

Using LARA for language learning: a pilot study for Icelandic

Branislav Bédi¹, Cathy Chua², Hanieh Habibi³,
Ruth Martinez-Lopez⁴, and Manny Rayner⁵

Abstract. This paper presents a brief overview of LARA (Learning And Reading Assistant), an open source online tool that has been under development since summer 2018. LARA currently⁶ contains a corpus of about 25 texts in ten languages and a crowdsourcing model used to expand the corpus. The central goal is to provide support for improving second language (L2) reading comprehension. The focus here is on the development of Icelandic content and its use during pilot testing amongst adult L2 learners of Icelandic. Preliminary feedback from users, while mostly positive, contained suggestions on how the tool might be improved.

Keywords: crowdsourcing, L2 learning, open source, pilot testing, reading.

1. Introduction

LARA is a free online tool which makes it easy to transform plain texts into a hyperlinked multimedia form designed to support non-native readers: it thus connects content providers, such as teachers, with readers/learners. When accessing LARA content, the user sees a split-screen with the text on one side and various options on the other. The unique feature of the interface is the personalised concordance created for each reader-user. That is, based on the reader's own history, words in the target language will be displayed in the contexts in which they previously appeared. Other functionalities include dictionary translations,

1. The Árni Magnússon Institute for Icelandic Studies, Reykjavík, Iceland; branislav.bedi@arnastofnun.is

2. Independent scholar, Geneva, Switzerland; cathyc@pioneerbooks.com.au

3. FTI/TIM University of Geneva, Geneva, Switzerland; hanieh.habibi@unige.ch

4. Seville University, Seville, Spain; ruthmart@gmail.com

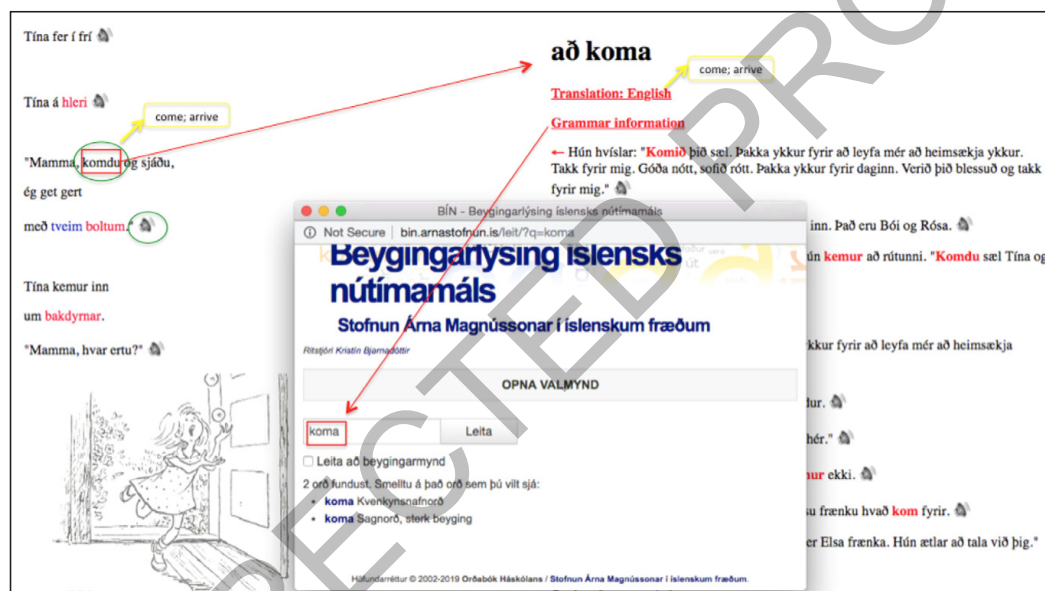
5. University of Geneva, Geneva, Switzerland; emmanuel.rayner@unige.ch

6. Mid-July 2019

How to cite this article: Bédi, B., Chua, C., Habibi, H., Martinez-Lopez, R., & Rayner, M. (2019). Using LARA for language learning: a pilot study for Icelandic. In F. Meunier, J. Van de Vyver, L. Bradley & S. Thouéšny (Eds), *CALL and complexity – short papers from EUROCALL 2019* (pp. 1-6). Research-publishing.net. <https://doi.org/10.14705/xxx-to-be-confirmed>

voice recordings of words and passages, links to online grammar resources, and vocabulary lists ordered both alphabetically and by frequency of occurrence in the text. All of these can be accessed by pointing and clicking. Figure 1 illustrates our initial Icelandic resource, *Tína fer í fri* (Skriver, 1981), a children's book containing about 2,700 words. LARA's content repository currently contains 25 texts in ten languages, for a total of about 140,000 words. A detailed overview of LARA can be found in Akhlaghi et al. (2019).

