# Automatic exercise generation to support macro-adaptivity in intelligent language tutoring systems

Tanja Heck[1], Detmar Meurers[2], and Florian Nuxoll[3]

**Abstract**. Foreign language teaching achieves best learning outcomes when individual differences of learners are taken into account. While it is difficult for teachers to support internal differentiation in the classroom, digital tools can adaptively propose individual learning paths through activities so that students can practice with appropriately challenging exercises. But how can sufficiently varied, systematically parametrized exercises be provided to enable a system to match them to individual learner needs? We present an approach for high-variability exercise generation that transforms a single specification into a multitude of exercises varying in complexity. The approach is currently being evaluated in a randomized controlled study in regular German seventh grade English classes, facilitating a systematic empirical exploration of adaptive exercise complexity in relation to learning outcomes.

**Keywords**: intelligent language tutoring systems, ICALL, exercise generation, macro-adaptivity.

## 1. Introduction

Language instruction is challenged by the need to accommodate heterogeneous groups of students who differ in optimal learning paths and required support to achieve the best possible learning outcomes. Individual differences play a major role

1. University of Tübingen, Tübingen, Germany; tanja.heck@uni-tuebingen.de; https://orcid.org/0000-0002-5091-1019
2. University of Tübingen, Tübingen, Germany; detmar.meurers@uni-tuebingen.de; http://purl.org/dm, https://orcid.org/0000-0002-9740-7442
3. University of Tübingen, Tübingen, Germany; florian.nuxoll@gmx.net

**How to cite this article**: Heck, T., Meurers, D., & Nuxoll, F. (2022). Automatic exercise generation to support macro-adaptivity in intelligent language tutoring systems. In B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouësny, & M. J. Whelpton (Eds), *Intelligent CALL, granular systems, and learner data: short papers from EUROCALL 2022* (pp. 162-167). Research-publishing.net. https://doi.org/10.14705/rpnet.2022.61.1452

in second language learning with students differing in a number of characteristics, including motivation, background, general language aptitude, and cognitive resources (Dörnyei, 2005). Yet instructors struggle with internal differentiation in classroom teaching as they generally lack time, suitable diagnostics, and appropriately adaptive materials. To speed up learning and improve learning outcomes, Intelligent Language Tutoring Systems (ILTS) aim to overcome these challenges by matching exercise properties to learner characteristics in an attempt to identify the best practice material at any time, for any student. Considering the vast space of exercise characteristics and combinations thereof, the availability of systematically parametrized learning material constitutes a major bottleneck for learner-adaptive digital systems.

A number of ILTS have been developed for foreign language learning. Examples successfully used in real life include Heift's (2010) E-Tutor, Nagata's (2009) Robo-Sensei, and Amaral and Meurers's (2011) Tagarela. Such systems can in principle support two types of adaptivity, micro- and macro-adaptivity. Most systems focus on providing scaffolding feedback that successively guides a learner toward the correct answer, i.e. micro-adaptivity. Adaptive exercise sequencing attempting to always select activities in the learner's zone of proximal development has received less attention. The challenge with such macro-adaptivity consists in a shortage of different activities which, if overcome, would allow adaptive systems to provide exercises whose nature and complexity matches a learner's characteristics and proficiency. Manually creating all these exercises quickly becomes unfeasible considering the number of possible combinations of exercise parameter settings. A systematic approach, however, could generate a multitude of activities from a single, well-defined specification. To this purpose, we present an approach for high-variability exercise generation with the goal of efficiently generating practice material for user-adaptive ILTS targeting beginning to intermediate learners.

## 2. System description

The approach focuses on exercises practicing grammatical concepts from the official English curriculum in Germany. The language-aware search engine FLAIR (Chinkina & Meurers, 2016) assists users in identifying texts rich in pedagogically relevant linguistic structures. Although, at this point, the high-variability exercise generation requires manually compiled specifications, FLAIR thus provides a natural system context for an exercise generation functionality. To use the new exercise generation in FLAIR, users upload a specification file

in Excel or JSON format and select the targeted grammatical concept and the output file format. The algorithm generates abstract exercise definitions from the specification for the entire space of pre-defined parameterizations. Parameters influencing exercise complexity include: (1) the exercise type, including memory, multiple choice, underline, jumbled sentences, categorization, and fill-in-the-blanks; (2) parameters specific to exercise types, such as the number and type of distractors for multiple choice activities or the nature and position of hints for fill-in-the-blanks exercises, e.g. lemmas above the exercise or behind the gaps; and (3) parameters specific to a grammatical concept, e.g. which clause is turned into the relative clause of a relative sentence, clause order in conditional sentences, or whether the exercise focuses on practicing verb forms in if-clauses or in main clauses.

