2.4. 17th century oaths and laws

The texts of oaths and laws are considered minor sources but this does not make them less important. They are particularly valuable because the Corpus of Early Written Latvian Texts contains only a limited number of secular texts. Among the sources

---

[9] As of early 2016.

formerly added to the corpus are court-martial laws (SKL1696_KB; SKL1696_RA) and a law prohibiting infanticide (SL1684). During this project, a number of other texts were added to the corpus: tradesman oaths, witness oaths, and several other laws, the original versions of which are kept in the State Archives of Latvia. They were deciphered and copied digitally. This was a rather difficult task, especially deciding how to render certain 17th century handwritten symbols in print.

At this moment, the Corpus of Early Written Latvian Texts contains eight hand-written documents: the statutes of linen-weavers from 1625 (LS1625); the informant's oath from the so-called plough revision (German: *Haken-Revision*) in the region of Vidzeme in 1638 (Zv1638_VAR); two witness oaths from 1681 (Zv1681_Liec_1; Zv1681_Liec_2); an oath of non-German timber sorters from 1681 (Zv1681_Kok); the tradesman oaths of the hemp manufacturers (Zv1689_Kan), the salt-carriers (Zv1689_Salsnes) and the warehouse workers (Zv1698_Lig).

The statutes of the Latvian linen-weavers' fraternity in Riga are particularly important because they contain both general vocabulary and specific professional and legal vocabulary. It is the earliest known Latvian text of this kind. Besides, its German original has also been preserved, which is very useful in the work on LVVV because it makes it possible to clarify problematic places in the Latvian text (the Ger-man text, however, is not included in the corpus). Most handwritten sources, includ-ing the statutes of the linen-weavers, are represented in the corpus by their scanned facsimiles as well, so that users can see their original appearance.

2.5. Dedication poems

In the 17th century a new genre of Latvian texts arose – the so-called dedication poetry. It was composed for all kinds of events – both in family life (weddings, baptisms, funerals), academic life (defenses of theses, acquisition of degrees), and literary life (publication of books). The dedication poems were composed in various languages and printed on separate sheets of papers or in sets of several pages. Quite often, the language of the poem was not the native language either of its author or the addressee. The Corpus of Early Written Latvian Texts contains some dedication poems that represent the secular poetry of the 17th century – e.g., a poem by Heinrich Fuhrmann from 1690 (Fuhr1690_LL), a poem by Michael Wittenburg from 1696 (Witt1696_MMID), and a poem from 1685 by an unknown author (ZP1685).

2.6. Possible future supplements to the corpus

The sources available in the Corpus of Early Written Latvian Texts of course do not cover all the data of 16th–17th-century Latvian. The corpus should eventually contain all the known texts, and will hopefully also be supplemented with new-found sources. The fact that formerly unknown material can sometimes come to light quite nearby is proved by the recently discovered Latvian text in the Catholic handbook *Rituale*

sacramentorum ac aliarum ecclesiae ceremoniarum for Lithuanian and Polish priests. Editions of this handbook containing also some Latvian texts date from 1675 and 1685, as well as from the 18th century. Although they contain only about 250 word usages, they represent a valuable addition to the 17th-century Catholic Latvian texts.

This case highlights one of the problematic issues: How should the compilators of the corpus handle repeated editions which are identical or almost identical to the first ones? The current principle – using only the first publication – is not applicable in all cases. Reprints that have not been revised do not cause a problem (e.g., the editions of G. Mancelius's Sermon book in the late 17th century) – they would not provide the corpus with new material, only add to the frequency of word usages. Thus it would be sufficient to upload only the earliest known edition of the above-mentioned Catholic texts. It is likewise clear that revised editions should be added to the corpus (as some already are) – e.g., the editions of M. Luther's Small Catechism, beginning with 1586, and the reprints edited by G. Mancelius (1631) and Heinrich Adolphi (1685).

Apart from printed texts, the corpus would also benefit from the inclusion of several handwritten sources which also would constitute an important supplement with regard to subject matter. Among them are the manuscripts of Andreas Gecelius' translations of the Psalms and Ecclesiastes, the anonymous late 17th-century manuscript of prayer texts (see Augstkalns, 2009, pp. 514–529) and Georg Elger's translations of the Gospels and Epistles from 1640.

A specific group of sources are Latvian inscriptions (separate words and phrases) in German and Latin texts, either printed or handwritten. These do not fully comply with the definition of 'early Latvian texts', and thus we are not currently considering uploading them to the corpus. Yet, in the future they might eventually become an interesting supplement. For instance, there are Latvian inscriptions in several late 17th and early 18th church metric books (such words as *Audſeckne* 'ward, foster-child', *Büſſeneek[s]* 'free peasant who has the right to carry weapons', etc.[10]), or in printed German texts (*maiſe ſemme* 'lit.: "bread land", black soil with clay', *Wavveering* 'wild rosemary'[11]).

3. The process of preparing source texts for inclusion in the corpus

Before uploading each source to the corpus, the texts are carefully prepared. They are scanned, OCR read with the program FineReader, and the digital versions are proofread, against the originals. During this process, problems sometimes arise with the identification of certain written symbols. For instance, in the text of the Old Tes-

---

[10] Examples from marriage records made by Nathaniel Pommer in the metric book of Mālpils parish church during 1697–1706 (LVVA, record group No 235, description No 3, file No 173).

[11] Examples from Salomon Gubertus' agriculture guidebook *Stratagema Oeconomicum oder Acker-student, denen jungen vngeübten Ackerleuten in Lieffland zum nöhtigen Vnterrichte …* (Gubertus, 1645).

tament it is sometimes unclear whether the diacritic above a letter is one big dot, two smaller dots, or a macron, or whether a letter is topped by a gravis or a half-erased caron. The texts contain several versions of diacritics which were employed without much consistency. A question arises: how precisely should all these nuances be reflected, and is it possible to do some unification? The current principle is to reflect them as precisely as possible, while also taking into account the peculiarities of each source. For instance, in E. Glück's Old Testament what appears as a macron over a vowel is most probably a diaeresis, i.e. an umlaut sign, where the two dots have merged. The deciphering of manuscripts highlights other types of problems, e.g.: 1) should we follow the inconsistent and often non-differentiated usage of small and capital letters? 2) how to interpret unclear parts of the text?

> *Beth Kad Wairack no SKaitieʃchen tohs Wahtzes gir,*
> *Buß Liedcz tems Wehlehtems buth.*
> (Linen-weaver statutes: LVVA, funds 4038, inventory 2, files 1124.)
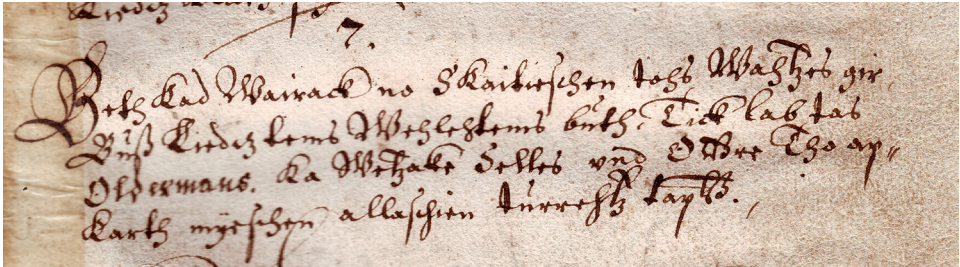

Figure 1. The statutes of linen-weavers (1625).

Due to the inconsistencies of the original, the text in the corpus retains capital letters only in word-initial position and only in proper names: *Beth kad wairack no ʃkaitieʃchen tohs wahtzes gir buß liedcz tems wehlehtems buth*[12]. (LS1625, 3r13.)

Before the texts are added to the corpus, structural marking is carried out. This, for example, entails marking words in other languages, so that they are not automatically included in the index of Latvian words. The majority of texts in the corpus are not monolingual. Different kinds of code-switching are very typical in the writings of early modern times. Until the late 17th century, the readership of Latvian books consisted of Baltic-German clergy, and Latvian text was basically inserted into German text, which functioned as a paratext. In such cases it is not difficult to distinguish the texts of different languages. Dictionary materials also do not cause problems. Problems arise, however, in cases of intrasentential code-switching. In such cases, it is not always

---

[12] If there are counted more Germans, then they should also be elected.

possible to tell whether a particular form is in Latvian or another language. A proper name in a Latvian text can be followed by a Latin inflectional ending, as in *ar weenu Mutt teitzeeta Deewu vnd to Thäwu muhſſa Kunga JEſu Chriſti* 'all together praise the God and the Father of our Lord Jesus Christ' (Manc1631_LVM, 2515). Here the gen. sg. *JEſu Chriſti* can be regarded as Latin forms. However, when the flectional forms of both languages look identical, as in *Vnd JEſus ſatziya vs to* 'and Jesus said to him' (Manc1631_LVM, 511), the word *JEſus* can be interpreted both as a Latin and a Latvian nom. sg. case form. In the corpus neither of these occasions is marked as non-Latvian text. Therefore they appear in the dictionary as well, including cases when the switched code is a single word, not a part of a sentence, e.g. *tas brähts / Abba / myļais Tähws* 'he cries / Abba/ dear Father' (Manc1631_LVM, 412). The word *Abba* 'my father' is left without marking in the corpus, and thus it also appears in the dictionary.

In the Latvian texts added to the corpus so far, one can observe words in German, Latin, Polish, Greek, Estonian, and Hebrew. The scale of structural annotation differs from source to source – in dictionaries it is more detailed than in other texts, e.g.

@v{Allein/} @l{adverb:} tickai to ween. (@v introduces the German text, @l – the Latin text).

Also in the Bible text a relatively detailed marking is used: each book and each chapter are marked, as well as comments on the text of the Holy Writ and notes, e.g.

@p{* Greek: Wallod: Apſmeets.}. Furthermore, indications to parallel passages in the end of a verse are marked, e.g. @t{w:3. 2.Moſ:Gr: 29,10.}.

This structural marking is also necessary for the determination of the exact location of a word form. The Corpus of Early Written Latvian Texts, unlike Modern Latvian text corpora, can give precise information about the source and the number of each page and line. This considerably alleviates the researcher's work, since it provides a simple way to check the correctness of the text by comparing it with the facsimile of the original.

Erroneous forms are also marked, i.e. obvious spelling mistakes or typos, e.g. *uud* instead of *und* 'and'. These errors are marked by first giving the reconstructed correct form and then the original erroneous form in curly brackets, e.g. *und{uud}*. On each such occasion a question arises: how far should the compilers of the corpus interfere by correcting the centuries-old text? The current principle is to mark only the obvious spelling mistakes. However, even with this principle there is a risk of acting incompetently or misleadingly.

Some publications, e.g. G. Mancelius's Lettus, contain a list of corrected errors that were noticed by author himself. Thus, G. Mancelius noticed that the form *attghadaht* 'to unravel' was misprinted as *attghaghat* and *Rahm* 'calm' as Rahn, and sometimes the author felt that a word should be replaced by another, more suitable one – e.g. in one case he wishes to replace *Smilltis* 'sand' by *Pieſchli* 'dust'. In the corpus, this is represented as *Pieſchli{Smilltis}*.
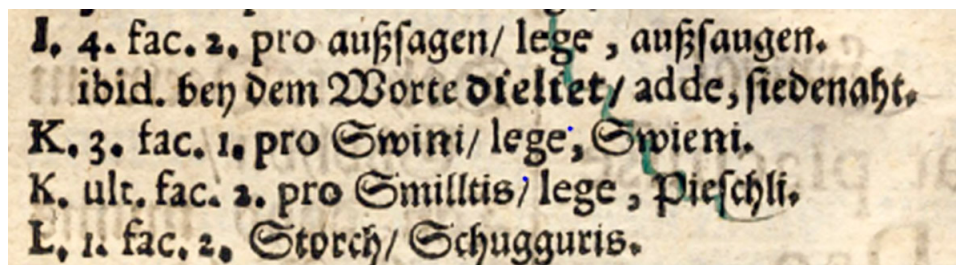
Figure 2. Sample from G. Mancelius's dicitonary Lettus (1638).

