

SENŪJŲ LIETUVIŠKŲ TEKSTŲ RENGIMAS MOKSLO TYRIMAMS

Vilniaus universitetas
2019 m. lapkričio 28–29 d.

Pranešimų tezės

MAKING OLD LITHUANIAN TEXTS USABLE FOR RESEARCH

Vilnius University
28–29 November 2019

Abstracts

Organizuoja
Organizer



Filologijos
fakultetas

Organizacinis komitetas
Organizing committee

Rėmėjas
With the support of



Birutė KABAŠINSKAITĖ
Gina KAVALIŪNAITĖ
Ernesta KAZAKĖNAITĖ

Turinys

Pranešimai:

| | |
|-----------------------------------|----|
| Everita ANDRONOVA..... | 3 |
| Paweł BRUDZYŃSKI | 4 |
| Jolanta BUDRIŪNIENĖ | 5 |
| Jolanta GELUMBECKAITĖ | 6 |
| Armin HOENEN..... | 7 |
| Gina KAVALIŪNAITĖ | 8 |
| Egidijus MILTAKIS..... | 9 |
| Mindaugas ŠINKŪNAS | 10 |
| Aurelija TAMOŠIŪNAITĖ | 11 |
| Felix THIES | 12 |
| Anta TRUMPA, Pēteris VANAGS | 14 |
| Andrius ZILNYS | 15 |
| Vilma ZUBAITIENĖ | 16 |

Kūrybinės dirbtuvės

| | |
|-------------------------|----|
| Ignas RUDAITIS | 17 |
| Wolf-Dieter SYRING..... | 18 |

Everita ANDRONOVA

*Institute of Mathematics and Computer Science, University of Latvia (IMCS UL),
St Petersburg University*

The Corpus of Early written Latvian *SENIE*: From short records to Latvian Bible translation

The Corpus of Early written Latvian comprises 16th–18th c. sources from different genres and of different length (e.g. text of the Lord's Prayer and the translation of the Holy Bible). The corpus is a valuable resource for the study of the history of Latvian writings and serves as the basis of the Historical Dictionary of the Latvian Language (www.tezaurs.lv/lvvv).

This presentation gives a general overview of the content of the corpus, with the main emphasis on shorter texts, namely, Latvian fragments in *Agenda Parva* (1622) and occasional poetry from the 17th and 18th c. The latter represents a new element in the corpus, which until recently included only 5 poems. Before the end of 2019 about 70 new poems from 15 sources will be added to the corpus. Among these is a unicum kept at the Russian National Library in St. Petersburg – the poem 'Mūsu visu upurs tai priecas dienā' (1791). Occasional poetry is of interest not only for specialists of philology, but also broadens our knowledge of culture and historical facts.

Paweł BRUDZYŃSKI
University of Warsaw

Old-Lithuanian irrealis in Chylinski's Bible

The aim of this presentation is to characterize tools and methods I used to describe the conditional, the Lithuanian instantiation of the cross-linguistic category of irrealis, in the Old Lithuanian text of Samuel Bogusław Chyliński's Bible translation.

One interesting and noteworthy fact is that there is a clearly discernible gap between the language of Chylinski's versions of the Old and New Testament. My assumption is that these differences stem from the way the translations were preserved. Published in 1660 The Old Testament most likely must have been edited before it was submitted for publication, although there is no evidence that anyone helped Chyliński with editorial work. I compared both Testaments with a number of books published in the 17th c. and one may assume that the language presented in the Old Testament corresponds to the spoken language of the period. The New Testament translation, of which only the manuscript has survived, is more puzzling, because the language preserved in the New Testament differs from that of the Old Testament in many ways.

The presentation focuses on the methods I used to obtain the data and some difficulties I faced in the initial stages of the research.

Jolanta BUDRIŪNIENĖ

Lietuvos nacionalinė Martyno Mažvydo biblioteka

Skaitmenintas dokumentinis paveldas nacionalinio / regioninio identiteto konstravimui: nacionalinių bibliotekų vaidmuo

Tautinės savivokos klausimas aktualus daugeliui pasaulio šalių, tarp jų ir Lietuvos gyventojams, pastaruoju laiku skatina tyrinėti, gilintis ir analizuoti šalies istorijos procesus, asmens ir jo vaidmens tautos formavimosi raidoje klausimus.

