A writing-support system utilizing 'Leichte Sprache' (LS; easy-to-read German) for people with low literacy skills
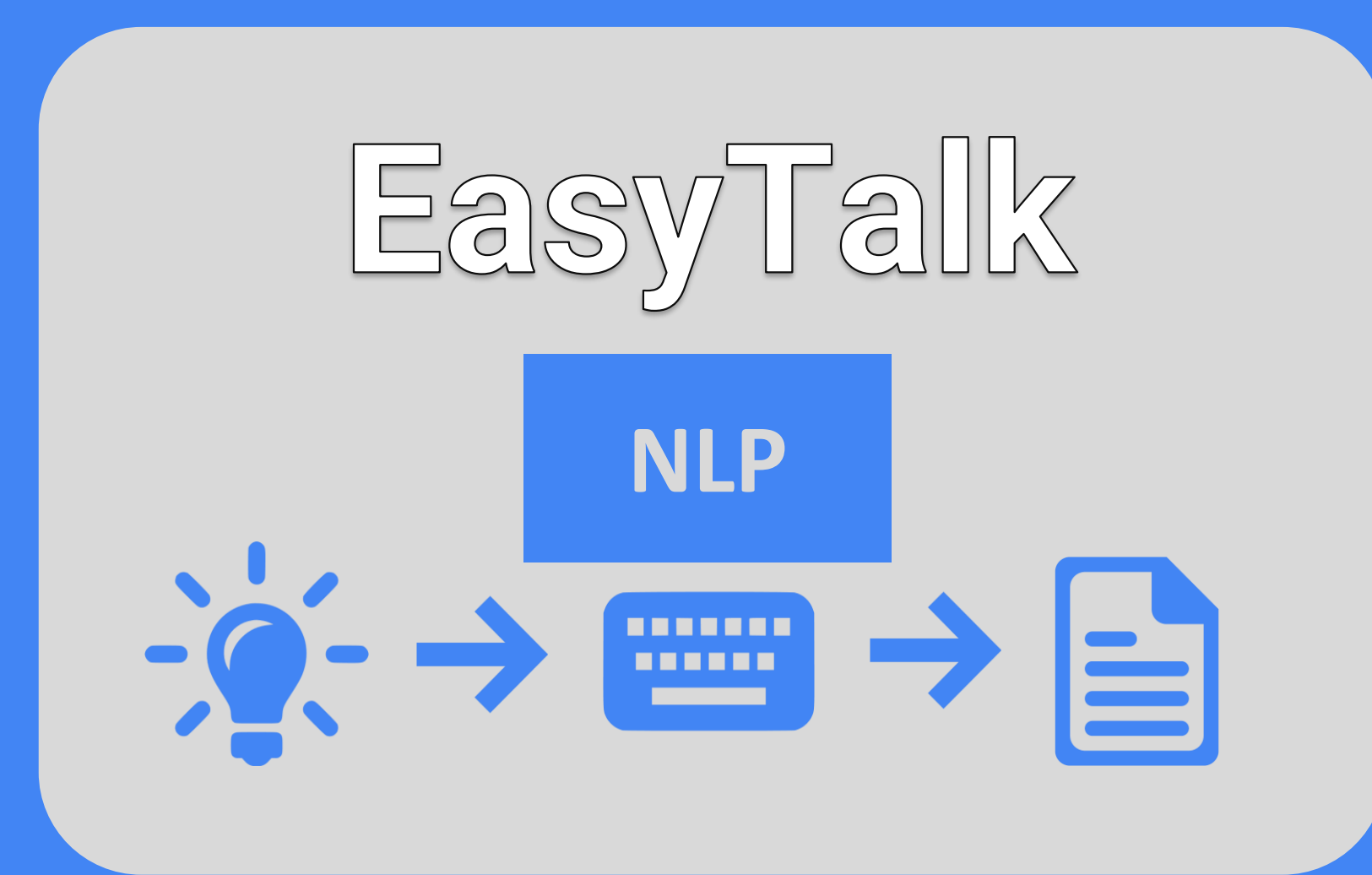
EasyTalk

NLP

UNIVERSITÄT KOBLENZ · LANDAU

Ina Steinmetz and Karin Harbusch
{inaschroeder | harbusch}@uni-koblenz.de,
Institute for Computational Visualistics, Computer Science Faculty

Target group

People with low literacy skills are excluded from the free exchange of written information in social and economic life. This also holds for many people with intellectual disabilities, learning difficulties or complex communication needs.

Prerequisites

- **Leichte Sprache (LS)** is a variety of German characterized by a small vocabulary and simplified constructions. It aims at inclusion by overcoming language barriers.
- LS rules[1,2,3] include e.g.: use only main clauses in active mode, indicative mood, present or present perfect tense; word order in declarative clauses is subject-verb-object (SVO)
- COMPASS is a natural language processing (NLG) paraphrase generator based on a lexicalized, unification-based Performance Grammar[4,5] using German CELEX[6].
- Symbols from Augmentative and Alternative Communication (AAC) remedy reading deficits.

