

# A flexible language learning platform based on language resources and web services

Elena Volodina\*, Ildikó Pilán\*, Lars Borin\*, Therese Lindström Tiedemann†

\*Språkbanken, University of Gothenburg, Sweden

Språkbanken, Institutionen för svenska språket, Göteborgs universitet, Box 200, 405 30 Göteborg, Sweden

†Department of Linguistics and Philology, Uppsala University, Sweden

Department of Linguistics and Philology, Uppsala University, Box 256, 751 05 Uppsala, Sweden

[elena.volodina@svenska.gu.se](mailto:elena.volodina@svenska.gu.se), [ildiko.pilan@svenska.gu.se](mailto:ildiko.pilan@svenska.gu.se), [lars.borin@svenska.gu.se](mailto:lars.borin@svenska.gu.se),  
[therese.lindstrom\\_tiedemann@lingfil.uu.se](mailto:therese.lindstrom_tiedemann@lingfil.uu.se)

## Abstract

We present Lärka, the language learning platform of Språkbanken (the Swedish Language Bank). It consists of an exercise generator which reuses resources available through Språkbanken: mainly Korp, the corpus infrastructure, and Karp, the lexical infrastructure. Through Lärka we reach new user groups – students and teachers of Linguistics as well as second language learners and their teachers – and this way bring Språkbanken's resources in a relevant format to them. Lärka can therefore be viewed as an case of real-life language resource evaluation with end users. In this article we describe Lärka's architecture, its user interface, and the five exercise types that have been released for users so far. The first user evaluation following in-class usage with students of linguistics, speech therapy and teacher candidates are presented. The outline of future work concludes the paper.

**Keywords:** exercise generator; ICALL; Lärka

## 1 Introduction

Lärka<sup>1</sup> is the ICALL platform of Språkbanken (the Swedish Language Bank). ICALL – Intelligent Computer-Assisted Language Learning – has as its main aim to draw on the opportunities offered by language resources, such as corpora, lexicons and natural language processing (NLP) components including lemmatizers, parsers, etc., to build more sophisticated and flexible applications for language learners and students of grammatical theory (Meurers, 2012; Amaral and Meurers, 2011; Heift and Schulze, 2007).

ICALL is an active area of research with such examples of end-user applications as E-tutor (Heift, 2010), Criterion and E-rater (Burstein et al., 2013), Werti (Meurers et al., 2010), TAGARELA (Amaral and Meurers, 2001). Research within ICALL is advancing along multiple dimensions, such as text and sentence readability (Pilán and Volodina, 2013; Shen et al., 2013), text simplification (Vajjala and Meurers, 2014), mother tongue identification (Li, 2013), learner error detection (Cahill et al., 2013), essay scoring (Östling et al., 2013), exercise generation (Dickinson and Herring, 2008), semantic analysis of learner production (King and Dickinson, 2013), etc. Functionality and appropriateness of ICALL applications depend linearly upon advances in these and other related areas.

Språkbanken has a long history of ICALL R&D, and Lärka combines and extends the capabilities of two earlier

applications, ITG (Saxena and Borin, 2002; Borin and Saxena, 2004) and SCORVEX (Volodina, 2010). The ITG platform was developed in the early 2000s and explicitly targeted students of linguistics. Its aim was to offer grammar – part-of-speech (POS) and syntactic-relation – exercises based on authentic language examples from annotated corpora. ITG was used on linguistics courses at several Swedish universities during the years 2005–2012 (Saxena and Lind, 2008), and sporadically alongside Lärka in 2013. SCORVEX was built as an ICALL application for students of Swedish as a foreign or second language, offering a variety of vocabulary exercises. In a rapidly evolving digital world, the technologies used in these applications have been superseded, and the development of Lärka was a natural consequence of a general restructuring of Språkbanken's language resource and technology infrastructure into one based on distributed (REST) web service components and web applications. At present, in addition to Lärka, this infrastructure comprises Korp<sup>2</sup> (Borin et al., 2012b) for text corpora, and Karp<sup>3</sup> (Borin et al., 2012a) for lexical resources.

1 <[spraakbanken.gu.se/larka/](http://spraakbanken.gu.se/larka/)>. Lärka is an acronym for *LÄR språket via KorpusAnalys* 'Learn language via corpus analysis'. The word itself also means 'lark', and it corresponds in English to *Language Acquisition Reusing Korp*.

