

A CAPT tool for training and research on lexical stress errors in German

Anjana Vakil
Computational Linguistics
Saarland University
anjanav@coli.uni-saarland.de



Overview

This poster presents the prototype Computer-Assisted Pronunciation Training (CAPT) tool **de-stress**: the German (**de**) System for **T**raining and **R**esearch on **E**rrors in **S**econd-language **S**tress [1].

de-stress targets lexical stress errors by non-native (L2) German speakers with French as their native language (L1). Its modular design incorporates various methods for diagnosing and presenting feedback on these errors, as described below.

Both instructional and research applications have motivated the development of **de-stress**:

- Learners can receive feedback without human instructor
- Teachers can create exercises matching individual student needs
- Researchers can study efficacy of various diagnosis/feedback types

Once more is known about which diagnosis/feedback types are most effective in which situations, this tool could become a useful component of an intelligent CAPT system (see fig. 1).

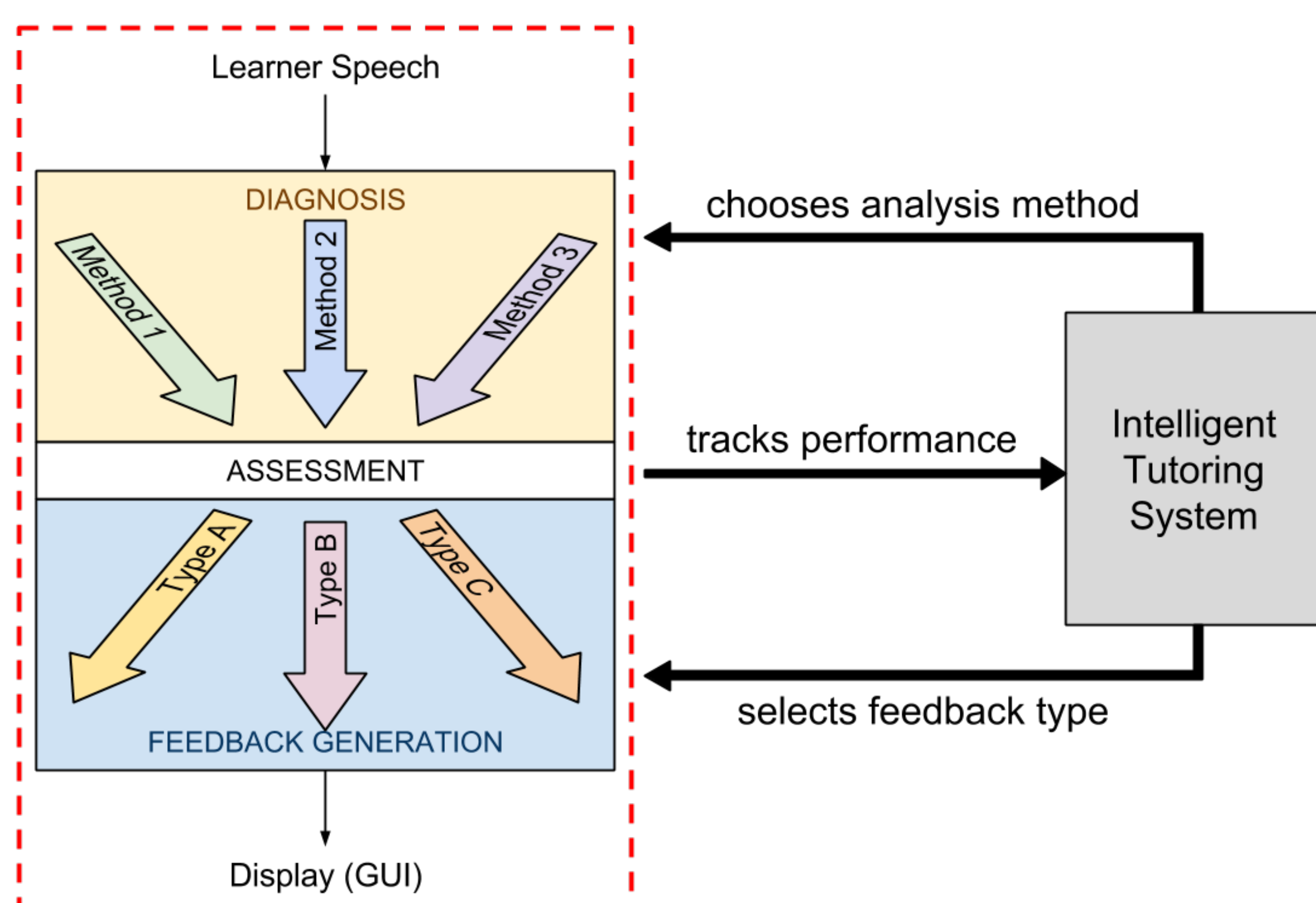


Figure 1: Conceptual diagram of **de-stress** (within dashed line) and its possible function in the context of an Intelligent Tutoring System (ITS).