The process would definitely by alleviated by a morphological markup of the early text corpus, which has not yet been performed. The Institute of Mathematics and Computer Science of the University of Latvia has a certain amount of experience with working on old texts and has developed a law-based system to improve the optical character recognition and to correct typical mistakes caused by this, so that the word forms in the old texts could later be mapped to Modern Latvian word forms (Pretkalniņa, Paikens, Grūzītis, Rituma, & Spektors, 2012). Of course, some preliminary work is necessary before this system can be emplyed to analyse 16th–17th century Latvian texts and in the work on the historical dictionary. Most probably, a description of the "grammar" of each of these sources should be prepared, so that it can be later used for the mark-up.

4. Writing and supplementing the LVVV entries

The central part of the project is the compilation of sample entries for the Historical Dictionary of Latvian. Therefore, along with adding new sources to the corpus, already in 2013 we supplemented the 150 previously compiled entries[13] with new information from sources added to the corpus during 2009–2012.[14] In the end of 2013, we uploaded 150 revised and 63 new entries to the LVVV website.

Since the beginning of 2014, intense work has been carried out on producing new entries as well as on supplementing the existing entries with materials from the sources that were recently added to the corpus. Because the project plan entails the preparation of a relatively small number of entries (ca. 1200, ca. 700 of which are already written and some of which have already been published on the above-mentioned website), the entries of common nouns are chosen according to the principle of representativity,[15]

---

[13] They had been published on the LVVV website at http://www.tezaurs.lv/lvvv/

[14] The second manuscript of C. Fürecker's dictionary (Fuer1650_70_2ms), The Catechism of the Catholic Church of 1585, (CC1585), Georg Dressell's Catechism for children (Dres1682_SBM).

[15] Entries of proper names are being prepared in alphabetical order.

i.e. so that they comprise all parts of speech and are different in size. Thus, the entries include both little-represented lexemes and *hapax legomena*, and also large entries where the frequency of the word usage reaches into the hundreds or thousands.

Certainly, this dictionary will not contain almost every written Latvian word of the 16th–17th century. The number of potential headwords is much larger. However, in order to prepare a complete dictionary, a larger team and much more time would be necessary. The 1,200 entries will serve as an example, a testing field and a basis for future work, but even such a limited-size dictionary can give some insight into the grammatical and semantic phenomena and processes of the language, as well as the general use of written Latvian during the 16th and 17th centuries. As noted in the beginning of this article, during the first stage of work on the Historical Dictionary of Latvian we elaborated its instructions,[16] based on the experience of historical dictionaries of other languages,[17] but also considering the specific character of the Latvian language and early Latvian texts. Still, it is impossible in advance to envisage all possible difficulties. Resuming the work on the dictionary entries in the new project, we encountered new problems which had to be solved. This was partly due to the new format of the dictionary, given the transition to the dictionary writing software *TLex Suit 2013* (Joffe, de Schryver, & Prinsloo, 2003) and partly due to the fact that, as the dictionary expands, the entries become more and more diverse.

Henceforth in this paper, the examples of LVVV entries of common nouns and proper names will be followed by a brief description, as well as an analysis of the problems encountered.

Example of a LVVV entry of a common noun:[18]

(I) dibins (57) s. m. dibbina (2), dibbinâ (1), dibbinà (33), dibbinam (4), dibbins (11), dibbinu (6)

(II) *der Boden/ Grund/ Dibbins.* Manc1638_L, 38A$_{20}$,
grundveſte/ dibbins. Manc1638_L, 194A7.
Dibbins, der boden. grund, Fuer1650_70_1ms, 58$_{30}$,
Dibbins, der boden. Grund. Fuer1650_70_2ms, 86$_3$.

---

[16] We plan to publish the improved and supplemented instructions on the project website (http://www.tezaurs.lv/lvvv) by the end of 2016.

[17] For instance, a dictionary of 11th–14th century Old Russian (СДЯ XI–XIV), a dictionary of Old Polish (SSP) and a historical dictionary of Slovakian (SHS).

[18] The sections of each entry are numbered with Roman numerals. The denotations used in the Corpus of Early Latvian Texts have been retained, as well as combinations of two characters that represent the characters found in the text. We plan to eventually transfer the corpus to Unicode symbols.

(III)1.'pamats'.

(IV) ..*ka wings mums arridſan ißwehleyis gir czaur to paſſchu / pirms tai Paſſaulei tas dibbins licktz tappa / ka mums by ſwähteems buht vnd nhe peemettigheems preekſcha winju eekſchan tahß miläſtibas.* Manc1631_LVM, 201$_{18}$, *Deewa ſtippris d i b b i n s palleek ſtippre / und nhe ſchaubahß / und tam gir taß Sehghels: taß Kungs paſieſt tohß / kattri winjam peedärr.* Manc1654_LP2, 53$_1$.

2. 'dziļums, apakša'.

Wings nhe ſaudſehya tohs wätzus gharyus Ļaudis / kattri ar ſawu Stipprumu d i b - b i n à grima. Manc1631_Syr, 544$_{22}$, *Winņs ißtetzeya no Thäwu / Und ghreeſehs attkal py Thäwu / Brautza Elles d i b b i n à / Und attkal py Deewa Krähßlu.* LGL1685_V5, 2$_{17}$.

(V) – **(elles) dibenā (no)iegrūst** (3) dibbinâ eeghruhſt (1), elles dibbinà.. ee=ghruhſch (1), elles dibbinà no=ghruhduſſi (1)

– **elles dibins** (25) elles dibbina (1), elles dibbinâ (1), elles dibbinà (21), elles dibbins (1), ellesdibbinu (1)

– **(ie)grimt (elles) dibinā** (7) elles dibbinà.. ee=ghrimmuſchi (1), elles dibbinà ee=ghrimbtam (1), ghrimmſim.. dibbinà (1), ghrimma elles dibbinà (1), dibbinà grima (1), grimſt .. dibbinà (1), dibbinà ghrimbt (1)

(VI) ◈: liet. *dùgnas* 'dibens, pamats' no **dubnas*. Saknes patskanis *i* sekundārs no sākotnējā *u*.

(VII) ↘ bezdibins.

↔ dibens, dubens.

Example of a LVVV entry of a proper name:

(I) Matužs (2) *s. m. npers.* Mattuſcha (2)

(III) 'Matejs; JD'

(IV) *IN JEſus tahlaku eedams / redſeja weenu Zilweku pee Muita=Buhdas ſehſchamu / Wahrda M a t t u ſ c h a / in ſazzija us wiņņu: Eij man pakkaļ: In wiņſch zehlehs / in gahja tam pakkaļ.* VLH1685, 113$_{13}$.

(V) Δ **Matuža diena** (1) Mattuſcha deenâ (1)

'Svētā apustuļa un evaņģēlista Mateja diena (kristīgās baznīcas svētki), 21. septembris'

Ś: *M a t t u ſ c h a   D e e n â.   21.* Septembr. VLH1685, 112$_{24}$.

(VI) ◈ < vc. Mattheus

(VII) ↔ Matejs; Mateuss; Matīss

4.1. Head of entry (headword, grammatical information, frequency of usage) (I)

The choice of form for headwords is usually one of the most discussed issues in historical lexicography. For the time being, there is no universal principle regulating this choice. The authors of each historical dictionary take into account the lexicographical tradition of their respective country, as well as the specific nature of the data.

In LVVV the headword is usually a single word (in exceptional cases two or more words) transcribed in contemporary spelling. It is the basic form of the respective word (e.g., **dibins** 'ground' or **Matužs** 'Matthew'). The early texts provide rich material for studying the adaptation process of borrowed lexemes, their phonological form and visual representation. This is particularly true in the case of proper names, therefore it is sometimes difficult to define the form of the headword. In most cases when there are two or more headwords, it is due to the spelling peculiarities of proper names (**Atenas**, **Atēnas** 'Athena') or unclear morphological stems (**Manase**, **Manases**, **Manasis**, **Manasus**). In some cases, however, the entries for adjectives also have more than one headword versions (**niknis**, **nikns**, **nikna** 'furious') — here the regular basic form is the masculine nikns, but alse the form niknis with an inserted vowel -i- and feminine form nikna are included as headwords. More than one headword is also found in entries for compounds with various connecting vowels (**dusmapūķis**, **dusmupūķis** '(lit.: anger dragon, angry person)') and entries for nouns that represent a *nomen agentis* if they have both gender forms in the corpus (**adītājs**, **adītāja** 'knitter').

In some cases, the entry contains only the headword and a cross-reference to another entry. This is often done where the spelling of a word in early texts or in a particular text deviates from the contemporary spelling, e.g. **badadzeguse** → **badadzeguze**, **badadzeguse** 'hoopoe' or **Pamphilia** → **Pamvilija**. The compilers of LVVV had to find a solution for representing variant forms in the entry or head of entry. The Corpus of Early Written Latvian Texts contains many spelling variants that cover at least two periods in the history of written Latvian (before and after G. Mancelius's reform in 1631), as well as examples of the individual spelling of some authors (e.g., G. Elger and J. Reuter). Moreover, the corpus includes printed texts and deciphered manuscripts. In LVVV it was decided to include different spelling and phonetic variants in one entry,[19] but different morphological stems in separate entries. Unclear or reconstructed endings (usually the basic forms of proper names) are italicized, e.g. **Ampli*a***, **Ampli*as*** (1) *s. m. npers.* Ampliu (1) 'Ampliats', or **pried*s*** (4) *s. m.* preedus (4) 'additional payment'. Italics are also used if the quality of a sound is unclear, e.g. **mīkčaula** (2) *gen.* mihkzaula (2) 'soft-shelled'.

---

[19] Other historical dictionaries (e.g., СОРЯМР XVI–XVII) follow a similar practice; so does the pilot volume of the Dictionary of Latvian Dialects (LVIV), which is currently in the process of making.

In the head of the entry, the headword is followed by information about the total frequency of the lemma in the corpus, grammatical information, and all the attested word forms with correspoding token frequency in the corpus. The electronic version of LVVV will be mapped to the corpus (with concrete location pointers (addresses) of word forms), which will allow the user to see the particular contexts. Following common practice, homonyms are placed in separate entries. In some cases it is difficult or impossible to determine the part of speech of a word if it is the result of conversion and the context does not give clear indications. For instance, *nabags* can be either the adjective 'poor' or the noun 'poor person'. For the distinction of this kind of homonyms we also consulted the 17th-century lexicographic sources. For instance, nowadays the word *bezdibenis* 'abyss' is a noun, but *bezdibens* was fixed in G. Mancelius's dictionary *Lettus* as an adjective meaning 'endlessly deep'.

In the course of the project we discussed several complex issues concerning headwords, sometimes modifying or changing the original principles we had set out.

4.2. Word fixation in lexicographic sources of the 17th century (II)

If a word appears in the lexicographic sources of the 17th century, the examples from the sources are given immediately after the head of the LVVV entry. Currently the examples are taken from three sources, but when the corpus is supplemented with new-found sources it will not be difficult to add the new information to the dictionary.

Usually these lexicographic sources contain common nouns but sometimes there are proper names as well. For instance, chapters Nr 49 and Nr 50 of G. Mancelius's *Phraseologia Lettica* contain place-names such as Liebaw/ Leepai 'Liepāja' (Manc1638_PhL, 411[14]), while C. Fürecker's dictionary manuscripts show sporadic occurrences of place-names, personal names and other proper names, such as **Eewa**, *Eva*, *it. faulbaum*, *Ewa kohks. id.* 'Ieva (proper name, and also name of a tree)' (Fuer1650_70_2ms, 1132), **Muhsa**, *die bach bei Baliske u. Namanna, die machen beide die Aa.* 'the Mūsa river, which together with the Mēmele creates the Lielupe river' (Fuer1650_70_1ms, 161[19]) or **Auſeklis** *Spulgis. der Morgen stern. deeniņņa ne tahļu*, *auſeklis jau uhslez*, *der tag ist õ weit*, *der morgen stern gehet schon auff.* 'Morning star (Venus)' (Fuer1650_70_1ms 14[28]).