Atminties institucijos, ypač nacionalinės bibliotekos, kurių archyvuose sukauptas išsamiausias kultūros paveldo artefaktų masyvas, yra vienos iš pagrindinių informaciją šios tematikos tyrimams galinčių teikti institucijų.

Skaitmenintas, patogiais būdais pateiktas kultūros paveldo turinys yra ta būtinoji ir pagrindinė sąlyga siekiant plačiau atverti ir paskatinti domėtis savąja praeitimi tiek įvairių sričių tyrėjus, tiek savo praeitimi ir tapatumu besidominčią jaunąją kartą.

Pranešime bus analizuojama, kokias skaitmeninto turinio pateikties paslaugas šiandien pasirengusios teikti bibliotekos, siekdamos priartinti dokumentinio paveldo turinį prie juo besidominčių visuomenės narių, kokie jos narių lūkesčiai. Bus remiamasi naujų šios srities projektų, vykdomų bibliotekose, patirtimis.

Jolanta GELUMBECKAITĖ

Goethe-Universität Frankfurt am Main, Institut für Empirische Sprachwissenschaft

**Altlitauisch Digital: Corpus des Kristijonas Donelaitis (1714–1780)
[CorDon]. Linguistische Annotation**

Das von der Fritz Thyssen Stiftung (Köln) geförderte Projekt *CorDon* (URL: <https://www.fritz-thyssen-stiftung.de/en/fundings/altlitauisch-digital-corpus-des-kristijonas-donelaitis-1714-1780-akronym-cordon/>) setzt sich zum Ziel, die gesamten litauischen Texte von Donelaitis (ca. 160 000 Textwörter) als tiefenannotiertes (d. h. mit einer linguistisch in die Tiefe gehenden Annotation) Corpus aufzubereiten, sie mit einer wechselseitigen Alignierung zu versehen, entsprechende textologische, philologische und linguistische Abfrage- und Analyseprozesse zu ermöglichen sowie eine benutzerfreundliche Suchoberfläche und Ausgabevisualisierung zu schaffen. Ein einheitliches grammatisches Modell für das Altlitauische (Tagset) und das einem historischen Corpus gerechte Annotationsverfahren wurde im Rahmen des SLIEKKAS-Projekts (2012–2014; URL: <http://titus.fkidg1.uni-frankfurt.de/sliekkas/>) entwickelt. Die Donelaitis-Texte werden als ein Corpus aufbereitet, das die folgenden metalinguistischen und linguistischen Informationen enthält: (1) Angaben zur Textstruktur; (2) paläographische bzw. typographische und sonstige textologische Merkmale (als eingeschlossener *apparatus criticus*); (3) lexikalische Angaben: Transliteration, Standardisierung, Lemmatisierung, Glossierung; (4) grammatische Angaben: PoS-Tagging; morphologische bzw. morphosyntaktische Informationen (die unveränderlichen morphologischen Kategorien der Lemmata sowie der konkreten Wortformen; die veränderlichen flexionsmorphologischen Charakteristika der konkreten im Text vorkommenden Wortform). Die Lexik und die Morphologie werden während des Annotationsprozesses zusätzlich eruiert. Die linguistische Annotation erweist sich keineswegs als ein automatisiertes Verfahren. Vielmehr entspricht dieser Prozess der integrierten Entwicklung von Wörterbuch und Grammatik, bei deren Erstellung alle Informationen der schon bestehenden Glossare, Wortindices und Texteditionen revidiert sowie notwendige Korrekturen unternommen werden müssen.