2 Korp (Borin et al., 2012b): <[spraakbanken.gu.se/korp/](http://spraakbanken.gu.se/korp/)>

3 Karp (Borin et al., 2012a): <[spraakbanken.gu.se/karp/](http://spraakbanken.gu.se/karp/)>

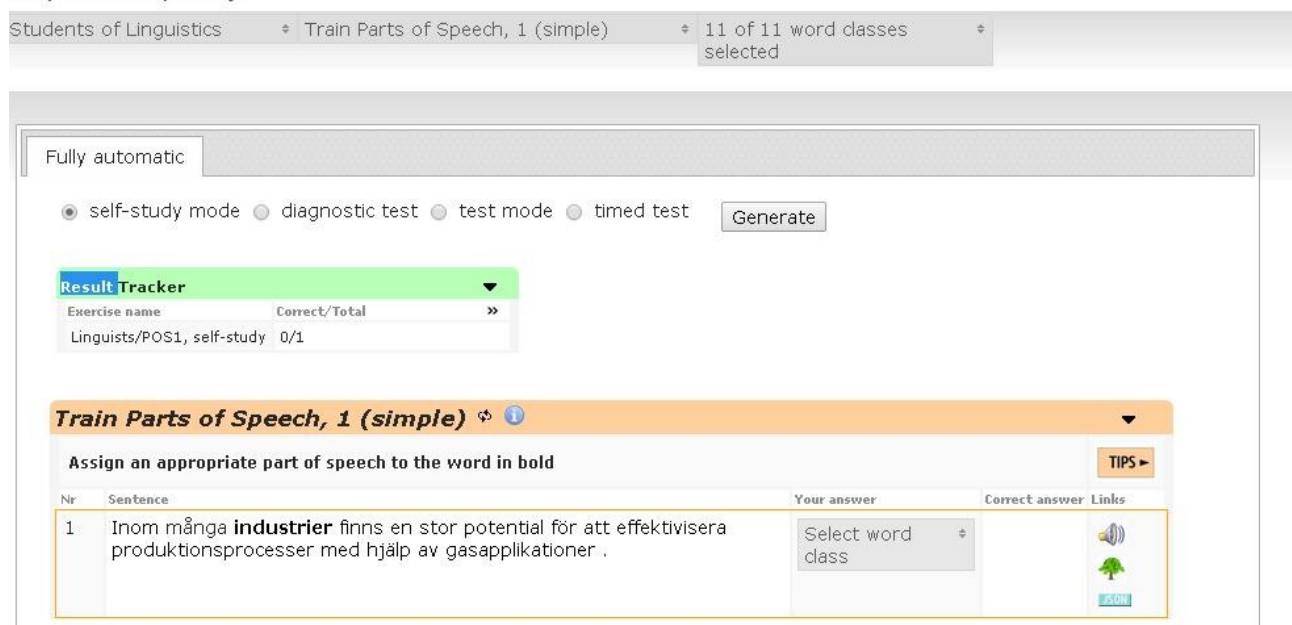


Figure 1. Lärka – user interface

The work on Lärka started in the project *Systems Architecture for ICALL* financed by NordPlus Sprog (2011–2013). Specified as a modular web-based exercise generator that reuses available annotated corpora and lexical resources, Lärka is freely available, targeting primarily learners of Swedish as a second/foreign language and students of (Swedish) linguistics. Being web-based, Lärka has advantages of accessibility and ease of use compared to its predecessors. With the release of Lärka 1.0 (October, 2013) ITG and SCORVEX have been “retired”.

Below, we give an overview of Lärka (section 2), summarize feedback from the first in-class uses (section 3) and conclude the paper outlining future plans (section 4).

## 2 Lärka in a nutshell

Lärka is designed as a *Service Oriented Architecture* based on web services. The platform comprises two main components – user interface and web services – where the web services can be reused by other applications (Volodina et al., 2012). Web services take care of exercise generation whereas the user interface collects user input, formats the web service output and assigns behavior to buttons and menus.

At the moment Lärka offers exercises for two target groups: students of linguistics and learners of Swedish. All available exercises share some **common features**, namely:

- *Training context*: sentence. The objective with the Lärka-based exercise generator has, from the onset, been to use real-life language examples from corpora. Possible

copyright issues are avoided by using only a single-sentence context.