4.3. Word meanings (III)

Here we conform to the principles elaborated during the previous projects – word meanings are deduced from the examples found in the corpus. They are explained as simply as possible, with the help of synonyms or periphrasis if necessary. For instance, the entry **čuska**, **čūška** (Modern Latvian *čūska* 'snake') contains the archaic transferred meaning 'devil'. Explanations of obsolete concepts sometimes include encyclopaedic information as well – e.g., the entry **nauda** 'money' contains the col-

location ∆ *skēpju nauda* (lit.: 'spear money', cf. Modern Latvian *šķēps* 'spear') and its explanation: 'the so-called *Alberta dālderi* ('Albert's thalers', named after the Dutch regent Albrecht II), used in 17th-century Livonia, Courland and Riga; they had a depiction of two crossed spears on the reverse'.

There are certain criteria patterns for interpreting parts of speech, criteria for distinguishing word meanings, and established methods for determining the meanings in unclear cases[20]. Still, the interpretation of meanings never becomes a routine because problematic and non-typical cases appear from time to time.[21] Some cases will be presented below.

Since this is a historical dictionary, the historical principle is observed, i.e. the meanings are given in chronological order of usage. The first one usually is the meaning fixed in the 16th or 17th century. Quite often, it does not coincide with the modern primary meaning of the word. For example, the word *skriet* in LVVV is first defined as 'to flow': *Te yſſtep' pe to kruſte kòke: Tas dárgas aſins ſkräie bes galle..* 'They drag (Him) to the crucifix, The dear blood flowed without end' (Elg1621_GCG, 67$_2$); the second meaning is 'to fly': *Vnd ka tee Puttni ſkreen / ta ghroſahs ſöw tee Wehyi..* 'And when the birds fly / then the winds change.' (Manc1631_Syr, 598$_{24}$), and only its third, most recent meaning (in the corpus first attested in a text by G. Mancelius from 1654) is the Modern Latvian meaning 'to run': *Weens labbs Ghanns / kad wings räds to Willku ſkreijam ſtarrpan tahms Ahweems / tad eeſahk taß ſaukt..* 'A good shepherd / when he sees the wolf running among the sheep / then he starts to call to them.' (Manc1654_LP1, 451$_{11}$). The search for meanings and their explanation is a very meticulous work demanding a lot of attention because even a seemingly clear usage can contain a meaning quite different from the contemporary one. For instance, the word *čakls* might seem to always have had the same meaning as nowadays, i.e. 'hard-working, industrious'. However, in G. Mancelius's dictionary the Latvian word *čakls* is used as an equivalent to German *behend*, *geschwind*, and *schnellfüssig*, which all mean 'quick, agile, nimble'. Besides, the comparison of the 16th–17th century Bible text translation with the newest Latvian translation of the Bible can also be helpful in interpreting the word meanings. For instance, the 1685 text uses the word *čakls* as follows: *Wiņņo Kahjas irr **tſchaklas** Aſſini isleet.* 'Their feet are swift to shed blood' (JT1685, Rm 3:15), while the 2012 text uses the word *žigls* 'quick' in the same verse: *viņu kājas ir **žiglas**, kad tie steidz izliet asinis* 'Their feet are swift to shed

---

[20]   In such cases we consult the following references: 18th–19th century dictionaries; the Mühlenbach-Endzelin Latvian dictionary (ME); Latvian dialect dictionaries (e.g., Ādamsons & Kagaine, 2000; Kagaine & Raģe, 1977–1983; Reķēna, 1998, etc.); literature about German loanwords in Latvian (e.g., Sehwers, 1918, 1953; Jordan, 1995); various German dictionaries; other kinds of specific literature, etc.

[21]   On such cases see also Siliņa-Piņķe and Trumpa (2014).

blood' (*Bībele*, 2012, p. 2437). Thus, the meaning Nr. 1 (i.e. the oldest meaning) in the LVVV entry **čakl(i)s**, **čakla** is: 'quick, agile, nimble'.

Nevertheless, the old dictionaries can not only shed light but sometimes also confuse the compilers of LVVV. Thus, the only two instances of the word *čākstēt* (with the German equivalent *knirschen*, *knarschen*, but without examples) are from the two manuscripts of C. Fürecker's dictionary: ***Zahksteht***, *knirschen*, *knarschen* (Fuer1650_70_1ms, 311$_6$), ***Zahksteht***, *Knirschen*, *Knarschen* (Fuer1650_70_2ms, 538$_5$). The German word *knirschen* nowadays is translated as 'to crunch, to squeak', and we have given this meaning in the LVVV entry **čākstēt**. However, due to the lack of examples we cannot be completely certain about the meaning of this word in those times.

Although in most cases meanings are deduced from the examples found in the corpus, in unclear cases we turn to other sources as well. For instance, the collocation *plikka cepure* 'lit.: naked hat' in C. Fürecker's manuscript is translated as *ein Hut* 'hat' Thus, if we were to rely only on C. Fürecker's translation, we would have to explain *plika cepure* simply as '*cepure*' 'hat', which would not be very helpful. Luckily, ME contains a compound *plikcepure* defined as 'hat made of hairless leather'. We give this definition in LVVV as well.[22] Since the meanings of collocations are sometimes interpreted on the basis of only a limited number of examples or unclear definitions in dictionaries, one can agree with O. Reichmann that in historical dictionaries the explanations of words sometimes are only "hypotheses about their meaning" (Reichmann, 2012, p. 251).

4.4. Examples (IV)

Since the online version of LVVV will be linked with the Corpus of Early Written Latvian Texts, making it possible to click on the word forms to see their usage in a broader context (this mapping is not yet provided in dictionary writing software, but is planned in the future), the dictionary text contains only two illustrative examples of each meaning – the oldest one and the most recent one (in the period from the early 16th century to the late 17th century). However, this approach has some drawbacks. Sometimes – especially in entries with a small number of usages – both examples are almost identical or very similar because they are from the translations of the same Bible verse or the same hymn or carol. In the period between these two there may have been examples from other sources as well, and sometimes they are more precise and illustrative, but due to the principle that we are following, we cannot include them in the dictionary. This problem will be partly solved by the already-mentioned mapping between the dictionary and the corpus – but only partly because the word forms are only given in the head of the entry, not in the sections presenting the separate meanings.

---

[22] For more on collocations and their meanings see below.

4.5. Collocations (V)

After the examples of each meaning we provide word collocations, if necessary. They are a significant element of LVVV as they show the words in actual use. The dictionary will distinguish three types of word collocations: 1) frequently used free collocations, which are included not because of their semantic opaqueness but due to their frequent use (in LVVV they are preceded by the sign –), 2) fixed collocations without transfer of meaning, i.e. collocations where the meaning of one of the components cannot be deduced from the semantics of the headword of the entry, but which still do not have a specific transfer of meaning (preceded by the sign Δ), 3) fixed collocations with transfer of meaning, i.e. idioms or phraseological units (preceded by the sign ◊). Both fixed collocations and idioms are provided with definitions and the oldest and most recent example of usage. Nevertheless, the identifaction and reflection of word collocations in LVVV highlight a number of problematic issues.

4.5.1. Collocations and compounds

The peculiarities of Latvian spelling in the 16th and 17th century often pose a question: how can one tell collocations from compounds? Compounds in early Latvian texts were written either as one word, e.g. *naudakalējs* (1) s. m. *naudakalleis* (1) '(lit.: money smith) moneyer', or with the double hyphen typical of Gothic script, e.g. *naudakaša* (1) s. com. nauda=kaſſcha (1) '(lit.: money-scraper) miser; greedy person'. Similarly, collocations can also be written in two versions: either separately or with a double hyphen, as in the following two examples of the collocation Δ *zobu nauda*: sohbu nauda (1), sohbu=nauda (1) '(lit.: teeth money) a gift money given by godparents at christening'.

Currently we follow several criteria for distinguishing collocations and compounds: 1) if the first component agrees in gender, number and case with the second component, it is most probably a collocation, not a compound, as in the following example: *Jo mums arri weens **Leelas=Deenas** Jehrs/ tas irr Kriſtus par mums uppurehts.* 'Because we also have a Paschal Lamb / that is Christ, sacrified for us' (VLH1685, 37₂₆); if the first component has no ending or the ending is reduced or changed, it can be regarded as a compound: *Mihļajs Eņģļis/ krahſchnajs Eņģļis/ All. All. Kur es atraſchu ſawu Kungu/ All. All. Tas irr augſcham zehlees no ta Kappa/ All. All. ſcho=deen ſchinnî ſwehtâ **Leeladeenâ**/ All. All.* 'Dear Angel/ magnificent Angel/ Hallelujah. Hallelujah. Where will I find my Lord/ Hallelujah. Hallelujah. He has raised from the tomb/ Hallelujah. Hallelujah. today in this holy Easter/ Hallelujah. Hallelujah.' (LGL1685_K1, 498); 2) if both components of a word combination (e.g., *allus=klappi* 'beer tankard') can be observed separately as well, the collocation is quoted in both respective entries, in this case ***alus*** 'beer' and ***klape*** 'tankard'; 3) if one of the components does not exist separately, as with *abla=sahle* 'clover' or *Baijero=semmes* 'Bavaria', we do not reconstruct forms such as **abls* or **baijeri*, but create separate entries for *abla zāle* and *Baijeru zeme*. But making distinctions based on intuition

would be even more arbitrary. Sometimes even in cases when two words are written without a space or double hyphen, it is doubtful whether it is a compound, e.g. *Waʃʃaraszäppure* 'summer hat'. Meanwhile, compounds are sometimes written with a space between their components, e.g. *Mihkst zauļis* most probably belongs in the entry *mīkstčaulis* 'soft-shelled egg'.

4.5.2. Phraseological units or idioms

An idiom is traditionally defined as "*a phrase etc. which is understood by speakers of a particular language despite its meaning not being predictable from that of the separate words*" (SOED, 1993, p. 1312). However, fixed collocations and idioms are not always attested frequently enough to make it clear whether they were typical of their time. This is due to the dominance of religious writings among the early Latvian texts. Thus, the idioms of this period are usually from the Bible or from spiritual poetry, and their frequency of usage may be due to the repeated editions, e.g.: ◊ *dusēt Dieva rokā* (2) 'to be under God's protection', ◊ *neturēt mēli iekškan iemautu* (2) 'not to keep silent', ◊ *akmins pol ūz dūšu* (2) 'unpleasant feelings begin', ◊ *Abraama klēpis* (3) 'the bosom of Abraham; paradise'. The language of G. Mancelius's texts and C. Fürecker's dictionary manuscripts is more diverse, with phraseological units that may stem from the actual Latvian language tradition of the time, e.g.: ◊ *mieles tapa*, *mieļa tapa* (5) 'drunkard', ◊ *putna nags* (2) 'miser', ◊ *naudiņa spiež maciņā* (1) 'said if somebody is wasteful', ◊ *Antiņš, kur stabulīte / stabuliņ?* (3) 'fool'; some of them are still used today. Nevertheless, even these fixed collocations with transferred meaning have a low frequency of usage in the Corpus of Early Written Latvian Texts. As pointed out by O. Reichmann in his *Historische Lexikographie*: the less examples of a possible idiom there are, the more difficult it is to define its meaning (Reichmann, 2012, p. 405). A question thus arises: can we regard as idioms collocations which obviously have a transferred meaning (e.g., *naudas žurka* 'miser') but very few or only one instance of usage? Until now, taking into account that the Corpus of Early Written Latvian Texts (and the number of Latvian texts in the 16th–17th century in general) is relatively small, such collocations were treated as idioms and are included in the dictionary. Even if they have a single instance of usage, we try to explain its meaning, judging from the context. The same refers to the fixed collocations.