Armin HOENEN

Goethe-Universität Frankfurt am Main, Institut für Empirische Sprachwissenschaft

Altlitauisch Digital: Corpus des Kristijonas Donelaitis (1714–1780) [CorDon]. Digitale Umsetzung

Das Projekt "Altlitauisch Digital: das Corpus des Kristijonas Donelaitis (1714–1780)" das kürzlich von der Fritz Thyssen Stiftung bewilligt wurde, erarbeitet ein linguistisch tiefenannotiertes Corpus des Altlitauischen bestehend aus Primärquellen sowie alignierten Teilen der Editionen Schleichers, Rhesas und Nesselmanns, letztere inklusive deutscher Übersetzung. Wie kann die linguistische Vielfalt am besten repräsentiert und explorabel über ein Web-Interface zur Verfügung gestellt werden? Welche Standards und welche Strukturen sind für eine nachhaltige Repräsentation nötig? Die Kurzpräsentation wird versuchen vorläufige Antworten auf einige dieser Fragen im Projektzusammenhang zu skizzieren. Begleitet wird die Präsentation durch ein funktionales Mockup. Dieses wird als HTML mit JavaScript zur Umsetzung der Nutzerinteraktionen umgesetzt. Die linguistischen Ebenen, die Annotationslayer des im ELAN-Format annotierten Corpus enthalten, werden in syntaktischen Einheiten übereinander visualisiert, wie dies auch im Lesekorpus Altdeutsch umgesetzt ist (siehe z.B. http://titus.uni-frankfurt.de/lea/i_defide_3.html). Daneben wird eine Suchmaske erlauben, die annotierten Texte nach bestimmten Merkmalen wie Kasus, Genus, Wortklasse, etc. und deren Kombinationen zu durchsuchen. Abgerundet wird die Darstellung durch interaktive Visualisierungen die u.a. einen alternativen Zugang zu Ergebnismengen liefern, die so groß sind (z.B. Funktionswörter, hochfrequente Merkmale), dass ein Close-Reading aller Einzelergebnisse mit großem Aufwand verbunden wäre. Schließlich wird eine TEI-konforme Codierung der Daten als Exportformat vorgestellt.

Gina KAVALIŪNAITĖ

Vilniaus universitetas

Samuelio Boguslavo Chylinskio Naujojo Testamento interneto svetainė

Pranešime bus pristatyta šiais metais startavusi Chylinskio Biblijai skirta internetosvetainė, pademonstruotos jos paieškos sistemos ir kitos funkcijos, pasidalinta kūrimo patirtimi. Svetainė www.chylinskibible@ff.vu.lt yra interaktyvi senųjų lietuviškų tekstų interneto svetainė, skirta pirmiausia mokslininkams lietuvių kalbos tyrėjams, tačiau aktuali ir skaitytojams, besidomintiems Biblijos vertimais. Joje pagal pasirinkimą galima naudotis pirmąja ar paskutine rankraščio teksto redakcijomis, išsirinkti vertėjo atliktus redakcinius morfologijos, leksikos, morfosintaksės ir sintaksės pakeitimus; sudaryta galimybė naudotis tiesioginiu, atgaliniu ir dažnumo žodžių formų sąrašais. Svetainėje galima vykdyti paieškas pagal Biblijos eilutę ir pagal rankraščio puslapį; pasirinktas žodis ar teksto fragmentas vienu mygtuko paspaudimu susiejamas su jo kontekstu, o kiekviena Biblijos eilutė susieta su atitinkama vieta Naujojo Testamento rankraštyje.

Egidijus MILTAKIS

Lietuvos kultūros tyrimų institutas

**Lituanistinių tyrimų potencialas ir praktika KU BRIAI skaitmeniniame
archyve – e-fontes.ku.lt**

e-fontes.ku.lt yra Klaipėdos universiteto Baltijos regiono istorijos ir archeologijos instituto kuriamas skaitmeninių istorinių dokumentų kopijų archyvas. Šioje virtualioje saugykloje yra kaupiami įvairaus pobūdžio Vakarų Lietuvos (Žemaitijos ir Mažosios Lietuvos) tyrimams aktualūs istoriniai šaltiniai, kurių chronologinės ribos XIV–XXI a. Šaltinių pobūdis įvairus – katalikų ir evangelikų bažnyčių archyvai, Žemaitijos pilies teismo aktų knygos, Žemaitijos dvarų, Klaipėdos krašto savivaldos archyvai ir kt. Skaitmenindamas senuosius dokumentus KU BRIAI bendradarbiauja su LVIA, LCVA, Lietuvos katalikų Bažnyčia ir kitomis archyvavimo veikla užsiimančiomis organizacijomis, tad e-fontes.ku.lt nuolat pildoma naujomis istorinių dokumentų skaitmeninėmis kopijomis. Lituanistiniams tyrimams šiuo metu labiausiai yra pritaikoma bažnytiniuose metrikuose esanti informacija apie asmenvardžius ir vietovardžius. Šiuos duomenimis KU ir kitų institucijų mokslininkai naudoja atlikdami mokslinius lokalsios istorijos, istorinės geografijos, prozopografinius tyrimus.