Figure 1 takes relative clauses as an example and outlines how different exercise types make use of different specification components for exercise generation. Memory activities, for example, use the information highlighted in bold font: the chunks of both clauses and the relative pronoun.

Figure 1. Parameter usage of different relative clause exercises

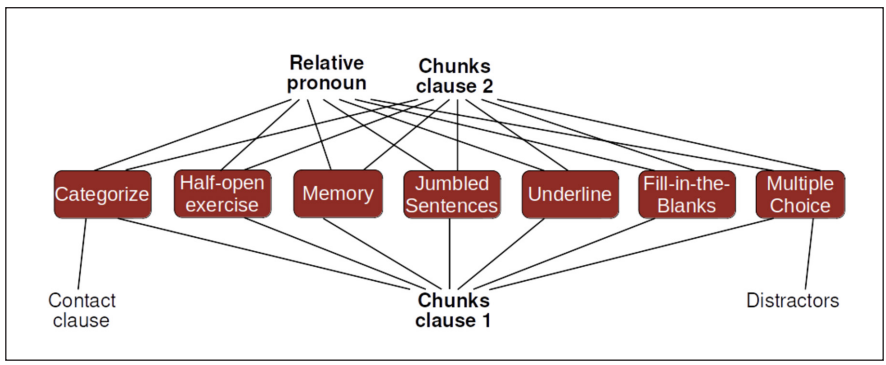
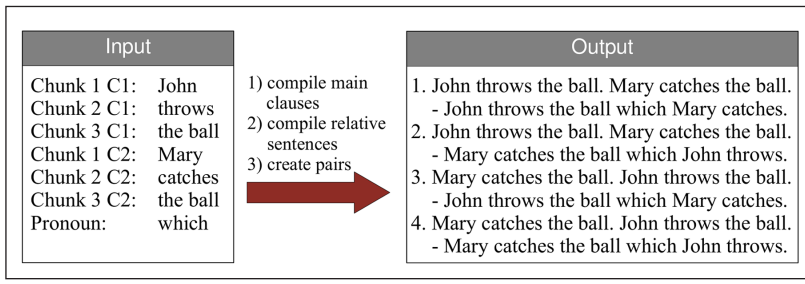



Figure 2 below illustrates how this subset of specification components is used to generate four distinct memory exercises for relative clauses. The parameters responsible for this variability consist in the order of the main clauses as well as in the choice of clause used for the relative clause in the relative sentence.

Depending on the selected output format, the algorithm transforms the generated abstract exercise definition into H5P files, which can be uploaded to any learner management system supporting the standard, or a proprietary XML format compatible with the FeedBook ILTS (Rudzewitz et al., 2018).

Figure 2. Generation of Relative clause Memory exercises from a specification




## 3. Evaluation and discussion

The approach targets grammatical concepts from the seventh grade English curriculum in Germany. The current implementation covers relative clauses with focus on pronoun usage or contact clauses, and conditionals practicing verb forms of a given conditional type or the differentiation between conditional types. The specifications for different foci on relative clauses differ slightly due to differing requirements of the resulting exercises. The specifications for conditionals can in principle be re-used if the grouping of specification items is altered in such a way that all items of a group belong to the same or to different conditional types. For relative pronoun practice, the algorithm supports automatic generation of 19 exercises from a single specification. Contact clauses yield three different exercises. For verbs of conditional sentences, the algorithm generates 48 exercises per specification. Exercises practicing the identification of conditional types come in 39 variants. This promising space of exercises is currently being empirically tested in a project related to the FeedBook focused on adaptive sequencing. In line with the setup advocated in Meurers et al. (2019), we are currently conducting a year-long randomized controlled field study in regular English classes in secondary schools.