- *Format*: multiple-choice. The target item (one word or phrase) is marked in the sentence. An accompanying drop-down menu contains several answer alternatives, only one of which is correct.

- *Reference materials*. Relevant articles are looked up in Wikipedia, Wiktionary and Karp, while a text-to-speech module provided by SitePal<sup>4</sup> offers pronunciation of relevant words and sentences. Reference materials are shown in a separate field that can be hidden when not wanted.

- *Training modes*: self-study, test and timed. The self-study mode reveals all clues (e.g. reference articles, syntactic tree structure, pronunciation, etc.) and also provides a possibility to try several answer options. In the test and timed test modes, the clues are not revealed until the answer is provided; and users cannot change their answer. In the timed test there is also an additional pressure of time set on the user.

- *Feedback* is offered in the form of immediate correct/incorrect symbols and a result tracker where information on correct/total number of answers is shown.

- *A new item* is generated as soon as the previous one is answered.

- To avoid *sentence duplicates*, the same sentence is never selected more than once during the

4 <sitepal.com/>

same exercise session.

- Finally, an *i-icon* provides information about each exercise type.

The screenshot shows the Lärka language learning interface. At the top, there is a navigation bar with 'Svenska | English | About Lärka'. Below that, a breadcrumb trail reads 'Language learners > Train inflections, multiple-choice > General purpose vocabulary > 3 of 11 word classes selected > 1 of 6 learner levels selected >'. The main content area is titled 'Fully automatic' and includes a 'Generate' button. A 'Result Tracker' shows 'Learners/inflections, test' with a score of '0/1'. The exercise title is 'Train inflections, multiple-choice'. The instruction is 'Choose an appropriate word form for the gap'. A table shows a sentence: '1 Hon är i alla fall inte \_\_\_\_\_.' A 'Select a word' dropdown menu is open, showing a list of Swedish inflections: 'bitter', 'bittert', 'bittra', 'bittrare', 'bittrast', and 'bittraste'. A 'TIPS' button is visible next to the dropdown.

Figure 2. Lärka - exercise with inflectional paradigms

The exercise repertoire for **students of linguistics** comprises exercises for training POS, syntactic relations and semantic roles. Common features for all the three exercise types include sentence filtering consisting in allowing only sentences of 5–20 tokens in length that are non-elliptic, i.e. contain a subject and a finite verb and which do not contain subordinate clauses, except when looking for sentences with subjunctions.

Users have an option of selecting one or a set of categories (e.g. POS) from a menu for training. Other features include statistics over the answers, diagnostic test and adaptivity which ensures that the categories that cause more difficulties return more often. A diagnostic test option gives the student a number of questions to test their knowledge of parts of speech, syntactic relations or semantic roles. After the test the student is given automatic feedback and recommendations what to focus on. From the result tracker the student can also choose to view more explicit feedback where the test sentences are listed, as well as the right and wrong answers.

The *POS* exercise (Figure 1) aims at students who want to practice differentiating between POS. The following POS and sub-categories are used in the exercise: adjective, adverb, conjunction, subjunction, determiner, nouns (incl. proper names), numerals, participles, prepositions (incl. particles), pronouns and verbs. All sentences are selected from SUC3.0 (Stockholm-Umeå Corpus – Ejerhed et al., 1992; Källgren, 2006), a corpus manually annotated for POS.

The *syntactic-relation* exercise is aimed at students who need to train and revise clause-level syntactic roles. The following 7 relations and sub-categories are used in the exercise: adverbial, finite and non-finite verbs, predicative,

direct and indirect objects, and subject. The sentences are selected from Talbanken (Teleman, 1974; Einarsson, 1976), a corpus manually annotated for syntactic relations.

The *semantic-role* exercise provides training for understanding the semantic relations in a sentence. There are 12 general roles, each of which encompasses a group of semantically related sub-roles coming from the role set used in the Swedish FrameNet (SweFN; Borin et al., 2010): Agent, Experiencer, Theme, Instrument, Location, Goal, Recipient, Origin, Time, Manner, Purpose, Cause (see Pilán and Volodina, 2014). All of the sentences come from SweFN, where they have been carefully selected from corpora and manually annotated for semantic roles.