Lietuvių senųjų raštų duomenų bazė ir internetiniai leidiniai

Lietuvių kalbos senųjų raštų duomenų bazė pradėta kurti 1996 m. Joje kaupiami XVI–XX amžiaus pradžios lituanistikos ir lituanikos šaltinių skaitmeniniai perrašai ir išvestiniai kūriniai. Skaitmeninimo tekstologinis principas – dokumentinis. Šaltiniai apima spaudinius, rankraščius ir įrašus (antkapiniuose ar kitokuose paminkluose, varpuose ir pan.).

2008 m. sukurta internetinė LKI svetainė „Senieji raštai“ <http://seniejirastai.lki.lt>, kurioje skelbiami šaltiniai, jų žodžių formų konkordancijos, indeksai. Nuo 2018 m. internete veikia duomenų baze paremta eksperimentinė paieškų sistema <http://sr.lki.lt>. Šiuo metu viešai prieinami 75 internetiniai leidiniai, duomenų bazė apima 13 šaltinių. Turimą skaitmeninių tekstų rinkinį iš viso sudaro 425 šaltiniai, kuriuose yra 5,7 mln. žodžių lietuvių kalba ir 1,1 mln. kitomis kalbomis.

Paieškai parengtą skaitmeninį vienetą sudaro šaltinio tekstas ir metainformacija. Žymima šaltinio ar jo fragmentų kalbos autorystė, rašymo ar spaudos vieta, laikas, bibliografinė informacija, puslapiai, spaudos lankai, turinio dalys (skyriai, giesmės, maldos, perikopės ir pan.), Biblijos eilutės, klaidų taisymai ir kt. Pagal šias žymas formuojama duomenų bazė, kurioje kiekvienas šaltinio žodis yra susietas su žymose aprašyta informacija. Atitikmuo dabartine rašyba generuojamas pagal transliteracijos taisykles.

Išvestiniai kūriniai yra tiesioginės ir atvirkštinės žodžių formų konkordancijos bei gramatiniai aprašai. Jų pagrindu rengiami žodžių indeksai ir kiti skaitmeniniai duomenų rinkiniai, skirti lingvistinei analizei. Sukurtos programinių modulių sistemos konkordancijų ir indeksų gamybai, rašybos normalizavimui, Biblijos fragmentų aptikimui, jų panašumo įvertinimui ir sąsajoms, ortografinių ypatybių distribucijos vizualizavimui, generuojami ir pildomi žodynai žodžių formų morfologinei analizei ir lemavimui.

Aurelija TAMOŠIŪNAITĖ

Johano Gutenbergo universitetas Mainz

Duomenynas „Mūsų laiškai“: rengimo principai, iššūkiai ir ateitis

Šio pranešimo tikslas – išsamiau pristatyti lietuviškų laiškų duomenyną „Mūsų laiškai“ (www.musulaiskai.lt). Prieš šešerius metus parengtas, mokslinei (ir ne tik) vartosenai atviras duomenynas yra platesnės, iki šiol vykdomos egodokumentų, visų pirma laiškų kaip empirinių duomenų, kaupimo iniciatyvos dalis. Didžioji dauguma sukauptų egodokumentų (ne tik laiškų, bet ir atvirukų, dienoraščių, užrašų knygelė, atminimų) yra surinkti ir suskaitmeninti iš privačių asmenų kolekcijų, t. y. viešai neprieinami ir nesaugomi archyvuose ar kitose atminties institucijose.