4.5.3. Interpretation of meanings in fixed collocations and idioms

The problems here are basically the same as with the interpretation of meanings of headwords. Whenever possible, they are deduced from the context. The task is easier if the respective collocation is given with a translation in one of the lexicographical sources (although the translations in 17th-century dictionaries cannot always be deciphered), or if the Latvian translation can be compared with the original text. In some cases, however, more detailed research is necessary. A good example

is the already-mentioned fixed collocation skēpju nauda 'spear money' from the court martial laws of 1696: *Kad nelaulahts Wihrs ar ohtra Wihra laulatas Śeewas/ jeb nelaulata Śeewa ar ohtras Śeewas laulata Wihra Laulibu pahrkahp/ tad buhs tam kas nelaulahts/ jo wiņſch pirmâ reiſi tohp atraſt/ 40. Dalderus Śuddraba=Naudu/ tas irr/ 20. Dalderus **Śkehpju=Naudu** ſtrahpi doht..* 'When an unmarried man with another man's spouse / or an unmarried wife with other wife's husband breaks the marriage / then the one who is not married / because it is first time (s)he has been found / must pay a penalty of 40 silver Thalers / that is/ 20 Thalers in spear money' SKL1696_KB, 91. At first, this concept, which was probably obvious to the audience of the 17th century, seemed obscure. Then it was clarified by the research of Konstance Kļava on the Latvian legal text vocabulary of the late 17th century, where she has explained the meaning of this collocation as follows: "the so-called *Alberta dālderi* 'Albert's thalers' (named after the Dutch regent Albrecht II), used in the 17th-century Livonia, Courland and Riga and evaluated higher than Swedish money; they had a depiction of two crossed spears on the reverse" (Kļava, 1989, p. 103). It turned out that this type of money is also mentioned by K. Mühlenbach as *šķēpu nauda* (ME IV, 1932, p. 33), and also in E. Dunsdorfs's book "*Latvijas vēsture 1600–1710*" (Dunsdorfs, 1962, pp. 303–305). Based on this information, we elaborated the definition in LVVV.

4.6. Explanation of origin (VI)

An important task of a historical dictionary is to provide new data and information relevant for the study of the development of the vocabulary. Certainly, this does not mean that a historical dictionary should replace the already existing etymological dictionaries – it should rather supplement them. The corpus makes it possible to pinpoint more precisely the time when a respective word entered the written language, which in many cases may be the time when the word actually appeared in Latvian.

Most Latvian texts from the 16th–17th century are religious texts. When it comes to the Christian vocabulary in Latvian, we have to keep in mind that it has been under a substantial degree of impact from other languages. Also, since Christianity was itself borrowed from another culture, we can say that the whole Christian discourse is borrowed. Nevertheless, it has been developed with the help of elements both from Latvian and from other languages. The Christian vocabulary of the Latvian language contains all possible kinds of borrowings: lexical, derivative, and semantic (according to W. Betz's terminology: *Lehnwörter*; *Lehnbildungen*; *Lehnbedeutungen* (Betz, 1959, p. 128)). There are also borrowings on the phraselogical level (German *Lehnwendungen*; see also Range, 1994). A historical dictionary might be of help to the research of this aspect of lexical history.

The explanations and references to the origins of words are based on various principles. Firstly, taking into account the digital format, it was decided to indicate

the origin after each headword, not just include a cross-reference. Therefore some lexemes in the dictionary may have identical references of origin. Secondly, these references are as short as possible, and various symbols are employed (◈ introduces explanation of origin; ↗ means "derived from"; : means "cf. "; < means "borrowed from"). The references include a limited amount of text – only in those cases where more detailed explanations are necessary. The indications of origin are of secondary importance in LVVV, since it is not an etymological dictionary. Therefore we basically use the alreadybexisting definitions. The sources most often employed are the following: ME; EH; Karulis 1992; Fraenkel, 1962–1965; Pokorny, 1959–1969; Sehwers, 1953; Summent, 1950. We also use the most recent etymological dictionaries of Lithuanian, such as Smoczyński, 2007 and ALEW (2015), and the Estonian etymological dictionary *Eesti etümoloogia sõnaraamat* (2012). The forms of loanword etymons (Estonian, Livonian, Lithuanian, (Middle) Low German, (Old) Russian words) are checked against recent dictionaries of the respective languages: SL; LB; СДЯ XI–XIV; LKŽ, the Livonian-Estonian-Latvian dictionary, etc. The sources of these explanations are not yet indicated but in the future they might be included in the log of theoretical sources.

The references might be divided into four groups. Firstly, the sign ◈↗ is followed by a reference to the word-building method – derivation or compounding, including also references to calque or loan translation, hybrid-coining, or others, such as

⌘ derivatives: *baudīšana* ◈↗ *baudīt* 'to enjoy'; *bodnieks* ◈↗*bode* 'small shop'; Jānīte ◈↗ *Jānis* 'John';

⌘ compounds: *baltvēdris* ◈↗ *balts* + *vēders*; *bezdelīgactiņa* ◈↗*bezdelīga* + *acs*; *Mežamuiža* ◈↗ *mežs* + *muiža* 'forest + estate';

⌘ calques: *apzābakots* ◈↗*ap-* + *zābaks*, from German *gestiefelt* 'with boots on';

⌘ hybrid words: *adatmanis* ◈↗ *adata* + *-manis* < Low German *mann* 'man' 'pin cushion';

⌘ various cases with comments: *aitiņa* ◈↗*avs* – with a missing *v* in the diminutive form *avitiņa* 'sheep'; *aicināt* ◈↗*ej/iet* or interj. *ai*.

In the second group, the sign ◈: is followed by a reference to related words in Latvian, in other Baltic languages, or in other Indo-European languages. Such references are given when the word-building process is obscure and outside the scope of Latvian language history, as in the following cases:

⌘ native words with direct equivalents in other Baltic languages: *adata* 'needle' ◈: Lithuanian *ãdata* 'needle';

⌘ native words with equivalents in other Indo-European languages: *agrs* 'early' ◈: Avestan *agrō* 'first', *agrəm* 'beginning'; *alot* 'to mistake' ◈: Greek αλάομαι 'I wander around'.

In the third group the sign ◈< is followed by a reference to the source, i.e. the etymon. Thus, it is used only with

⌘ lexical borrowings: *boķēt* 'to thresh' ◈< Middle Low German *bōken* 'to beat'; *balandas* 'goosefoot' ◈< Curonian or Lithuanian *balanda* 'goosefoot'; *aba* 'or' ◈< Belarusian or Polish *abo* 'or'.

This group covers most entries of proper names. These are mostly personal names and place-names from religious texts, and their original form in the source language is added. Usually this is from Luther's Bible translation (Luth1545), e.g., *Laodikeja* ◈< Greek *Laodíkeia*; *Liflante* ◈< German *Livland*; *Miha* ◈< German *Micha*.

The fourth group consists of words denoted by the sign ◈. They have unclear or unknown origin. These entries contain more text describing these individual cases. Sometimes parallels in other languages are shown as well:

⌘ *abuls* 'clover' ◈ From *dābuols* 'clover' with change of root and suffix.

⌘ *bokstīties* 'to stagger' ◈ Unclear origin.

⌘ *burkāns* 'carrot' ◈ Unclear origin; cf. Baltic-German *burkan*, Estonian *porgand*, Russian dialects *burkan*, *borkan* 'carrot'.

Of course, such a brief description gives rise to problems that cannot be solved within the LVVV project. They refer either to etymological studies that are outside the scope of the compilation of LVVV, or to theories on morphology that may change in the course of time. It is crucial to use the same reference types both for common names and proper names.

4.7. Cross-references (VII)

At the end of the entry there are cross-references to other entries: ↗ points from a derivation to the basic word, ↘ denotes a reference concerning word-derivation, ↔ denotes a reference of comparison, → denotes a reference to an entry where the respective word can be found. When necessary, comparative references are given from a fixed collocation or idiom to another entry. Due to technical issues that are currently not yet solved, it is not possible to create references in the opposite direction, i.e. from one entry to a collocation quoted in another entry.

4.8. References to literature (VIII)

This section is included only in relatively few entries, and is not present in the entry examples shown in the beginning of this article. In cases where specific literature has been used, e.g. a book on plant names in Latvian (Ēdelmane & Ozola, 2003), the reference to this source is given at the end of the entry.

5. Dictionary writing software

As the project title shows, this dictionary is in a digital format. In order to facilitate the process of its compilation and uploading to the internet, where the dictionary will be supported with various types of search engines,[23] we employ the so-called dictionary writing software. The compilation of new entries is closely linked with mastering the dictionary compilation software *TLex Suit 2013* and its adaptation for the specific needs of the Historical Dictionary of Latvian.[24]

The dictionary writing system helps to structure the entry, to automatize various actions (thus, one no longer has to take care of the visual appearance of the entry each time, or to type the titles of sources, which are instead chosen from a menu). It also helps preparing the dictionary for uploading to the internet with various search options (e.g., the opportunity to search for grammatical parameters, for a particular source, for etymological references, etc.). Nevertheless, this program has its drawbacks. For instance, during the work process one cannot mark problematic places that will need to be revisited. There is a comment log in the end of the entry, but it is not very suitable for our particular needs. Therefore the entries are first prepared in *Word* format, discussed in team meetings, and only then uploaded in the system. Still, the advantages of this lexicographic tool outweigh its disadvantages, and its usefulness is indisputable.

6. Conclusion

Although the four-year long project 'Corpus-based Electronic Historical Dictionary of Latvian' is nearing its end, the work on the dictionary is far from finished. By the end of the project, only 1,200 sample entries will be completed, a small number compared to the number of all the possible entries. However, a large portion of the work will be accomplished, which will make it possible later to resume the compilation of this dictionary – on a new level of quality, within the framework of new projects, with a much larger text corpus as the basis, with new experience with entries of various parts of speech and various sizes, and with modern lexicographic tools. And, most importantly – with the possibility to publish the entries online, so that they are freely available to users and comfortable to use.

---

[23] The possibility of searching and linking is undoubtedly an advantage with electronic dictionaries compared to printed ones (Svensén, 2009, pp. 441–445).

[24] Just like with the creation of the text corpus, the adaptation of the dictionary writing system was done with valuable help from experts from the Institute of Mathematics and Computer Science of the University of Latvia – particularly *Dr.sc.comp.* Normunds Grūzītis and *M.sc.comp.* Ilmārs Poikāns.

Acknowledgements

Abbreviations

ALEW — Hock Wolfgang et al. *Altlitauisches etymologisches Wörterbuch*. 3 Bände. (Studien zur historisch-vergleichenden Sprachwissenschaft 7). Hamburg: Baar-Verlag, 2015.

Br1520_PN — Bruno tēvreize. http://korpuss.lv/senie/static/Br1520_PN.html

CC1585 — *Catechismvs Catholicorum…* Ehsprestcz Vilne Pille Littoutte Semmen pi Danielem Lanciciensem expan tho gadde 1585.

Dres1682_SBM — Dressell Georg. *Swähta Bährno=Mahziba…* Riga: Drukkts und Verlegts Heinrich Bessemesser, 1682.

Elg1621_GCG — Elger Georg. *Geistliche Catholische Gesänge…* Gedruckt zu Braunßberg bey George Schönfels, Anno 1621.

ETS — Metsmägi Iris, Sedrik Meeli, Soosaar, Sven-Erik. *Eesti etümoloogiasõnaraamat*. Tallinn: Eesti Keele Sihtastus, 2012.

Fuer1650_70_1ms — *Lettisches und Teutsches Wörterbuch…* http://korpuss.lv/senie/source.jsp?codificator=Fuer1650_70_1ms

Fuer1650_70_2ms — *Lettisches und Teutsches Wörterbuch…* http://korpuss.lv/senie/source.jsp?codificator=Fuer1650_70_2ms

Fuhr1690_LL — Fohrmann Hinrich. *Laimiga Lauliba…* Riga: Georg. Matth. Nöller, 1690.