**Learners of Swedish** are offered two exercise types: *training vocabulary knowledge* and *training inflectional paradigms*. Figure 2 shows an example of the latter. *Both exercise types* offer the option of selecting a domain for target vocabulary as well as learner proficiency levels. The selection of target vocabulary comes from three main sources:

- From *frequency-based lists*, such as the Swedish Kelly list (Volodina and Johansson Kokkinakis, 2012) and the Swedish Academic Word List (AO; Sköldberg and Johansson Kokkinakis, 2012). In this case words may combine with POS restrictions to get a subset of vocabulary from this resource, e.g. general purpose vocabulary (Kelly list) + verbs. The inflectional paradigm exercise, however, targets only three POSs, namely nouns, verbs and adjectives. Kelly words are assigned to 6 different proficiency levels according to the Common European Framework of

Reference (CEFR; COE, 2012), which also makes it possible to select vocabulary from a certain proficiency level/frequency band. Words from AO always correspond to advanced proficiency (C2).

- From 30 LEXIN (Gellerstam, 1999) domain lists. In this case we do not offer an option of filtering target vocabulary items for POSs as domain vocabulary in Lexin Picture Series contains mainly nouns, and in certain domains some verbs, adjectives and adverbs.

In both exercise types sentences are selected according to the proficiency level specified by the user. For that, a special Lärka-based sentence readability module, *HitEx* “*Hit Examples*”, currently available for intermediate level [B1] and above (Pilán et al., 2013; Pilán, 2013). The module selects and ranks corpus hits either based on heuristic rules only or using a combination of rules and classification with machine learning. To assess the readability of sentences, a number of morpho-syntactic (e.g. average dependency length) and lexical-semantic features (e.g. CEFR level and frequency of words) are taken into consideration. The rules offer the possibility also to filter sentences containing certain linguistic elements including, among others, abbreviations, negative formulations and participles.

Sentences are selected from three different corpora to cater for a combination of different genres, namely SUC3.0 (a balanced corpus with texts from various genres), GP2012 (newspaper texts) and ROM99 (novels). Sentences for training vocabulary coming from AO are selected from specialized corpora comprising academic texts in the areas of the humanities and social sciences (Sköldbberg and Johansson Kokkinakis, 2012).

The two exercise types differ in how the *distractors* (incorrect alternatives) are selected. For vocabulary training, words of the same morphosyntactic tag are selected, whereas in the inflectional paradigm exercise a morphology web service provides different word forms.

### 3 Lärka in use – initial user experiences

Lärka has been used by Uppsala University for the past three terms (2013-2014) and by the University of Gothenburg during Spring Term 2014, to teach first-year students of linguistics, speech therapy and language technology, as well as trainee teachers of upper-secondary Swedish and Swedish as a Second Language.

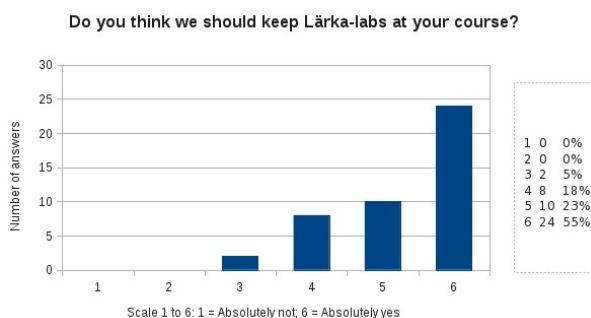


Figure 3. Evaluation results 1

The aim of the general introductory course to Linguistics<sup>5</sup> was to enable students to identify, among others, POS, syntactic relations, main/embedded clauses and semantic roles. During three laboratory sessions, one for POSs, the other for syntactic relations, and the last for semantic roles, students have been working individually and in small groups with the exercises. Although lab sessions were optional, student attendance was very high. Students were also encouraged to continue training with Lärka at home.