Pirmajame (2013 m. publikuotame) duomenyno leidime mokslinei analizei pateikiama daugiau nei 1300 laiškų, atspindinčių lietuvių rašomosios kalbos vartoseną ir jos kaitą per visą XX amžių. Pranešimo metu trumpai pristatysiu šios lietuviškų laiškų kaupimo iniciatyvos ištakas, didžiausią dėmesį skirsiu duomenyno rengimo principams, jo sandarai bei teikiamoms kalbinių, visų pirma sociolingvistinių, tyrimų galimybėms aptarti. Sociolingvistinės analizės galimybes iliustruosiu *nuoširdumo* leksemų vartosenos atvejo analize. Atsižvelgdama į tarptautinį kontekstą, taip pat šešerių metų darbo su duomenynu patirtį, bandysiu kritiškai (ir retrospektyviai) įvertinti duomenyno rengimo, sandaros, tekstų pateikimo privalumus bei trūkumus, aptarsiu galimas tolimesnes šio empirinio duomenų šaltinio tobulinimo galimybes.

Felix THIES

Halės prie Zalės universitetas

Halės universiteto „Liuterio“ projekto duomenų bazės pristatymas

Projekte *Die Ausstrahlung der Schriften Luthers auf die Sprachen Mittel- und Osteuropas* („Liuterio raštų poveikis Vidurio ir Rytų Europos kalboms“), vykdytame 2017–2019 m. Vokietijos Halės prie Zalės universiteto Germanistikos katedroje, buvo analizuota stiprią protestantišką tradiciją turinčių Rytų Europos kalbų medžiaga. Tarp jų abiejų išlikusių baltų kalbų – lietuvių ir latvių (senosios prūsų kalbos tekstai projektui buvo netinkami) – ir keturių slavų kalbų – aukštutinių ir žemutinių sorbų, kroatų ir slovėnų; taip pat rumunų, estų ir vengrų kalbų tekstai. Korpusą sudaro XVI–XVIII a. Liuterio tekstų vertimai, svarbiausi – Mažasis katekizmas (*Enchiridionas*), Biblija ir postilė.

Kadangi projektas truko gana trumpą laiką ir nebuvo galima išsamiai ištirti visų šaltinių, jo tikslas buvo apžvelgti įvairias Rytų Europos kalbų vertimo strategijas, atlikti kiekybinę analizę ir nustatyti Liuterio įtakos mastą skirtingoms sritims bei kalboms.

Tyrimo rezultatai bus paskelbti internetinėje duomenų bazėje su įvairiomis paieškos funkcijomis.

1) Žodžių daryba

Buvo parinkti 26 būdvardžiai ir 35 daiktavardžiai, kurie germanistų laikomi specifiniais Liuterio dariniais, pvz., dūrinys *Menschenkinder* („žmonių vaikai“) arba *kleingläubig* („netvirtai tikintis“, pažodžiui „mažatikis“).

Dūriniai, pvz., *Menschenkinder*, į lietuvių kalbą dažniausiai išversti junginiais su atributiniu daiktavardžiu (*žmonių vaikai, vaikai žmonių*), retai su atributiniu būdvardžiu (*nepakajingos dvasios*, Liuterio *Flattergeist* „plazdanti dvasia?“). 44% atvejų genityvas eina po valdančio daiktavardžio (*vaikai žmonių*), 30% eina prieš jį (*žmonių vaikai*), o 6% pavyzdžių su daiktavardžiu derintas būdvardis. Kitaip latvių kalboje, kurioje visais atvejais genityvas eina prieš valdantį žodį (*Žilweko Behrņi*).

Priesaginių būdvardžių atvejais buvo pastebėtos trys pagrindinės vertimo strategijos:

- a) kaip pirminis (nepriesaginis) būdvardis (*bos-haft-ig* → *piktas*),
- b) kaip priesaginis būdvardis (*fried-lich* → *pakaj-ingas*),
- c) kaip dūrinys (*leicht-herz-ig* → *lengva-šird-ingas*).

Paskirstymas lietuvių kalboje – a) 21%, b) 38%, c) 13 %, latvių – a) 11%, b) 46%, c) 14%.

Duomenų bazė suteiks galimybę ieškoti specifinės priesagos arba žodžio struktūros.