Gis1507_PN — *Gisberta tēvreize*. http://korpuss.lv/senie/static/Gis1507_PN.html

Gr1520_PN — *Grūnava tēvreize*. http://korpuss.lv/senie/static/Gr1520_PN.html

Has1550_PN — *Hazentētera tēvreize*. http://korpuss.lv/senie/static/Has1550_PN.html

ILV — Ernštreits Valts, Muzikante Marika, Grīnberga Maima. *Igauņu-latviešu vārdnīca*. Rīga: Latviešu valodas aģentūra; Tallinn: Eesti Keele Sihtastus, 2015.

JT1685 — *Tas Jauns Testaments…* Riga: Johann Georg Wilcken Königl. Buchdr., 1685.

Laz1557_PN — *Lācija tēvreize*. http://korpuss.lv/senie/static/Laz1557_PN.html

LB — Lasch Agathe, Borchling Conrad u. a. *Mittelniederdeutsches Handwörterbuch*. Bd. 1–. Neumünster: Wachholtz, 1928–

LETLS — Viitso Tiit-Rein, Ernštreits Valts. *Līvõkīel-ēstikīel-leţkīel sõnārõntõz*. Rīga: Latviešu valodas aģentūra; Tartu: Tartu Ülikool, 2012–2013.

LGL1685_K1   – *Lettische geistliche Lieder und Kollekten…* Mitau: George Radetzky, 1685.

LGL1685_V5   – *Lettische geistliche Lieder und Psalmen…* Riga: Bey und auff Kosten Georg Matth. Nöllern, 1685.

LKŽ   – *Lietuvių kalbos žodynas.* T. 1–20. Vilnius: Mintis, Mokslas, Lietuviu kalbos instituto leidykla, 1941–2002.

LS1625   – *Linaudēju šrāga.* http://korpuss.lv/senie/static/LS1625.html

Luth1545   – *Luther's Biblia 1545 Original-Text. Ausgabe letzter* Hand. http://lutherbibel. net/

LVIV   – *Latviešu valodas izlokšņu vārdnīca. Prospekts.* Rīga: LU Latviešu valodas institūts, 2005.

LVVA   – Latvijas Valsts vēstures arhīva materiāli.

Manc1631_LVM   – Mancelius Georgius. *Lettisch Vade mecum.* Gedruckt zu Riga durch vnd in Verlegung Gerhard Schröders, 1631.

Manc1631_Syr   – *Das Haus=, Zucht= vnd Lehrbuch Jesu Syrachs* [Tulkojis Georgs Mancelis]. Gedruckt zu Riga in Lieffland durch vnd in Verlegung Gerhard Schröders, 1631.

Manc1637_Sal   – *Die Sprüche Salomonis…* [Tulkojis Georgs Mancelis]. Riga: Gedruckt durch vnd in verleg: Gerhard Schrödern, 1637.

Manc1638_L   – Mancelius Georgius. *Lettus…* Gedruckt vnnd verlegt zu Riga durch Gerhard. Schröder, 1638.

Manc1638_PhL   – Mancelius Georgius. *Phraseologia Lettica…* Riga: Gedruckt vnnd Verlegt durch Gerhard. Schröder, 1638.

Manc1638_Run   – Mancelius Georgius. Zehen Gespräche Deutsch und Lettisch… – *Phraseologia Lettica…* Riga: Gedruckt vnnd Verlegt durch Gerhard. Schröder, 1638.

Manc1654_LP1   – Mancelius Georgius. *Lang=gewünschte Lettische Postill…* I. Riga: Durch Gerhard Schrödern gedruckt und verlegt, 1654.

Manc1654_LP2   – Mancelius Georgius. *Lang=gewünschte Lettische Postill…* II. Riga: durch Gerhard Schrödern gedruckt und verlegt, 1654.

ME   – Mīlenbahs Kārlis. *Latviešu valodas vārdnīca.* Rediģējis, papildinājis, turpinājis J. Endzelīns, I–IV sēj. Rīga: Kultūras fonds, 1923–1932.

Meg1593_PN   – *Megisera tēvreize.* http://korpuss.lv/senie/static/Meg1593_PN.html

SKL1696_KB   – *Sawadi Karra=Teesas Likkumi.* Rigâ, 1696.

SKL1696_RA   – *Sawadi Karra=Teesas Likkumi.* Rigâ, 1696.

SL   – Schiller Karl, Lübben August. *Mittelniederdeutsches Wörterbuch.* Bd. 1–6. Bremen: Vlg. von J. Kühtmann's Buchhandlung, 1875–1881.

SL1684   – *Muhso Augsta un Warrena Kehnina Sohdu=Likkums prett to Behrno=Muschinaschanu.* Riga: Gedruckt bey Georg Matth. Nöllern, 1684.

SOED   – Shorter Oxford English Dictionary. First edition. Volume 1. A.–M. Oxford University Press. New York, 1993.

SSP   – *Słownik staropolski.* Red. S. Urbańczyk. Zesz. 1–76. Wrocław –Warszawa – Kraków – Gdańsk, 1953–2002.

SXVII	– *Elektroniczny słownik języka polskiego XVII i XVIII wieku.* http://sxvii.pl/
VLH1685	– Vermehretes Lettisches Hand=Buch… Mitau: George Radetzky, 1685.
VLH1685_Cat	– Luther, Martin. *Der kleine Catechismus…* durch den Seel: Hrn: Georgium Mancelium… In Lettischer Sprache außgegeben… Mitau: George Radetzky, 1685.
VLH1685_Sal	– *Die Sprüche Salomonis…* Mitau: George Radetzky, 1685.
VLH1685_Syr	– *Das Hauß=Zucht=und Lehr=Buch Jesus Syrachs…* Mitau: George Radetzky, 1685.
WBN	– *Wörterbuchnetz.* http://woerterbuchnetz.de/
Witt1696_MMID	– M.W. In: *Dissertatio Philosophica: De Trinitate Platonis et platonicorum Quam Adspirante divinâ gratia et Consentiente Amplissimó Philosophorum Ordine in almâ ad Emmam Academiâ Gustaviano-Carolina, Sub Praesidio viri celeberrimi Michaelis Dau, Elog. et Poës.Prof.Ordin.u.h.t. Facultatis Philosophicae Decano, Publicè ventilandam proponit Heinrich Christian Wagener Saxo.* In Auditorio Majori MDCXCVI.d.2.Maji Horis&Loco consvetis. Dorpati, Excudit Johannes Bredeken, Acad. Typogr.
ZP1685	– *Semmiga Paklannischana…* [Unknown author's dedication poem]. Drukketa no Jurra Radezka zeeniga Leela Kunga Grahmatu=Drukkera, 1685.
Zv1638_VAR	– Ziņu sniedzēja zvērests Vidzemes arklu revīzijā. http://www.korpuss.lv/senie/static/Zv1638_VAR.html
Zv1681_Kok	– Nevācu koku šķirotāju zvērests. LVVA 1382. fonds, 2. apraksts, 7. lieta, 1a. lpp.
Zv1681_Liec_1	– Liecinieku zvērests LVVA 1379. fonds, 1. apraksts, 27. lieta, 1. lpp.
Zv1681_Liec_2	– *Liecinieku zvērests.* http://korpuss.lv/senie/static/Zv1681_Liec_2.html LVVA 1382. fonds, 2. apraksts, 7. lieta, 1b. lpp.
Zv1698_Lig	– Liģeru amata zvērests. LVVA 1382. fonds, 2. apraksts, 11. lieta, 261.–263. lpp.
Zv1689_Kan	– Kaņepāju kulstītāju amata zvērests. LVVA 1382. fonds, 2. apraksts, 11. lieta, 251.–253. lpp.
Zv1689_Salsnes	– Sālsnesēju amata zvērests. LVVA 1382. fonds, 2. apraksts, 11. lieta, 249.–250. lpp.
СДЯ XI–XIV	– *Словарь древнерусского языка (XI–XIV вв.)* (Vol. 1–10). Москва: Рус. яз., Азбуковник, 1988–2013.
СОРЯМР XVI–XVII	– *Словарь обиходного русского языка Московской Руси XVI–XVII веков,* Санкт-Петербург: Наука, 2004.
СРЯ XI–XVII	– *Словарь русского языка XI–XVII вв.* Москва: Наука, 1975–

Bibliography

Andronova, E. (2007). The Corpus of Early Written Latvian: Current state and future tasks. In *Proceedings of 'Corpus Linguistics 2007'*. Birmingham, UK. Retrieved 20 July 2016, from http:ucrel.lancs.ac.uk/publications/CL2007/paper/245_Paper.pdf

Andronova, E., Siliņa-Piņķe, R., Trumpa, A., & Vanags, P. (2008). Latviešu valodas vēsturiskās vārdnīcas (16.–18. gs.) projekts: pirmā posma rezultāti un turpmākie uzdevumi. In *Letonikas otrais kongress: Valodniecības raksti-1* (pp. 74–89). Rīga: LU Latviešu valodas institūts.

Andronova, E., Trumpa, A., & Vanags, P. (2006). Latviešu valodas vēsturiskās vārdnīcas (16.–18. gs.) projekts: iestrādes un problēmas. In *Letonikas pirmais kongress: Valodniecības raksti* (pp. 9–17). Rīga: Latvijas Zinātņu akadēmija.

Andronova, E., Trumpa, A., & Vanags, P. (2012). Latviešu valodas vēsturiskās vārdnīcas (16.–17. gs.) projekts: Problēmas un risinājumi. In *Apvienotais Pasaules latviešu zinātnieku III kongress un Letonikas IV kongress „Zinātne, sabiedrība un nacionālā identitāte": Valodniecības raksti* (pp. 196–209). Rīga: LU Latviešu valodas institūts.

Augstkalns, A. (2009). *Darbu izlase*. Rīga: LU Latviešu valodas institūts.

Ādamsons, E., & Kagaine, E. (2000). *Vainižu izloksnes vārdnīca* (Vols. A–M, N–Ž). Rīga: LU Latviešu valodas institūts.

Betz, W. (1959). Lehnwörter und Lehnprägungen im Vor- und Frühdeutschen. In F. Maurer & F. Stroh (Eds.), *Deutsche Wortgeschichte* (neubearb. Aufl., Vol. 1, pp. 127–147). Berlin: de Gruyter.

Biezais, H. (1973). Geistliche Lieder von Salomon Gubertus. In *Beiträge zur lettischen Kultur- und Sprachegeschichte* (pp. 69–78). Åbo: Åbo Akademi.

Bībele. (2012). Rīga: Latvijas Bībeles biedrība.

Breidaks, A. (1994). H. Megisers un veclatviešu rakstu valodas sākotne. In *Baltu filoloģija IV. Zinātniski informatīvi ziņojumi* (pp. 15–18). Rīga: Latvijas Universitāte.

Bruno tēvreize [Br1520_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/Br1520_PN.html

Catechismvs Catholicorum… [CC1585]. (1585). Ehsprestcz Vilne Pille Littoutte Semmen pi Danielem Lanciciensem expan tho gadde 1585.

Das Haus=, Zucht= vnd Lehrbuch Jesu Syrachs [Manc1631_Syr]. (1631). Riga: Gerhard Schröder.

Das Hauß=Zucht=und Lehr=Buch Jesus Syrachs… [VLH1685_Syr]. (1685). Mitau: George Radetzky.

Die Sprüche Salomonis… [Manc1637_Sal]. (1637). Riga: Gerhard Schröder.

Die Sprüche Salomonis… [VHL1685_Sal]. (1685). Mitau: George Radetzky.

Dressell, G. (1682). *Swähta Bährno=Mahziba…* [Dres1682_SBM]. Riga: Heinrich Bessemesser.

Dunsdorfs, E. (1962). *Latvijas vēsture 1600.–1710.* Stokholma: Daugava.