The labs have been very positively accepted by students (Figure 3). Out of 45 answers, 34 students (78%) have commented in favour of using Lärka as part of the courses (scores 5–6 on the scale of 1 to 6), while 10 students (22%) have been reserved about it (scores 3–4). Similarly to written evaluations of ITG (Saxena and Lind, 2008), students found the labs fun and instructive, and they appreciated the opportunity to get individual help from the teacher as well as receiving instantaneous feedback from the program. They trained primarily in self-study mode, with initial and final diagnostic tests, and appreciated the real-life challenge presented by the sentences coming from authentic texts and the possibility to consult various sources of reference. Students as well as teachers, who were initially skeptical, found Wikipedia articles useful. The existence of contradictory views on certain aspects of linguistics sparked some lively discussions which proved to lead to a better understanding of the complexity of language (and linguistics).

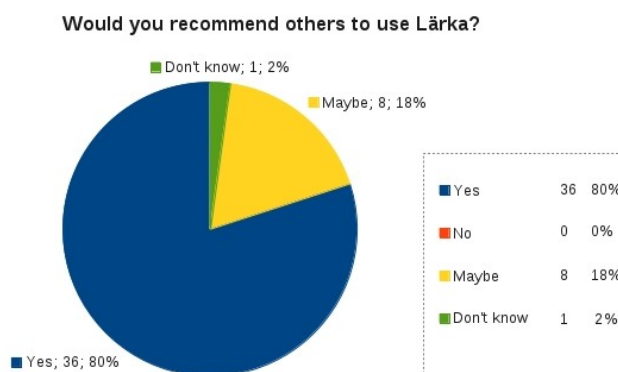


Figure 4. Evaluation results 2

Lärka was generally perceived by students as a very useful and instructive complimentary tool. 80% of the students would recommend the tool for others to use, followed by 18% of those who might recommend it, compared to 2% who were uncertain (Figure 4). In a number of written (optional) comments students wrote that they would definitely continue practicing with Lärka outside the classroom, that it was a very easy-to-use tool and that they appreciated an opportunity to step away from paper-bound training to a more fun way of

5 <[www.uu.se/utbildning/utbildningar/selma/kursplan/?kpid=28202&lasar=13%2F14&tyt=1](http://www.uu.se/utbildning/utbildningar/selma/kursplan/?kpid=28202&lasar=13%2F14&tyt=1)>

testing and improving their linguistic knowledge. Teachers have appreciated having an inexhaustible source of real-life sentences demonstrating different linguistic features. They have specifically valued that these labs made students think critically and contributed to the acquisition of a deeper and more long-term knowledge (cf. Saxena and Lind, 2008). In addition, teachers got insight into which elements students found difficult and, thus, had an opportunity to identify concepts in need of clarification. We received comments, however, concerning the terminology used in Lärka which would need to be explicitly explained to avoid clashes between course book terminology and categories used in Lärka (e.g. *articles* versus *determiners*; *prepositional object* versus *object adverbial*; *predicate* versus *finite and non-finite verbs*). Other desirable features that have been named are: easier sentences and an option to have an arbitrary sentence analyzed. Exercises for morphological and phonological analysis are two other points that have been added to our to-do-list after these labs.

#### 4 Concluding remarks

In this paper we have presented the Lärka version that was released in October 2013 and summarized the first experiences of using it. Lärka has become a real-life test for a number of Language Technology resources (SweFN, Korp corpora, Karp lexicons, text-to-speech module) and stimulated development of new language technology algorithms (e.g. first experiments with HitEx).

A number of additional modules are under active development:

- a sentence readability module, HitEx, to automatically determine the difficulty-level of sentences with planned extension to all CEFR levels;
- a corpus editor that is being used for the annotation of CEFR-based course books (Volodina et al., 2013);
- a dictation and spelling exercise where target items at different linguistic levels (lemmas, inflected word forms, phrases and sentences) are pronounced using text-to-speech technology whereas the user needs to write down what he/she hears. Spelling errors are anonymously logged into a special database for later analysis and generation of useful feedback. (Volodina et al. 2013; Pijetlovic 2013)

We also plan to extend Lärka with new exercises and a number of new exercise formats, e.g. wordbank, gap cloze and free answer. Spell-checking needs to be added in case of a free-answer format. There have been requests for a grammar exercise format where the whole sentence would be analyzed for POSs, syntactic relations or semantic roles.

Other features that we plan to add include:

- addition of statistics logs to cover exercises aimed at second language learners, in the same fashion as for students of linguistics;
- an option to add user-created word lists for language learners;
- an option to choose among alternative sets of

linguistic terms to cater for users coming from different terminological traditions.