Research objective

- To what extent can **natural language processing** (NLP) **support users with low literacy skills to produce correct and coherent text** at their personal level of proficiency in spelling and clause construction?
- Which **grammatical support** is needed?
- What does it mean to **make the interface intuitive and easy-to-use**?
- EasyTalk[7] supports readers of LS to produce their own text.

Writing support in EasyTalk

EasyTalks helps users to practice text production at the personal skill level by writing sentences of varying complexity and prompting the user to add important information for the reader[8] and create text coherence.

On the wordform level

EasyTalk offers a personalized vocabulary. Each word form can be supplemented with a customizable symbol from the users' preferred AAC symbol set (ARASAAC symbols[9] are preset). All words and commands in *EasyTalk* can be read aloud on demand.
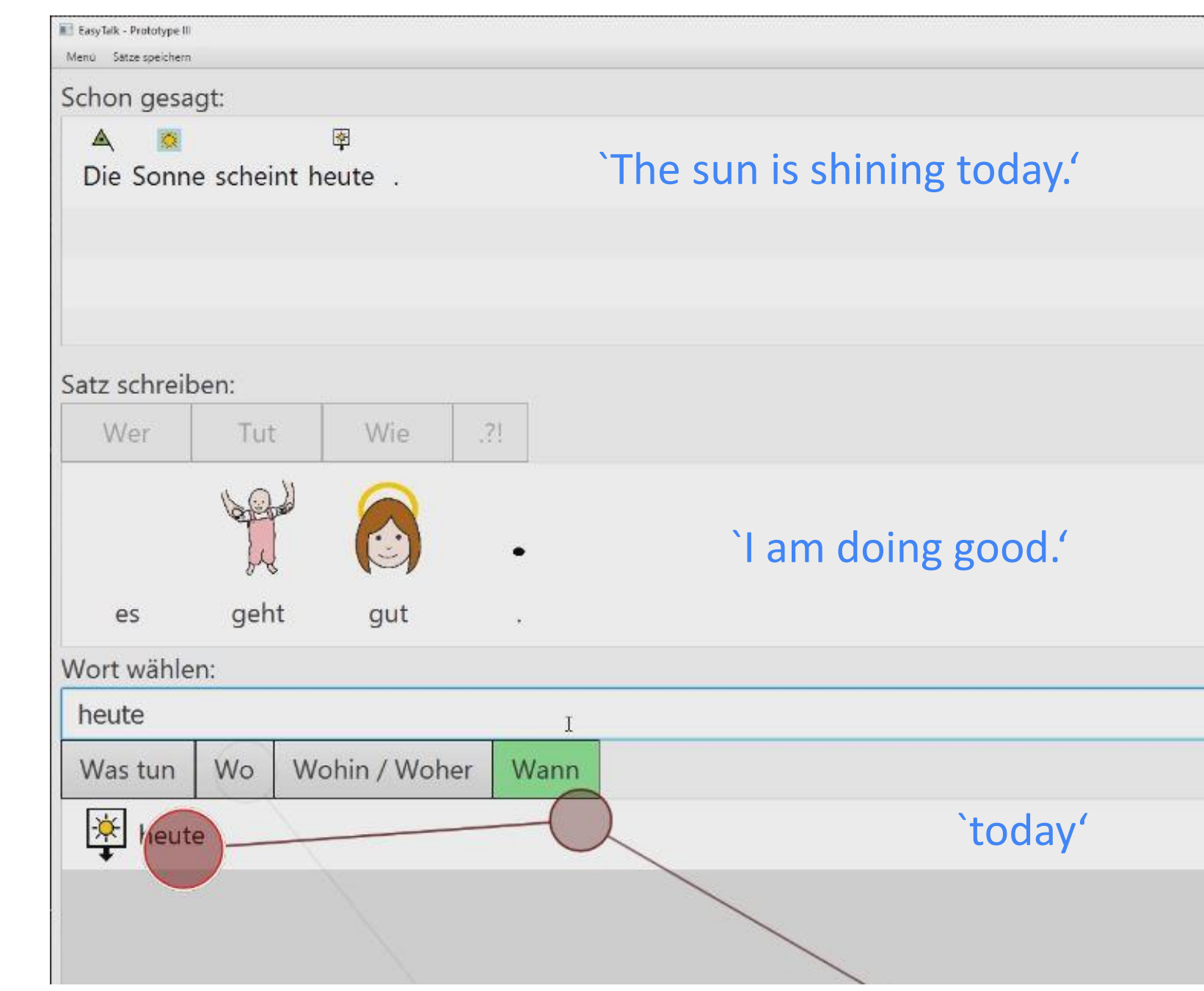
Within a sentence

EasyTalk maintains the syntactic correctness by building up a derivation tree in so called *scaffolded writing*[10] in a dialog with the user based on wh-cues (Who, When,…) to provide a context-sensitively filtered list of wordforms.