Elektroniczny słownik języka polskiego XVII i XVIII wieku [SXVII]. (n.d.). Retrieved 20 July 2016, from http://sxvii.pl/

Elger, G. (1621). *Geistliche Catholische Gesänge…* [Elg1621_GCG]. Braunßberg: George Schönfels.

Ernštreits, V., Muzikante, M., & Grīnberga, M. (2015). *Igauņu-latviešu vārdnīca* [ILV]. Rīga: Latviešu valodas aģentūra.

Ēdelmane, I., & Ozola, Ā. (2003). *Latviešu valodas augu nosaukumi*. Rīga: SIA „Augsburgas institūts".

Fohrmann, H. (1690). *Laimiga Lauliba…* [Fuhr1690_LL]. Riga: Georg. Matth. Nöller.

Fraenkel, E. (1962–1965). *Litauisches etymologisches Wörterbuch* (Vols. 1–2). Heidelberg: Carl Winter.

Fürecker, C. (1650–1670). *Lettisches und Teutsches Wörterbuch…* (The first manuscript) [Fuer1650_70_1ms]. Retrieved 20 July 2016, from http://korpuss.lv/senie/source. jsp?codificator=Fuer1650_70_1ms

Fürecker, C. (1650–1670). *Lettisches und Teutsches Wörterbuch…* (The second manuscript) [Fuer1650_70_2ms]. Retrieved 20 July 2016, from http://korpuss.lv/senie/source. jsp?codificator=Fuer1650_70_2ms

Gisberta tēvreize [Gis1507_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/ static/Gis1507_PN.html

Grūnava tēvreize [Gr1520_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/ static/Gr1520_PN.html

Gubertus, S. (1645). *Stratagema Oeconomicum Oder Ackerstudent Denen jungen Vngeübten Ackerleuten in Lieffland zum nöhtigen Vnterricht…* Riga: Gerhard Schröder.

Hazentētera tēvreize [Has1550_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/ senie/static/Has1550_PN.html

Hock, W. et al. (2015). *Altlitauisches etymologisches Wörterbuch* [ALEW] (Vols. 1–3). Hamburg: Baar-Verlag. (Studien zur historisch-vergleichenden Sprachwissenschaft, 7).

Joffe, D., de Schryver, G.-M., & Prinsloo, D. J. (2003). Introducing TshwaneLex – A New Computer Program for the Compilation of Dictionaries. In G.-M. de Schryver (Ed.), *TAMA 2003 South Africa: Conference Proceedings* (pp. 97–104). Pretoria: (SF)2 Press. Retrieved 20 July 2016, from http://tshwanedje.com/publications/tama-TL.pdf

Jordan, S. (1995). *Niederdeutsches im Lettischen. Untersuchungen zu den mittelniederdeutschen Lehnwörtern im Lettischen*. Bielefeld: Verlag für Regionalgeschichte.

Kagaine, E., & Raģe, S. (1977–1983). *Ērģemes izloksnes vārdnīca* (Vols. 1–3). Rīga: Zinātne.

Kaņepāju kulstītāju amata zvērests [Zv1689_Kan]. (n.d.). LVVA funds 1382, inventory 2, files 11, pp. 251–253.

Karulis, K. (1992). *Latviešu etimoloģijas vārdnīca* (Vols. 1–2). Rīga: Avots.

Kļava, K. (1989). *Die Sprache der lettischen juristischen Dokumente vom Ende des 17. Jahrhunderts: Ein Beitrag zur lettischen Sprachgeschichte*. Stockholm: Almqvist & Wiksell International.

Lasch, A., & Borchling, C. et al. (1928). *Mittelniederdeutsches Handwörterbuch* [LB]. (Vol. 1). Neumünster: Wachholtz.

Latviešu valodas izlokšņu vārdnīca. Prospekts [LVIV]. (2005). Rīga: LU Latviešu valodas institūts.

Lācija tēvreize [Laz1557_PN]. Retrieved 20 July 2016, from http://korpuss.lv/senie/static/ Laz1557_PN.html

Lettische geistliche Lieder und Kollekten… [LGL1685_K1]. (1685). Mitau: George Radetzky.

Lettische geistliche Lieder und Psalmen… [LGL1685_V5]. (1685). Riga: Georg Matth. Nöller.

Liecinieku zvērests 1 [Zv1681_Liec_1]. (n.d.). LVVA funds 1379, inventory 1, files 27, p. 1.

Liecinieku zvērests 2 [Zv1681_Liec_2]. (n.d.). LVVA funds 1382, inventory 2, files 7, p. 1b.

Lietuvių kalbos žodynas [LKŽ]. (1941–2002) (Vols. 1–20). Vilnius: Mintis, Mokslas, Lietuviu kalbos instituto leidykla.

Liģeru amata zvērests [Zv1698_Lig]. (n.d.). LVVA funds 1382, inventory 2, files 11, pp. 261–263.

Linaudēju šrāga [LS1625]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/LS1625.html

Luther, M. (1685). *Der kleine Catechismus… durch den Seel: Hrn: Georgium Mancelium… In Lettischer Sprache außgegeben…* [VLH1685_Cat]. Mitau: George Radetzky.

Luther's Biblia 1545 Original-Text. Ausgabe letzter Hand [Luth1545]. (n.d.). Retrieved 20 July 2016, from http://lutherbibel.net/

M.W. (1696). In *Dissertatio Philosophica: De Trinitate Platonis et platonicorum Quam Adspirante divinâ gratia et Consentiente Amplissimó Philosopho rum Ordine in almâ ad Emmam Academiâ Gustaviano-Carolina, Sub Praesidio viri celeberrimi Michaelis Dau, Elog. et Poës.Prof.Ordin.u.h.t. Facultatis Philosophicae Decano, Publicè ventilandam proponit Heinrich Christian Wagener Saxo* [Witt1696_MMID]. Dorpati: Excudit Johannes Bredeken, Acad.Typogr.

Mancelius, G. (1631). *Lettisch Vade mecum* [Manc1631_LVM]. Riga: Gerhard Schröder.

Mancelius, G. (1638). *Lettus…* [Manc1638_L]. Riga: Gerhard Schröder.

Mancelius, G. (1638). *Phraseologia Lettica…* [Manc1638_PhL]. Riga: Gerhard Schröder.

Mancelius, G. (1638). *Zehen Gespräche Deutsch und Lettisch… – Phraseologia Lettica…* [Manc1638_Run]. Riga: Gerhard Schröder.

Mancelius, G. (1654). *Lang=gewünschte Lettische Postill…* [Manc1654_LP1] (Vol. 1). Riga: Gerhard Schröder.

Mancelius, G. (1654). *Lang=gewünschte Lettische Postill…* [Manc1654_LP2] (Vol. 2). Riga: Gerhard Schröder.

Megisera tēvreize [Meg1593_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/Meg1593_PN.html

Metsmägi, I., Sedrik, M., & Soosaar, S.-E. (2012). *Eesti etümoloogiasõnaraamat* [ETS]. Tallinn: Eesti Keele Sihtastus.

Milčonoka, E. (2003). Latviešu valodas 17. gadsimta teksti internetā. *Baltu filoloģija*, *12*(1), 139–150.

Mīlenbahs, K. (1923–1932). *Latviešu valodas vārdnīca* [ME] (Vols. 1–4). (J. Endzelīns, Ed.). Rīga: Kultūras fonds.

Mīlenbahs, K. (1932). *Latviešu valodas vārdnīca* [ME IV] (Vol. 4). (J. Endzelīns, Ed.). Rīga: Kultūras fonds.

Muhso Augsta un Warrena Kehnina Sohdu=Likkums prett to Behrno=Muschinaschanu [SL1684]. (1684). Riga: Georg Matth. Nöller.

Nevācu koku šķirotāju zvērests [Zv1681_Kok]. (n.d.). LVVA funds 1382, inventory 2, files 7, p. 1a.

Paikens, P., Auzina, I., Garkaje, G., & Paegle, M. (2012). Towards named entity annotation of Latvian National Library corpus. In *Human language technologies – the Baltic perspective* (pp. 169–175). (Frontiers in Artificial Intelligence and Applications, 247). http://doi.org/10.3233/978-1-61499-133-5-169

Pokorny, J. (1959–1969). *Indogermanisches etymologisches Wörterbuch* (Vols. 1–2). Bern: Francke.

Pokrotniece, K. (2012). Grēki un parādi – pamest vai piedot? (Ieskats latviešu senajās tēvreizēs). *Baltistica*, (8 priedas), 111–122. http://dx.doi.org/10.15388/baltistica.0.8.2115

Pretkalniņa, L., Paikens, P., Grūzītis, N., Rituma, L., & Spektors, A. (2012). Making historical Latvian texts more intelligible to contemporary readers. In *Proceedings of LREC 2012 Workshop "Adaptation of Language Resources and Tools for Processing Cultural Heritage Objects"*, Istanbul, Turkey, 2012 (pp. 29–35). Retrieved 20 July 2016, from http://www.lrec-conf.org/proceedings/lrec2012/workshops/13.ProceedingsCultHeritage.pdf

Range, J. D. (1994). Zur Erforschung des Lehngutes in den Baltischen Sprachen. *Linguistica Baltica*, *3*, 211–231.

Reichmann, O. (2012). *Historische Lexikographie*. Berlin: De Gruyter.

Reķēna, A. (1998). *Kalupes izloksnes vārdnīca* (Vols. 1–2). Rīga: Latviešu valodas institūts.

Ross, K., & Vanags, P. (Eds.). (2008). *Common roots of the Latvian and Estonian literary languages*. Frankfurt am Main: Peter Lang.

Rūķe-Draviņa, V. (1977). *The standardization process in Latvian 16ᵗʰ century to the present*. Stockholm: Almquist & Wiksell. (Acta Universitatis Stockholmiensis. Stockholm Slavic Studies, 11)

Sawadi Karra=Teesas Likkumi [SKL1696_KB]. (1696). Rigâ.

Sawadi Karra=Teesas Likkumi [SKL1696_RA]. (1696). Rigâ.

Sālsnesēju amata zvērests [Zv1689_Salsnes]. (n.d.). LVVA funds 1382, inventory 2, files 11, pp. 249–250.

Schiller, K., & Lübben, A. (1875–1881). *Mittelniederdeutsches Wörterbuch* [SL] (Vols. 1–6). Bremen: J. Kühtmann's Buchhandlung.

Sehwers, J. (1918). *Die deutschen Lehnwörter im lettischen* (Inaugural-Dissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät I der Universität Zürich). Zürich: Buchdruckerei Berichthaus.

Sehwers, J. (1953). *Sprachlich-kulturhistorische Untersuchungen vornehmlich über den deutschen Einfluss im Lettischen*. Berlin: Harrassowitz.

Semmiga Paklannischana… [Unknown author's dedication poem] [ZP1685]. (1685). Drukketa no Jurra Radezka zeeniga Leela Kunga Grahmatu=Drukkera.

Shorter Oxford English Dictionary [SOED]. (1993) (1st ed., Vol. 1). New York: Oxford University Press.

Siliņa-Piņķe, R., & Trumpa, A. (2014). Das Historische Wörterbuch der Lettischen Sprache (16.–18. Jahrhundert): Vorstellung und eine Fallstudie. In *B. Bock & M. Kozianka (Eds.), Weiland Wörter-Welten – Akten der 6. Internationalen Konferenz zur Historischen Lexikographie und Lexikologie (Jena, 25.–27. Juli 2012.)* [*Whilom Worlds of Words – Proceedings of the 6th International Conference on Historical Lexicography and Lexicology (Jena, 25–27 July 2012)*] (pp. 261–272). Hamburg: Verlag Gr. Kovač.

Smoczyński, W. (2007). *Słownik etymologiczny języka litewskiego*. Wilno: Universitas.

Summent, A. (1950). *Unbeachtete slawische Lehnwörter im Lettischen* (Inaugural-Dissertation zur Erlangung der Doktorwürde). Göttingen: Georg-August-Universität.