We are also considering to offer the 5 exercises in a simplified mobile app version.

#### 5. References

- Amaral, Luiz and Detmar Meurers (2011) On Using Intelligent Computer-Assisted Language Learning in Real-Life Foreign Language Teaching and Learning. *ReCALL*. Vol 23, No 1. 2011.
- Amaral, Luiz and Detmar Meurers (2007). Putting activity models in the driver's seat: Towards a demand-driven NLP architecture for ICALL. *EUROCALL 2007, Symposium on NLP in CALL*. Sept, 5-8, 2007. University of Ulster, Coleraine Campus, Ireland.
- Borin, Lars, Dana Dannélls, Markus Forsberg, Maria Toporowska Gronostaj and Dimitrios Kokkinakis (2010). The past meets the present in Swedish FrameNet++. *14th EURALEX International Congress*. Leeuwarden: EURALEX. 269–281.
- Borin, Lars, Markus Forsberg, Leif-Jöran Olsson and Jonatan Uppström (2012a). The open lexical infrastructure of Språkbanken. *Proceedings of LREC 2012*. Istanbul: ELRA. 3598–3602.
- Borin, Lars, Markus Forsberg and Johan Roxendal (2012b). Korp – the corpus infrastructure of Språkbanken. *Proceedings of LREC 2012*. Istanbul: ELRA. 474–478.
- Borin, Lars and Anju Saxena (2004). Grammar, incorporated. *CALL for the Nordic languages*, ed. by Peter Juel Henriksen. (Copenhagen Studies in Language 30.) Copenhagen: Samfundslitteratur. 125–145.
- Burstein, Jill, Joel Tetreault and Nitin Madnani (2013). The E-rater Automated Essay Scoring System. In Shermis, Mark D.; Burstein, Jill (eds.) *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. New York: Routledge, 2013, p55-67.
- Cahill, Aiofe, Martin Chodorow, Susanne Wolff and Nitin Madnani (2013). Detecting Missing Hyphens in Learner Text. *Proceedings of the 8th Workshop on Innovative Use of NLP for Building Educational Applications*, NAACL-HLT 2013, Atlanta, USA.
- COE (2012). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*. Council of Europe. <[http://www.coe.int/t/dg4/linguistic/Cadre1\\_en.asp](http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp)>.
- Dickinson, Markus and Joshua Herring (2008). Developing Online ICALL Exercises for Russian. *Proceedings of the 3rd Workshop on Innovative Use of NLP for Building Educational Applications (ACL08-NLP-Education)*. Columbus, OH.
- Einarsson, Jan (1976). *Talbankens skriftspråkskonkordans*. Lund University: Department of Scandinavian Languages.
- Ejerhed, Eva, Gunnel Källgren, Ola Wennstedt and