Between the sentences

EasyTalk prompts the user to add explicit Rhetorical Relation Theory[11] (RST)-inspired coherence specifications to connect sentences – thus, expressing the intended communicative goal of the sentence.

System evaluation

In a series of Case studies conducted with different target groups (L2 learners with different literacy and computer skills respectively, AAC experts, LS experts, and individuals with intellectual disabilities or autism spectrum disorder) we received positive feedback. For instance, first evaluations of eye-tracking and gaze-focus recordings rom recent text-writing sessions indicate that the wh-cues provide supportive orientation and help to focus.

Support while writing a sentence

Panel 1 Previously typed text

Panel 2 Switches between (a) write sentence & (b) connect sentences

Panel 3 Inflected suggestions for current cue or connector

Support for producing coherent text

Row 1: `And´ | `Or´ | `But´ | `Colon´
Row 2: `Because´ | `Therefore´ | `If´ | `Choose other´

The five essential steps of the text-production process in EasyTalk:

1. In empty Panel 2 (a), clue ``.?!" is provided (click button to change; declarative is the default);
2. Add next word to Panel 2(a) by selecting word by word from the suggestion list offering inflected forms;
3. Choose ✓ to finish the sentence (X deletes last word) → Step 3 in parallel to switching Panel 2 to (b) (cf. step 4);
4. The completed sentence moves to Panel 1 (it can be read out loud and/or exported for further use);
5. Select connection of the next sentence → The system switches back to Panel 2(a) for the next sentence.

System demonstration:
http://inasteinmetz.de/easytalk/

Screenshot from a text writing session in EasyTalk. The red circles indicate the user's recorded gaze.

Future Work

- **Long-term testing** with readers of LS with low literacy skills to evaluate whether EasyTalk can help individuals to improve their literacy skills.
- **Adding supportive typing beyond the scope of pure LS** rules commonly used in LS texts, like simple subordinate clauses and simple past tense for modal verbs.
- **Development of an app version** optimized for handheld mobile devices like smartphones and tablets.

References

1. Barrierefreie Informationstechnik-Verordnung – BITV 2.0. (2011). Accessible at: https://www.gesetze-im-internet.de/bitv_2_0/BJNR184300011.html
2. Netzwerk Leichte Sprache e.V. (2013). Die Regeln für Leichte Sprache. Netzwerk Leichte Sprache e.V., Berlin, Germany.
3. Inclusion Europe. (2014). Informationen für alle – Europäische Regeln, wie man Informationen leicht lesbar und leicht verständlich macht. Inclusion Europe, Brussels, Belgium.
4. Harbusch, K. & Kempen, G. (2002). A quantitative model of word order and movement in English, Dutch and German complement constructions. In Proceedings of the 19th international conference on Computational Linguistics – Volume 1 (COLING '02), Association for Computational Linguistics, USA, 1-7. DOI: https://doig.org/10.3115/1072228.1072262
5. Kempen, G. & Harbusch, K. (2002). Performance Grammar: A declarative definition. In Computational Linguistics in the Netherlands 2001. Brill, Leiden, Niederlande. 148–162. DOI: https://doi.org/10.1163/9789004334038_013
6. Gulikers, L., Rattnik, G., & Piepenbrock, R. (1995). German Linguistic Guide of the CELEX lexical database. Tech. rep., Linguistic Data Consortium, Philadelphia, MA, USA.
7. Steinmetz, I. & Harbusch, K. (2020). Enabling fast and correct typing in 'Leichte Sprache' (Easy Language). The Fourth Widening Natural Language Processing Workshop. Seattle, CA, USA. 64– 67. DOI: http://dx.doi.org/10.18653/v1/2020.winlp-1.17
8. Bell, A. (1984). Language style as audience design. Language in society, 13(2), 145-204. DOI: https://doi.org/10.1017/S004740450001037X
9. ARASAAC Symbols: www.arasaac.org
10. Harbusch, K, & Kempen, G. (2011). Automatic online writing support for L2 learners of German through output monitoring by a natural-language paraphrase generator. Levy, M., Blin, F., Bradin Siskin, C., & Takeuchi, O. (Eds.), WorldCALL - International Perspectives on Computer-Assisted Language Learning. Routledge/Taylor&Francis Group, New York, NY, USA. 128-143.
11. Mann, W. C. & Thompson, S. A. (1988). Rhetorical Structure Theory: Toward a theory of text organization. Text — Interdisciplinary Journal for the Study of Discourse. 8(3). 243–281. DOI: https://doi.org/10.1515/text.1.1988.8.3.243