Limitations of the current approach include the handling of discourse phenomena such as coreference and cohesion. Example 1 illustrates that for conditional sentences, the semantic validity is challenged if the first clause contains pronouns referring to an antecedent in the second clause. Also related to clause order, Example 2 illustrates how cohesion can become an issue with de-contextualized relative clauses if the clauses have a natural order which is violated by the generated relative sentence. An additional limitation is introduced by the naive nature of the exercise generation algorithm which cannot resolve co-references unless the surface forms are identical, so that Example 3a cannot be transformed into a relative sentence but instead requires a specification such as given in Example 3a'.

(1) a. If Tommy found his glasses, he would wear them.
b. # He would wear them if Tommy found his glasses.
b'. Tommy would wear his glasses if he found them.

(2) a. The boy who came in greeted everyone.
b. # The boy who greeted everyone came in.

(3) a. The boy came in. He greeted everyone.
a'. The boy came in. The boy greeted everyone.

We are therefore currently introducing further natural language processing analyses to overcome these limitations and to make it possible to add parameters that are characteristic of developmental sequences discussed in the second language acquisition literature (e.g. Pienemann & Johnston, 1986), such as the use of negation or interrogative structures. In addition, we are implementing a user interface to facilitate specification authoring and to avoid formatting issues of the input files. This interface will also merge different specifications of a grammatical concept and thus allow even greater exercise variability per specification. Since the exercise generation is already embedded into a language-aware search engine, another strand of current work focuses on the generation of exercises from authentic texts.

## 4. Conclusions

We have presented an approach to support macro-adaptivity in ILTS through high-variability exercise generation. In future work, we plan to extend the approach to a broader range of grammatical concepts. We also intend to evaluate the influence of different parameters and their interaction on exercise difficulty, approximated by student performance. To this purpose, we hope to gain first insights from the ongoing field study which we will then broaden through analyses based on controlled lab studies on exercises with systematically manipulated parameters.

## References

Amaral, L. A., & Meurers, D. (2011). On using intelligent computer-assisted language learning in real-life foreign language teaching and learning. *ReCALL*, *23*(1), 4-24. https://doi.org/10.1017/S0958344010000261

Chinkina, M., & Meurers, D. (2016). Linguistically aware information retrieval: providing input enrichment for second language learners. *BEA 2016* (pp. 188-198). https://doi.org/10.18653/v1/W16-0521

Dörnyei, Z. (2005). The psychology of the language learner: individual differences in second language acquisition (1st ed.). *Routledge*. https://doi.org/10.4324/9781410613349

Heift, T. (2010). Developing an intelligent language tutor. *CALICO journal*, *27*(3), 443-459. https://doi.org/10.11139/cj.27.3.443-459

Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics*, *39*, 161-188. https://doi.org/10.1017/S0267190519000126

Nagata, N. (2009). Robo-Sensei's NLP-based error detection and feedback generation. *Calico Journal*, *26*(3), 562-579. https://doi.org/10.1558/cj.v26i3.562-579

Pienemann, M., & Johnston, M. (1986). An acquisition-based procedure for second language assessment (ESL). *ARAL*, *9*(1), 92-122. https://doi.org/10.1017/S0272263100007324

Rudzewitz, B., Ziai, R., De Kuthy, K., Möller, V., Nuxoll, F., & Meurers, D. (2018). Generating feedback for English foreign language exercises. *BEA 2018* (pp. 127-136). https://doi.org/10.18653/v1/W18-0513

Published by Research-publishing.net, a not-for-profit association
Contact: info@research-publishing.net



**Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022**
**Edited by Birna Arnbjörnsdóttir, Branislav Bédi, Linda Bradley, Kolbrún Friðriksdóttir, Hólmfríður Garðarsdóttir, Sylvie Thouësny, and Matthew James Whelpton**

**Publication date: 2022/12/12**



Typeset by Research-publishing.net
Cover photo by © 2022 Kristinn Ingvarsson (photo is taken inside Veröld – House of Vigdís)
Cover layout by © 2022 Raphaël Savina (raphael@savina.net)

ISBN13: 978-2-38372-015-7 (PDF, colour)

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

**Legal deposit, France**: Bibliothèque Nationale de France - Dépôt légal: décembre 2022.