2) Sintaksė

Sintaksės srities pagrindinis šaltinis – pasakojimas apie Jėzaus gimimą (Lk 2,1–20). Be žodžių tvarkos, buvo analizuotas parodomųjų įvardžių vartojimas (lyginant su vokiečių kalbos artikeliu) ir dalyvių konstrukcijų vartojimas (pvz., lot. *laudantium et dicentium* „pagiriančių ir sakančių“ – vok. *die lobten und sprachen* „tie pagyrė ir sakė“, liet. [VEE] *garbinancziu Diewa / ir bilancziu*, bet lat. [GB] *tee flaweja un la33ija* „tie pagyrė ir sakė“).

Kiekybinės analizės rezultatai vizualizuoti tinklalapyje.

3) Metrinė struktūra ir rimas

Korpusas apima ir rinktinės Liuterio sukurtas giesmes, tarp jų *Vom Himmel hoch*. Lietuvių kalbos medžiaga imta iš Mažvydo, Vilento, Zengštoko ir Kleino giesmynų. Duomenų bazėje aprašyta metrinė struktūra ir skiemenų skaičius, rimo schema.

Anta TRUMPA¹, Pēteris VANAGS²

¹*Latvijas universiteto Latvių kalbos institutas,*

²*Latvijas universitetas, Stokholmo universitetas*

„Latvių kalbos istorinis žodynas (XVI–XVII a.)“: atliktieji ir šio meto darbai

2016 m., baigiantis ketverių metų Latvijos mokslo tarybos projektui, buvo publikuotas elektroninis „Latvių kalbos istorinis žodynas (XVI–XVII a.)“ („Latviešu valodas vēsturiskā vārdnīca (16.–17. gs.)“) (<https://tezaurs.lv/lvvv/>), kurio pagrindu sudaro „Latvių kalbos senųjų tekstų korpuso“ („Latviešu valodas seno tekstu korpus“, (<http://senie.korpuss.lv/>)) medžiaga. Žodyne šiuo metu publikuoti 1182 pavyzdiniai žodyno straipsniai, tai, žinoma, nėra didelis kiekis tokio tipo leidiniui. Vis dėlto svarbu yra tai, kad kuriant straipsnius buvo parengta ir elektroninio žodyno internetinė svetainė, teikianti plačias paieškos galimybes. Joms užtikrinti žodynas rašomas leksikografams skirta programa *TLex 2013*, specialiai pritaikyta darbui su „Latvių kalbos istoriniu žodynu“.

Šiomet po trejų metų pertraukos „Latvių kalbos istorinio žodyno“ kūrimas buvo atnaujintas remiant Valstybinei tyrimų programai. Šiuo metu žodyno autoriai dėmesį daugiausiai kreipia į pačius straipsnius, mat žodyno sandara, straipsnio struktūros kūrimas ir leksikografo darbo vietos pritaikymas buvo ankstesnio etapo uždavinys. Pradėta dirbti nuo pačios žodyno pradžios, a raidės, toliau numatyta rašyti straipsnius iš eilės, neatsižvelgiant į kalbos dalį ar straipsnio apimtį. Kuriant pavyzdinius straipsnius buvo laikytasi „lizdų“ principo, šalia pagrindinio žodžio straipsnio pridedant ir visus XVI–XVII a. tekstuose pasitaikiusių šio žodžio vedinių straipsnius.

Remiantis naująja tvarka, t. y. rengiant visus straipsnius iš eilės, autoriai iš pradžių bandė atsisakyti „lizdų“ principo, vis dėlto apdorojant medžiagą pasirodė, kad šio principo iš dalies pravartu laikytis, nes vienoje vietoje apžvelgiant visą vienos šaknies žodžių vartoseną senuosiuose tekstuose, neretai atskleidžiamos tikslesnės jų reikšmės. Pranešime bus smulkiau pažvelgta į keletą žodyno straipsnių kūrimo problemų, jos bus iliustruojamos pavyzdžiais.