Svensén, B. (2009). *A handbook of lexicography: The theory and practice of dictionary-making.* Cambridge: Cambridge University Press.

Tas Jauns Testaments… [JT1685]. (1685). Riga: Johann Georg Wilcken Königl. Buchdr.

Urbańczyk, S. (Ed.). (1953–2002). *Słownik staropolski* [SSP] (Vols. 1–76). Wrocław, Warszawa, Kraków, Gdańsk.

Vanags, P. (2009). Die Literatur der Letten im Zeichen von Reformation und Konfessionalisierung. In M. Asche, W. Buchholz, & A. Schindling (Eds.), *Die baltische Lande im Zeitalter der Reformation und Konfessionalisierung* (Pt. 1, pp. 263–305). Münster: Aschendorf.

Vanags, P. (2014). Latviešu valodas vēsturiskās vārdnīcas projekts. In *Valodas prakse: Vērojumi un ieteikumi* (pp. 97–109). Rīga: Latviešu valodas aģentūra.

Vermehretes Lettisches Hand=Buch… [VLH1685]. (1685). Mitau: George Radetzky.

Viitso, T.-R., & Ernštreits, V. (2012–2013). *Līvõkīel-ēstikīel-leţkīel sõnārõntõz* [LETLS]. Rīga: Latviešu valodas aģentūra.

Wörterbuchnetz [WBN]. (n.d.). Retrieved 20 July 2016, from http://woerterbuchnetz.de/

Ziņu sniedzēja zvērests Vidzemes arklu revīzijā [Zv1638_VAR]. (n.d.). Retrieved 20 July 2016, from http://www.korpuss.lv/senie/static/Zv1638_VAR.html

Словарь древнерусского языка (XI–XIV вв.) [СДЯ XI–XIV]. (1988–2013) (Vols. 1–10). Москва: Рус. яз., Азбуковник.

Словарь обиходного русского языка Московской Руси XVI–XVII веков [СОРЯМР XVI–XVII]. (2004). Санкт-Петербург: Наука.

Словарь русского языка XI–XVII вв. [СРЯ XI–XVII]. (1975–). Москва: Наука.

Bibliography (Transliteration)

Andronova, E. (2007). The Corpus of Early Written Latvian: Current state and future tasks. In *Proceedings of 'Corpus Linguistics 2007'.* Birmingham, UK. Retrieved 20 July 2016, from http:ucrel.lancs.ac.uk/publications/CL2007/paper/245_Paper.pdf

Andronova, E., Siliņa-Piņķe, R., Trumpa, A., & Vanags, P. (2008). Latviešu valodas vēsturiskās vārdnīcas (16.–18. gs.) projekts: pirmā posma rezultāti un turpmākie uzdevumi. In

Letonikas otrais kongress: Valodniecības raksti-1 (pp. 74–89). Rīga: LU Latviešu valodas institūts.

Andronova, E., Trumpa, A., & Vanags, P. (2006). Latviešu valodas vēsturiskās vārdnīcas (16.–18. gs.) projekts: iestrādes un problēmas. In *Letonikas pirmais kongress: Valodniecības raksti* (pp. 9–17). Rīga: Latvijas Zinātņu akadēmija.

Andronova, E., Trumpa, A., & Vanags, P. (2012). Latviešu valodas vēsturiskās vārdnīcas (16.–17. gs.) projekts: Problēmas un risinājumi. In *Apvienotais Pasaules latviešu zinātnieku III kongress un Letonikas IV kongress „Zinātne, sabiedrība un nacionālā identitāte": Valodniecības raksti* (pp. 196–209). Rīga: LU Latviešu valodas institūts.

Augstkalns, A. (2009). *Darbu izlase*. Rīga: LU Latviešu valodas institūts.

Ādamsons, E., & Kagaine, E. (2000). *Vainižu izloksnes vārdnīca* (Vols. A–M, N–Ž). Rīga: LU Latviešu valodas institūts.

Betz, W. (1959). Lehnwörter und Lehnprägungen im Vor- und Frühdeutschen. In F. Maurer & F. Stroh (Eds.), *Deutsche Wortgeschichte* (neubearb. Aufl., Vol. 1, pp. 127–147). Berlin: de Gruyter.

Biezais, H. (1973). Geistliche Lieder von Salomon Gubertus. In *Beiträge zur lettischen Kultur- und Sprachegeschichte* (pp. 69–78). Åbo: Åbo Akademi.

Bībele. (2012). Rīga: Latvijas Bībeles biedrība.

Breidaks, A. (1994). H. Megisers un veclatviešu rakstu valodas sākotne. In *Baltu filoloģija IV. Zinātniski informatīvi ziņojumi* (pp. 15–18). Rīga: Latvijas Universitāte.

Bruno tēvreize [Br1520_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/Br1520_PN.html

Catechismvs Catholicorum… [CC1585]. (1585). Ehsprestcz Vilne Pille Littoutte Semmen pi Danielem Lanciciensem expan tho gadde 1585.

Das Haus=, Zucht= vnd Lehrbuch Jesu Syrachs [Manc1631_Syr]. (1931). Riga: Gerhard Schröder.

Das Hauß=Zucht=und Lehr=Buch Jesus Syrachs… [VLH1685_Syr]. (1685). Mitau: George Radetzky.

Die Sprüche Salomonis… [Manc1637_Sal]. (1637). Riga: Gerhard Schröder.

Die Sprüche Salomonis… [VHL 1685_Sal]. (1685). Mitau: George Radetzky.

Dressell, G. (1682). *Swähta Bährno=Mahziba…* [Dres1682_SBM]. Riga: Heinrich Bessemesser.

Dunsdorfs, E. (1962). *Latvijas vēsture 1600.–1710*. Stokholma: Daugava.

Elektroniczny słownik języka polskiego XVII i XVIII wieku [SXVII]. (n.d.). Retrieved 20 July 2016, from http://sxvii.pl/

Elger, G. (1621). *Geistliche Catholische Gesänge…* [Elg1621_GCG]. Braunßberg: George Schönfels.

Ernštreits, V., Muzikante, M., & Grīnberga, M. (2015). *Igauņu-latviešu vārdnīca* [ILV]. Rīga: Latviešu valodas aģentūra.

Ēdelmane, I., & Ozola, Ā. (2003). *Latviešu valodas augu nosaukumi*. Rīga: SIA „Augsburgas institūts".

Fohrmann, H. (1690). *Laimiga Lauliba…* [Fuhr1690_LL]. Riga: Georg. Matth. Nöller.

Fraenkel, E. (1962–1965). *Litauisches etymologisches Wörterbuch* (Vols. 1–2). Heidelberg: Carl Winter.

Fürecker, C. (1650–1670). *Lettisches und Teutsches Wörterbuch…* (The first manuscript) [Fuer1650_70_1ms]. Retrieved 20 July 2016, from http://korpuss.lv/senie/source. jsp?codificator=Fuer1650_70_1ms

Fürecker, C. (1650–1670). *Lettisches und Teutsches Wörterbuch…* (The second manuscript) [Fuer1650_70_2ms]. Retrieved 20 July 2016, from http://korpuss.lv/senie/source. jsp?codificator=Fuer1650_70_2ms

Gisberta tēvreize [Gis1507_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/ static/Gis1507_PN.html

Grūnava tēvreize [Gr1520_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/ static/Gr1520_PN.html

Gubertus, S. (1645). *Stratagema Oeconomicum Oder Ackerstudent Denen jungen Vngeübten Ackerleuten in Lieffland zum nöhtigen Vnterricht…* Riga: Gerhard Schröder.

Hazentētera tēvreize [Has1550_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/ senie/static/Has1550_PN.html

Hock, W. et al. (2015). *Altlitauisches etymologisches Wörterbuch* [ALEW] (Vols. 1–3). Hamburg: Baar-Verlag. (Studien zur historisch-vergleichenden Sprachwissenschaft, 7).

Joffe, D., de Schryver, G.-M., & Prinsloo, D. J. (2003). Introducing TshwaneLex – A New Computer Program for the Compilation of Dictionaries. In G.-M. de Schryver (Ed.), *TAMA 2003 South Africa: Conference Proceedings* (pp. 97–104). Pretoria: (SF)2 Press. Retrieved 20 July 2016, from http://tshwanedje.com/publications/tama-TL.pdf

Jordan, S. (1995). *Niederdeutsches im Lettischen. Untersuchungen zu den mittelniederdeutschen Lehnwörtern im Lettischen*. Bielefeld: Verlag für Regionalgeschichte.

Kagaine, E., & Raģe, S. (1977–1983). *Ērģemes izloksnes vārdnīca* (Vols. 1–3). Rīga: Zinātne.

Kaņepāju kulstītāju amata zvērests [Zv1689_Kan]. (n.d.). LVVA funds 1382, inventory 2, files 11, pp. 251–253.

Karulis, K. (1992). *Latviešu etimoloģijas vārdnīca* (Vols. 1–2). Rīga: Avots.

Kļava, K. (1989). *Die Sprache der lettischen juristischen Dokumente vom Ende des 17. Jahrhunderts: Ein Beitrag zur lettischen Sprachgeschichte*. Stockholm: Almqvist & Wiksell International.

Lasch, A., & Borchling, C. et al. (1928). *Mittelniederdeutsches Handwörterbuch* [LB]. (Vol. 1). Neumünster: Wachholtz.

Latviešu valodas izlokšņu vārdnīca. Prospekts [LVIV]. (2005). Rīga: LU Latviešu valodas institūts.

Lācija tēvreize [Laz1557_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/ Laz1557_PN.html

Lettische geistliche Lieder und Kollekten… [LGL1685_K1]. (1685). Mitau: George Radetzky.

Lettische geistliche Lieder und Psalmen… [LGL1685_V5]. (1685). Riga: Georg Matth. Nöller.

Liecinieku zvērests 1 [Zv1681_Liec_1]. (n.d.). LVVA funds 1379, inventory 1, files 27, p. 1.

Liecinieku zvērests 2 [Zv1681_Liec_2]. (n.d.). LVVA funds 1382, inventory 2, files 7, p. 1b.

Lietuvių kalbos žodynas [LKŽ]. (1941–2002) (Vols. 1–20). Vilnius: Mintis, Mokslas, Lietuviu kalbos instituto leidykla.

Liģeru amata zvērests [Zv1698_Lig]. (n.d.). LVVA funds 1382, inventory 2, files 11, pp. 261–263.

Linaudēju šrāga [LS1625]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/LS1625.html

Luther, M. (1685). *Der kleine Catechismus… durch den Seel: Hrn: Georgium Mancelium… In Lettischer Sprache außgegeben…* [VLH1685_Cat]. Mitau: George Radetzky.

Luther's Biblia 1545 Original-Text. Ausgabe letzter Hand [Luth1545]. (n.d.). Retrieved 20 July 2016, from http://lutherbibel.net/

M.W. (1696). In *Dissertatio Philosophica: De Trinitate Platonis et platonicorum Quam Adspirante divinâ gratia et Consentiente Amplissimó Philosopho rum Ordine in almâ ad Emmam Academiâ Gustaviano-Carolina, Sub Praesidio viri celeberrimi Michaelis Dau, Elog. et Poës.Prof.Ordin.u.h.t. Facultatis Philosophicae Decano, Publicè ventilandam proponit Heinrich Christian Wagener Saxo* [Witt1696_MMID]. Dorpati: Excudit Johannes Bredeken, Acad.Typogr.

Mancelius, G. (1631). *Lettisch Vade mecum* [Manc1631_LVM]. Riga: Gerhard Schröder.

Mancelius, G. (1638). *Lettus…* [Manc1638_L]. Riga: Gerhard Schröder.

Mancelius, G. (1638). *Phraseologia Lettica…* [Manc1638_PhL]. Riga: Gerhard Schröder.

Mancelius, G. (1638). *Zehen Gespräche Deutsch und Lettisch…* – Phraseologia Lettica… [Manc1638_Run]. Riga: Gerhard Schröder.