- Magnus Åström (1992). *The linguistic annotation system of the Stockholm-Umeå Corpus project. Report No 33*. University of Umeå: Department of Linguistics. Department of Scandinavian Languages.
- Gellerstam, Martin (1999). LEXIN - lexikon för invandrare. *LexicoNordica* 6. 3-18.
- Heift, Trude (2010). Developing an intelligent language tutor. *CALICO Journal* 27(3), 443-459.
- Heift, Trude and Mathias Schulze (2007). *Errors and Intelligence in Computer-Assisted Language Learning: Parsers and Pedagogues*, London: Routledge.
- Källgren, Gunnel (2006). Documentation of the Stockholm - Umeå Corpus. *Manual of the Stockholm Umeå Corpus version 2.0*. Sofia Gustafson-Capková and Britt Hartmann (eds). Stockholm University: Department of Linguistics.
- King, Levi and Markus Dickinson (2013). Shallow Semantic Analysis of Interactive Learner Sentences. *Proceedings of the 8th Workshop on Innovative Use of NLP for Building Educational Applications*, NAACL-HLT 2013, Atlanta, USA.
- Li, Baoli (2013). Recognizing English Learners' Native Language from Their Writings. *Proceedings of the 8th Workshop on Innovative Use of NLP for Building Educational Applications*, NAACL-HLT 2013, Atlanta, USA.
- Meurers, Detmar (2012) Natural Language Processing and Language Learning. *Encyclopedia of Applied Linguistics*, edited by Carol A. Chapelle. Blackwell.
- Meurers, Detmar, Ramon Ziai, Luiz Amaral, Adriane Boyd, Aleksandar Dimitrov, Vanessa Metcalf and Niels Ott (2010). Enhancing Authentic Web Pages for Language Learners. *Proceedings of the 5th Workshop on Innovative Use of NLP for Building Educational Applications*, NAACL-HLT 2010, Los Angeles.
- Östling, Robert, André Smolentzov, Björn Tyrefors Hinnerich and Erik Höglin (2013). Automated Essay Scoring for Swedish. *Proceedings of the 8th Workshop on Innovative Use of NLP for Building Educational Applications*, NAACL-HLT 2013, Atlanta, USA.
- Pijetlovic, Dijana (2013). *Swedish spelling game: Developing Swedish spelling exercises on the ICALL platform Lärka using Text-to-Speech*. Master Thesis. University of Gothenburg.
- Pilán, Ildikó (2013). *NLP-based approaches to sentence readability for second language learning purposes*. Master Thesis. University of Gothenburg.
- Pilán, Ildikó and Elena Volodina (2014). Reusing Swedish FrameNet for training semantic roles. LREC 2014.
- Pilán, Ildikó, Elena Volodina and Richard Johansson (2013). Automatic selection of suitable sentences for language learning exercises. *Proceedings of Eurocall 2013*.
- Saxena, Anju and Lars Borin (2002). Locating and reusing sundry NLP flotsam in an e-learning application. *LREC 2002. Workshop Proceedings. Customizing knowledge in NLP applications: strategies, issues, and evaluation*. Las Palmas: ELRA. 45-51.
- Saxena, Anju and Mikaëla Lind (2008). Corpora in Grammar Learning. Evaluation of ITG. *Resourceful Language Technology: Festschrift in Honor of Anna Sågvall Hein*. 149-158. Studia Linguistica Upsaliensia. Acta Universitatis Upsaliensis: Uppsala.
- Shen, Wade, Jennifer Williams, Tamas Marius and Elizabeth Salesky. (2013) A Language-Independent Approach to Automatic Text Difficulty Assessment for Second-Language Learners, *Second Workshop for Predicting and Improving Text Readability for Target Reader Populations (PITR)*, Association for Computational Linguistics (ACL), Sofia, Bulgaria, August 2013.
- Sköldberg, Emma and Sofie Johansson Kokkinakis (2012). A och O om akademiska ord. Om framtagning av en svensk akademisk ordlista. *Nordiska studier i lexikografi 11. Rapport från Konferensen om lexikografi i Norden, Lund, maj 2011*.
- Teleman, Ulf (1974). *Manual för grammatisk beskrivning av talad och skriven svenska*. Lund: Liber.
- Vajjala, Sowmya and Detmar Meurers (2014). Readability Assessment for Text Simplification: From Analyzing Documents to Identifying Sentential Simplifications. to appear in the *International Journal of Applied Linguistics, Special Issue on Current Research in Readability and Text Simplification* edited by Thomas François and Delphine Bernhard.
- Volodina, Elena (2010). *Corpora in Language Classroom: Reusing Stockholm Umeå Corpus in a vocabulary exercise generator*. Saarbrücken: Lambert Academic Publishing.
- Volodina, Elena, Lars Borin, Hrafn Loftsson, Birna Arnbjörnsdóttir and Guðmundur Örn Leifsson (2012). Waste not, want not: Towards a system architecture for ICALL based on NLP component re-use. Workshop on NLP in Computer-Assisted Language Learning. *Proceedings of the SLTC 2012 workshop on NLP for CALL. Linköping Electronic Conference Proceedings 80*. 47-58.
- Volodina, Elena and Sofie Johansson Kokkinakis (2012). Introducing Swedish Kelly-list, a new lexical e-resource for Swedish. *Proceedings of LREC 2012*, Istanbul: ELRA.
- Volodina, Elena, Dijana Pijetlovic, Ildikó Pilán and Sofie Johansson Kokkinakis (2013). Towards a gold standard for Swedish CEFR-based ICALL. *Proceedings of the Second Workshop on NLP for Computer-Assisted Language Learning. Nodalida 2013*, Oslo, Norway.