Andrius ZILNYS

Lietuvos vyriausiojo archyvaro tarnyba

**Projekto „Elektroninės paslaugos EAIS ‘Skaitmeninė skaitykla’ sukūrimas“
pristatymas**

Projekto tikslas – Elektroninio archyvo informacinės sistemos infrastruktūroje sukurti tris elektronines paslaugas: „Skaitmeninė skaitykla“, „Genealogija“ ir „Asmenvardžių archyvas“, kurios užtikrintų vieningą prieigą prie valstybės archyvuose sukauptų skaitmeninių išteklių, didintų valstybės archyvuose saugomų dokumentų pasiekiamumą, vystytų ryšį su suinteresuotųjų naudotojų grupėmis.

XVIII a. Mažosios Lietuvos žodynai ir Biblija

Pranešime bus tiriama, kokia Biblijos leksika yra patekusi į tris spausdintus Mažosios Lietuvos lietuvių–vokiečių ir vokiečių–lietuvių kalbų žodynus: F. V. Hako „Vocabularium“ (1730), P. Ruigio „Lexicon“ (1747) ir K. G. Milkaus „Wörterbuch“ (1800) ir vieną rankraštinį vadinamąjį Krauzės žodyną. Hako žodyne nėra pateikta 1727 m. Naujojo Testamento ir 1728 m. Psalmyno nuorodų, todėl šaltinių įtaką galima atsekti tik lyginamosios ir tekstologinės analizės būdu. Ruigio ir Milkaus žodynuose prie lietuvių kalbos žodžių ir junginių vokiečių–lietuvių kalbų dalyse yra nurodoma konkrečiai vartosenos 1735 m. Biblijoje vieta. Deja, nuorodos dažnai klaidingos. Kai kur žodynuose pateikiamos kitokios nei Biblijoje žodžių formos. Vadinamajame Krauzės žodyne galima rasti senesnių Biblijos vertimų žodžių. Biblijos ir žodynų tarpusavio tyrimus palengvino suskaitmeninti 1727 m., 1728 m., 1735 m. Biblijos ir minėtų žodynų leidimai bei automatinis eilučių gretinimo įrankis, sukurtas Lietuvių kalbos institute.

Kūrybinės dirbtuvės

Ignas RUDAITIS
Vilniaus universitetas

TEI ir XML dirbtuvės

TEI (*Text Encoding Initiative*) yra skaitmeninis formatas, skirtas perteikti tekstus ir juos anotuoti įvairiais filologiniais aspektais. Šis formatas yra tapęs standartu skaitmeninės humanitarikos srityje. Seminaro metu vyks dviejų dienų dirbtuvės, per kurias dalyviai susipažins su TEI formatu – išmėgins jį su pavyzdiniais tekstais, susipažins su jo techninėmis prielaidomis bei išmoks atpažinti jo teikiamą naudą.

Pirmąją dirbtuvių dieną taikydami lingvistines analogijas susitelksime ties hierarchinėmis (medžio pavidalo) struktūromis, vadinamosiomis jų schemomis, ir TEI kaip konkrečia medžių struktūrų schema. Paprastus tekstus reprezentuosime TEI medžiais. Antroji diena bus skirta XML kalbai, kuri leis šiuos (ir bet kokius kitus) medžius įvesti klaviatūra ir reprezentuoti kaip paprastą skaitmeninį tekstą. Taip būsime išgavę keletą paprastų, bet visaverčių TEI dokumentų. Baigsime aptardami įvairios paskirties programinę įrangą, kuri apdoroja būtent TEI formato dokumentus.

Wolf-Dieter SYRING
Buxtehude, Vokietija

Essentials while preparing electronic editions of historical texts

Within digital projects based on historical texts, two sides are usually involved, on the one side content-oriented (philology, linguistics, paleography) and on the other side data-oriented professions. The sessions of the workshop will rise the challenge of a "real" project, i.e. the attendents will work on structuring and planning a concrete project, one on a manuscript and one on an early print. In both sessions, after an introduction the audience will be split into two groups (for each type of the project). The essential tasks during the process of structuring and planning the project have to be identified and to be managed. The first session will focus on structuring the content and on calculating the schedule of the project; the second one will add more detailed elements of source representation and professional cooperation. In the end of each session, the results are summarized and weighted regarding their importance.