Figure 1. LARA functionality for Icelandic text in a split-screen view. Featured are: translation tags, grammar information, audio icons, contextual word reference in the corpus



LARA's pedagogical model is built on Palincsar and Brown's (1984) interactive teaching to promote independent learning from reading comprehension, Oxford's (1990) reading strategy, and on the early example of Johns's (2002) Data-Driven Learning (DDL). Based on these sources, we consider L2 learning by reading comprehension can be facilitated by: (1) providing learners with texts that are suitable to their level, (2) making the content compatible with the reader's knowledge, and (3) supporting active strategies which readers may employ to enhance understanding and retention, and to facilitate comprehension (Palincsar & Brown, 1984, p. 118). DDL has from the start built on the affordances of text retrieval software in L2 education. In particular, it provides a framework for supporting scaffolding and learner autonomy (Boulton & Tyne, 2013; Corino & Onesti, 2019).

The use of corpora in L2 education is still relatively rare. Some of the best-known online corpus analysis tools with open access are the NoSketchEngine (Rychlý, 2007), SkELL (Baisa & Suchomel, 2014), AntConc (Anthony, 2019), and some others that are included in LexTutor (Cobb, 2019). The commonality of these tools is that they contain crowdsourced corpora consisting of online texts in various languages, provide information about the frequency of words, and support language learning by giving access to numerous occurrences of words, phrases, idioms, and other language expressions in context.

Organisationally, LARA is a part of the open source CALlector project⁷, which in turn is closely linked to the enetCollect⁸ COST network. Ethical issues are central to both LARA and CALlector. The users, be they teachers/content developers or learners/content users, are in control of their data, and the platform/tools are designed to be self-sufficient through best practices of crowdsourced free models. In terms of software engineering, the key medium-term goals are to make the content development process simple and user-friendly, with only basic computer skills required, and embed the platform in a dedicated social network which connects content developers and content users. Although initial LARA content is already being trialled by classroom teachers in Iceland and elsewhere, these goals have still only been partially achieved. In particular, the social network level is not yet available; the tools still need to be downloaded and installed on a local machine; and for some languages, currently including Icelandic, the process of annotating the text still requires laborious manual tagging of surface words by head-word (while automatic tagging is supported for many other languages). Work on all these issues is progressing. A first version of a LARA web portal is currently being tested, which will obviate the need to install software. Taggers for other languages, in particular, an open source Icelandic tagger, are being integrated, and we expect the social network to go live towards the end of 2019.

2. Pilot testing of LARA for learning Icelandic

The pilot test was organised around an anonymous questionnaire in March 2019. There were 47 voluntary participants consisting of learners enrolled in the Icelandic Practical Diploma Course at the University of Iceland: 21 female, 23 male, one ‘other’, two ‘prefer not to say’; and 39 beginners, eight intermediate, coming from 17 countries. After a short introduction to the tool, they spent 60 minutes using it by reading the Icelandic text on their laptops while using headphones for listening.

7. <https://www.unige.ch/callector/>

8. <http://enetcollect.eurac.edu>

The instructor only assisted with technical issues. The learners were requested to use the tool on their own to achieve optimal first-time user experience results. Then they were given 20 minutes to respond to the questionnaire which included 31 questions about learners' background and perceived usefulness with Likert-scale answers (Davis, 1989), and open-ended questions about the assessment of digital tools (Nesbitt, 2013). The answers about the users' general experience and learning effect are summarised in Table 1. Most of the learners either agreed or strongly agreed with positive statements about LARA. Compared with the traditional way of reading a book, most users said that the tool's functionalities increased their efficiency in reading with understanding, learning vocabulary, and (to some extent) learning grammar. Nearly all students said that the integrated audio helped to learn more about pronunciation.

Results from open-ended questions suggest that over 90% of the learners liked the application. Problems experienced included: poor internet connection, missing or inaccurate word translations, and lack of support for mobile devices. Currently Chrome is the only browser permitting full functionality. About 70% suggested improvements, e.g. including instructions before the task, on-screen instructions to help navigate, options for both female and male voices, adding a bookmark, and options for changing font/background colour. About the same proportion of learners liked the design. About 81% of the learners said that the application met their needs; 11% suggested improvements like adding more explanations about grammar, sentence structure, vocabulary, and phrases. Eight percent said it did not meet their needs.