Mancelius, G. (1654). *Lang=gewünschte Lettische Postill…* [Manc1654_LP1] (Vol. 1). Riga: Gerhard Schröder.

Mancelius, G. (1654). *Lang=gewünschte Lettische Postill…* [Manc1654_LP2] (Vol. 2). Riga: Gerhard Schröder.

Megisera tēvreize [Meg1593_PN]. (n.d.). Retrieved 20 July 2016, from http://korpuss.lv/senie/static/Meg1593_PN.html

Metsmägi, I., Sedrik, M., & Soosaar, S.-E. (2012). *Eesti etümoloogiasõnaraamat* [ETS]. Tallinn: Eesti Keele Sihtastus.

Milčonoka, E. (2003). Latviešu valodas 17. gadsimta teksti internetā. *Baltu filoloģija*, *12*(1), 139–150.

Mīlenbahs, K. (1923–1932). *Latviešu valodas vārdnīca* [ME] (Vols. 1–4). (J. Endzelīns, Ed.). Rīga: Kultūras fonds.

Mīlenbahs, K. (1932). *Latviešu valodas vārdnīca* [ME IV] (Vol. 4). (J. Endzelīns, Ed.). Rīga: Kultūras fonds.

Muhso Augsta un Warrena Kehnina Sohdu=Likkums prett to Behrno=Muschinaschanu [SL1684]. (1684). Riga: Georg Matth. Nöller.

Nevācu koku šķirotāju zvērests [Zv1681_Kok]. (n.d.). LVVA funds 1382, inventory 2, files 7, p. 1a.

Paikens, P., Auzina, I., Garkaje, G., & Paegle, M. (2012). Towards named entity annotation of Latvian National Library corpus. In *Human language technologies – the Baltic perspec-*

tive (pp. 169–175). (*Frontiers in Artificial Intelligence and Applications*, 247). http://doi.org/10.3233/978-1-61499-133-5-169

Pokorny, J. (1959–1969). *Indogermanisches etymologisches Wörterbuch* (Vols. 1–2). Bern: Francke.

Pokrotniece, K. (2012). Grēki un parādi – pamest vai piedot? (Ieskats latviešu senajās tēvreizēs). *Baltistica*, (8 priedas), 111–122. http://dx.doi.org/10.15388/baltistica.0.8.2115

Pretkalniņa, L., Paikens, P., Grūzītis, N., Rituma, L., & Spektors, A. (2012). Making historical Latvian texts more intelligible to contemporary readers. In *Proceedings of LREC 2012 Workshop "Adaptation of Language Resources and Tools for Processing Cultural Heritage Objects"*, Istanbul, Turkey, 2012 (pp. 29–35). Retrieved 20 July 2016, from http://www.lrec-conf.org/proceedings/lrec2012/workshops/13.ProceedingsCultHeritage.pdf

Range, J. D. (1994). Zur Erforschung des Lehngutes in den Baltischen Sprachen. *Linguistica Baltica*, 3, 211–231.

Reichmann, O. (2012). *Historische Lexikographie*. Berlin: De Gruyter.

Reķēna, A. (1998). *Kalupes izloksnes vārdnīca* (Vols. 1–2). Rīga: Latviešu valodas institūts.

Ross, K., & Vanags, P. (Eds.). (2008). *Common roots of the Latvian and Estonian literary languages*. Frankfurt am Main: Peter Lang.

Rūķe-Draviņa, V. (1977). *The standardization process in Latvian 16ᵗʰ century to the present*. Stockholm: Almquist & Wiksell. (Acta Universitatis Stockholmiensis. Stockholm Slavic Studies, 11)

Sawadi Karra=Teesas Likkumi [SKL1696_KB]. (1696). Rigâ.

Sawadi Karra=Teesas Likkumi [SKL1696_RA]. (1696). Rigâ.

Sālsnesēju amata zvērests [Zv1689_Salsnes]. (n.d.). LVVA funds 1382, inventory 2, files 11, pp. 249–250.

Schiller, K., & Lübben, A. (1875–1881). *Mittelniederdeutsches Wörterbuch* [SL] (Vols. 1–6). Bremen: J. Kühtmann's Buchhandlung.

Sehwers, J. (1918). *Die deutschen Lehnwörter im lettischen* (Inaugural-Dissertation zur Erlangung der Doktorwürde der hohen philosophischen Fakultät I der Universität Zürich). Zürich: Buchdruckerei Berichthaus.

Sehwers, J. (1953). *Sprachlich-kulturhistorische Untersuchungen vornehmlich über den deutschen Einfluss im Lettischen*. Berlin: Harrassowitz.

Semmiga Paklannischana… [Unknown author's dedication poem] [ZP1685]. (1685). Drukketa no Jurra Radezka zeeniga Leela Kunga Grahmatu=Drukkera.

Shorter Oxford English Dictionary [SOED]. (1993) (1st ed., Vol. 1). New York: Oxford University Press.

Siliņa-Piņķe, R., & Trumpa, A. (2014). Das Historische Wörterbuch der Lettischen Sprache (16.–18. Jahrhundert): Vorstellung und eine Fallstudie. In B. Bock & M. Kozianka (Eds.), *Weiland Wörter-Welten – Akten der 6. Internationalen Konferenz zur Historischen Lexikographie und Lexikologie (Jena, 25.–27. Juli 2012.)* [*Whilom Worlds of Words – Proceedings of the 6ᵗʰ International Conference on Historical Lexicography and Lexicology (Jena, 25–27 July 2012)*] (pp. 261–272). Hamburg: Verlag Gr. Kovač.

Slovar' drevnerusskogo iazyka (XI–XIV vv.) [SDIA XI–XIV]. (1988–2013) (Vols. 1–10). Moskva: Rus. iaz., Azbukovnik.

Slovar' obikhodnogo russkogo iazyka Moskovskoĭ Rusi XVI–XVII vekov [SORIAMR XVI–XVII]. (2004). Sankt-Peterburg: Nauka.

Slovar' russkogo iazyka XI–XVII vv. [SRIA XI–XVII]. (1975–). Moskva: Nauka.

Smoczyński, W. (2007). *Słownik etymologiczny języka litewskiego*. Wilno: Universitas.

Summent, A. (1950). *Unbeachtete slawische Lehnwörter im Lettischen* (Inaugural-Dissertation zur Erlangung der Doktorwürde). Göttingen: Georg-August-Universität.

Svensén, B. (2009). *A handbook of lexicography: The theory and practice of dictionary-making*. Cambridge: Cambridge University Press.

Tas Jauns Testaments… [JT1685]. (1685). Riga: Johann Georg Wilcken Königl. Buchdr.

Urbańczyk, S. (Ed.). (1953–2002). *Słownik staropolski* [SSP] (Vols. 1–76). Wrocław, Warszawa, Kraków, Gdańsk.

Vanags, P. (2009). Die Literatur der Letten im Zeichen von Reformation und Konfessionalisierung. In M. Asche, W. Buchholz, & A. Schindling (Eds.), *Die baltische Lande im Zeitalter der Reformation und Konfessionalisierung* (Pt. 1, pp. 263–305). Münster: Aschendorf.

Vanags, P. (2014). Latviešu valodas vēsturiskās vārdnīcas projekts. In *Valodas prakse: Vērojumi un ieteikumi* (pp. 97–109). Rīga: Latviešu valodas aģentūra.

Vermehretes Lettisches Hand=Buch… [VLH1685]. (1685). Mitau: George Radetzky.

Viitso, T.-R., & Ernštreits, V. (2012–2013). *Līvõkīel-ēstikīel-leţkīel sõnārōntõz* [LETLS]. Rīga: Latviešu valodas aģentūra.

Wörterbuchnetz [WBN]. (n.d.). Retrieved 20 July 2016, from http://woerterbuchnetz.de/

Ziņu sniedzēja zvērests Vidzemes arklu revīzijā [Zv1638_VAR]. (n.d.). Retrieved 20 July 2016, from http://www.korpuss.lv/senie/static/Zv1638_VAR.html

Elektroniczny historyczny słownik łotewski
oparty na korpusie wczesnych tekstów łotewskich

Streszczenie

Artykuł poświęcony jest powstawaniu *Elektronicznego historycznego słownika łotewskiego* (http://www.tezaurs.lv/lvvv) w oparciu o korpus wczesnych tekstów łotewskich (http://www.korpuss.lv/senie/). Omówiono niektóre kwestie odnoszące się do opracowania i przetwarzania danych korpusowych. Ponadto opisano główne źródła dodane do korpusu w okresie czteroletniej pracy nad nim: *Modlitewnik* z XVI w., słowniki, teksty ślubowań i statutów, teksty religijne i tzw. poezja dedykowana z XVII w. Celem projektu jest opracowanie pilotażowego elektronicznego słownika szesnasto- i siedem-

nastowiecznego języka łotewskiego, w którym hasła obejmą wszystkie części mowy. Słownik będzie zawierał około 1200 haseł, w tym nazwy własne i rzeczowniki pospolite.

Główny nacisk położono na opis haseł słownikowych, zawierający istotne uwagi praktyczne i teoretyczne. Omówiono poszczególne części hasła słownikowego, po czym umieszczono komentarz odnoszący się do różnych kwestii związanych z daną częścią (np. wybór hasła wyrazowego i przedstawienia wersji pisowni) i do przyjętych rozwiązań. Szczególną uwagę poświęcono główce hasła, objaśnieniu znaczenia wynikającego z przykładów występujacych w korpusie, różnym rodzajom kolokacji i ich przedstawieniu w słowniku, jak też informacjom etymologicznym. Na końcu zamieszczono zwięzły przegląd oprogramowania słownikowego *TLex 2013*, oparty na doświadczeniu autorów, zdobytym podczas pracy z tym narzędziem.

Słowa kluczowe: *Elektroniczny historyczny słownik łotewski*; korpus tekstów wczesnołotewskich; hasło słownikowe; kolokacje i idiomy; etymologia; odnośniki w słownikach; oprogramowanie słownikowe; *TLex Suit 2013*

The Electronic Historical Latvian Dictionary Based on the Corpus of Early Written Latvian Texts

Abstract

This article deals with the development of the Electronic Historical Latvian Dictionary (http://www.tezaurs.lv/lvvv) based on the Corpus of Early Written Latvian Texts (http://www.korpuss.lv/senie/). Some issues concerning the compilation and processing of the corpus data are discussed and the main sources added to the Corpus during the four-year project are described: the 16th c. Lord's Prayers, 17th c. dictionaries, texts of oaths and laws, religious texts and so-called dedication poetry. The aim of the project is to compile a pilot electronic dictionary of 16th–17th century Latvian where all parts of speech are represented among the entries. This dictionary will contain ca. 1,200 entries, including both proper names and common nouns.

The main emphasis is on the description of the dictionary entries supplied with relevant practical and theoretical observations. Each part of the dictionary entry is discussed, followed by comments on various issues pertaining to that part (e.g., the choice of headword and the representation of spelling versions) and how these were resolved. Special attention is paid to the head of entry, explanation of meaning deduced from the examples found in the corpus, different types of collocations and their representation in the dictionary, as well as etymological information. Finally,

we present a brief review of the dictionary writing software *TLex 2013* based on our experience with this tool.

Keywords: Corpus-based historical dictionary of Latvian; the Corpus or Early Latvian Texts; dictionary entry; collocations and idioms; explanation of origin; cross-references in dictionaries; dictionary writing software; *TLex Suit 2013*

Correspondence:

Everita Andronova, Institute of Mathematics and Computer Science, University of Latvia, Riga, Latvia, e-mail: everita.andronova@lumii.lv

Renāte Siliņa-Piņķe, Latvian Language Institute of the University of Latvia, Riga, Latvia, e-mail: silaziiile@gmail.com

Anta Trumpa, Latvian Language Institute of the University of Latvia, Riga, Latvia, e-mail: antat@latnet.lv

Pēteris Vanags, Latvian Language Institute of the University of Latvia, Riga, Latvia, e-mail: pvanags@latnet.lv