Table 1. Learners' perceptions regarding usefulness

No.	Question	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1.	The reading task would be difficult to perform without this application.	19%	49%	19%	11%	2%
2.	Using this application gives me greater control over my reading task.	38%	59%	9%	4%	-
3.	Using this application improves my understanding of the Icelandic text.	49%	47%	4%	-	-
4.	The application addresses my learning-related needs in this course.	40%	55%	2%	2%	-
5.	Using this application saves me time.	34%	55%	11%	-	-
6.	This application enables me to accomplish reading tasks more quickly.	34%	51%	15%	-	-

7.	This application supports critical aspects of my learning.	23%	53%	17%	6%	-
8.	Using this application allows me to accomplish more reading tasks than would otherwise be possible.	36%	51%	11%	2%	-
9.	Using this application reduces the time I spend on unproductive activities.	26%	38%	26%	11%	-
10.	Using this application enhances my effectiveness in reading.	32%	60%	6%	2%	-
11.	Compared to using books, using this application improves the quality of reading Icelandic texts.	45%	40%	11%	2%	2%
12.	Using this application increases my learning productivity.	28%	53%	19%	-	-
13.	Using this application makes it easier for me to read Icelandic texts.	38%	55%	4%	2%	-
14.	Using this application makes it easier for me to learn vocabulary.	40%	47%	13%	-	-
15.	Using this application makes it easier for me to learn grammar.	28%	38%	26%	9%	-
16.	Using this application makes it easier for me to learn pronunciation.	49%	43%	4%	4%	-
17.	Overall, I find this application useful in my Icelandic course.	55%	38%	4%	-	2%

3. Conclusions and future work

Based on the above results, the tool appeals to learners: it assists them in the development of reading and associated skills including vocabulary, grammar, listening comprehension, and pronunciation. Learners find the current functionalities, i.e. dictionary translations, voice recordings of words and passages, links to online grammar resources, and vocabulary lists ordered both alphabetically and by frequency of occurrence in the text, practical and useful.

In the future, the crowdsourcing approach will enable content providers to add more texts in Icelandic, as is the case for other languages. The personalised concordance facility that allows learners to answer questions about vocabulary by reading concordance lines (Hunston, 2002) will become part and parcel of the learning process. The pilot test also included several suggestions from learners on how to improve the tool. Some important improvements have already been implemented and others are in the pipeline.

4. Acknowledgements

This article is based upon work from COST Action enetCollect (CA16105) supported by COST (European Cooperation in Science and Technology). Work at the University of Geneva was funded by the Swiss National Science Foundation under grant IZCOZ0_177065.

References

- Akhlaghi, E., Bédi, B., Butterweck, M., Chua, C., Gerlach, J., Habibi, H., Ikeda, J., Rayner, M., Sestigiani, S., & Zuckermann, G. (2019). Overview of LARA: a learning and reading assistant. In *Proceedings SLATE 2019*, Graz, Austria. <https://doi.org/10.21437/slate.2019-19>
- Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Waseda University. <https://www.laurenceanthony.net/software>
- Baisa, V., & Suchomel, V. (2014). SkELL – web interface for English language learning. In *Eighth Workshop on Recent Advances in Slavonic Natural Language Processing* (pp. 63-70). Tribun EU. <https://nlp.fi.muni.cz/raslan/2014/12.pdf>
- Boulton, A., & Tyne, H. (2013). Corpus linguistics and data-driven learning: a critical overview. In *Bulletin VALS-ASLA 97*, 97-118. <https://core.ac.uk/download/pdf/20661944.pdf>
- Cobb, T. (2019). *Complete leixical tutor v8.3* [computer program]. <https://www.lex tutor.ca>
- Corino, E., & Onesti, C. (2019). Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning. In *Frontiers of Education*, 4(7), 1-12. <https://doi.org/10.3389/feduc.2019.00007>
- Davis, F. D. (1989). Percieved usefulness, perceived ease of use, and user acceptance of information technology. In *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Hunston, S. (2002). *Corpora in applied linguistics*. Oxford University Press.
- Johns, T. (2002). Data-driven learning: the perpetual challenge. In *Teaching and Learning by Doing Corpus Analysis* (pp. 105-117). Brill Rodopi. https://doi.org/10.1163/9789004334236_010
- Nesbitt, D. (2013). Student evaluation of CALL tools during the design process. *Computer Assisted Language Learning*, 26(4), 371-387. <https://doi.org/10.1080/09588221.2012.680471>
- Oxford, R. L. (1990). *Language learning strategies: what every teacher should know*. Heinle & Heinle.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175. https://doi.org/10.1207/s1532690xci0102_1
- Rychlý, P. (2007). Manatee/Bonito – a modular corpus manager. In *1st Workshop on Recent Advances in Slavonic Natural Language Processing* (pp. 65-70). Masaryk University. <https://nlp.fi.muni.cz/trac/noske>
- Skriver, E. (1981). *Tína fer í frí. Námsgagnastofnun*. Prent húsið.