Error diagnosis

A simple web interface presents a learner with a German sentence to read aloud, with one word highlighted as the target for that exercise. The learner submits an utterance of the sentence for assessment. The learner's realization of the target word's lexical stress pattern is diagnosed via one of the following options:

- **Classification** using machine learning [3]. Possible feature sets:
 - Syllable-level prosodic features (extracted with JSnoori [2]):
 - Duration
 - Fundamental frequency (F0)
 - Intensity
 - Word uttered
 - Speaker age/gender/proficiency
- **Comparison** to reference (native-speaker) utterance(s). Options:
 - One-to-one learner-to-reference comparison using JSnoori [2]
 - One-to-many comparison (averaging one-to-one results)
 - Manual reference selection by either instructor or student
 - Automatic reference selection based on F0 mean and range

Feedback delivery

Based on the error diagnosis, one or more of the following types of feedback are presented to the learner via the web interface (see fig. 2).

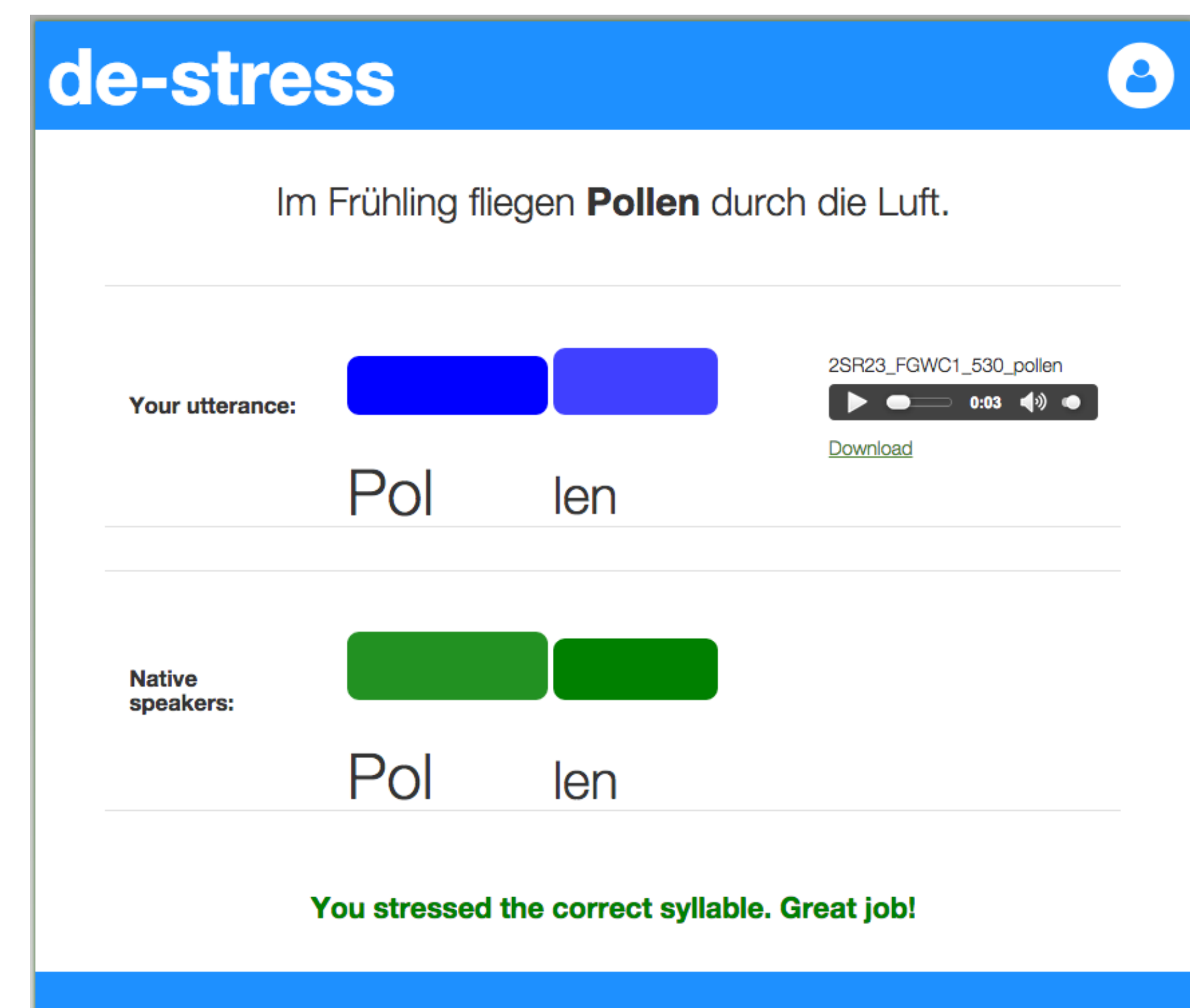


Figure 2: Feedback via graphical visualization (blue/green rectangles), text stylization (syllable text below rectangles), learner utterance playback (audio controls), and verbal message (green text).

- **Explicit feedback:**
 - Verbal error/success messages (see fig. 2)
 - Graphical "skill bars"

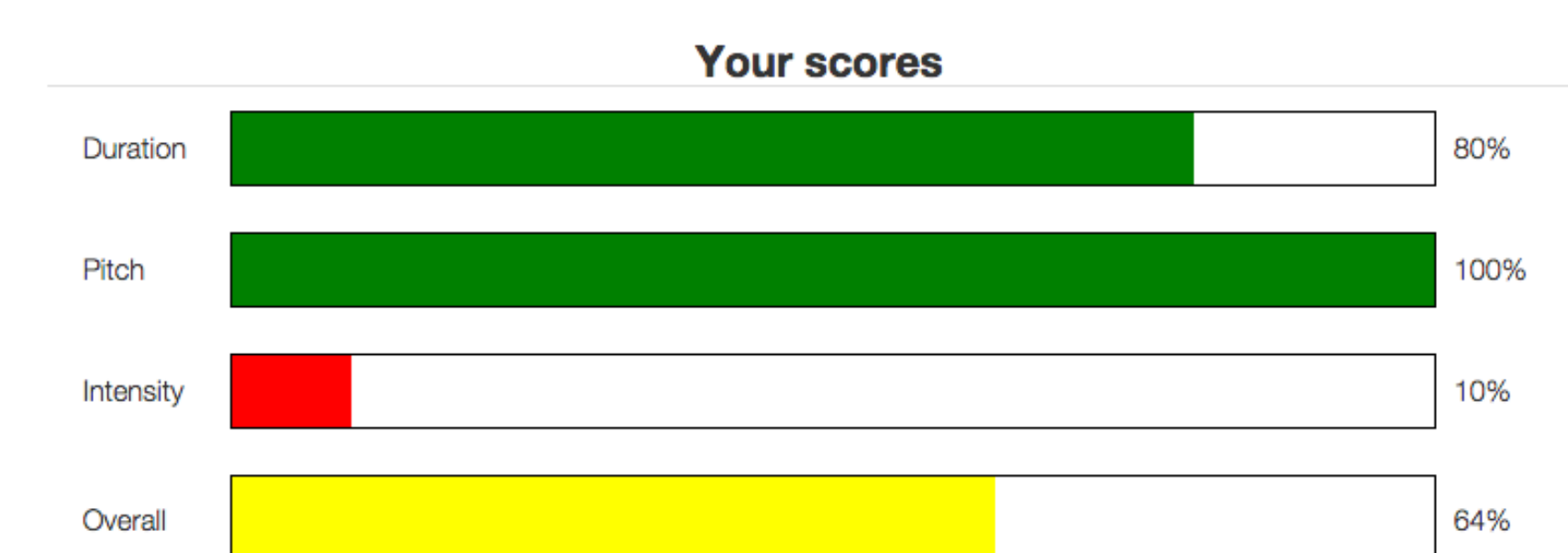


Figure 3: Feedback via skill bars.

- **Implicit visual feedback** (see fig. 2):
 - Graphical visualization of syllable prosody
 - Text stylization reflecting syllable duration
- **Implicit auditory feedback:**
 - Original learner and reference utterances (see fig. 2)
 - Prosodically modified learner utterance (using JSnoori [2])
- **Self-assessment** questionnaire for learner to complete

Administrative interface for teachers/researchers

This simple graphical interface allows a language teacher or CAPT researcher to create new exercises for students to complete. Each exercise features a specific combination of the various diagnosis and feedback options described above.

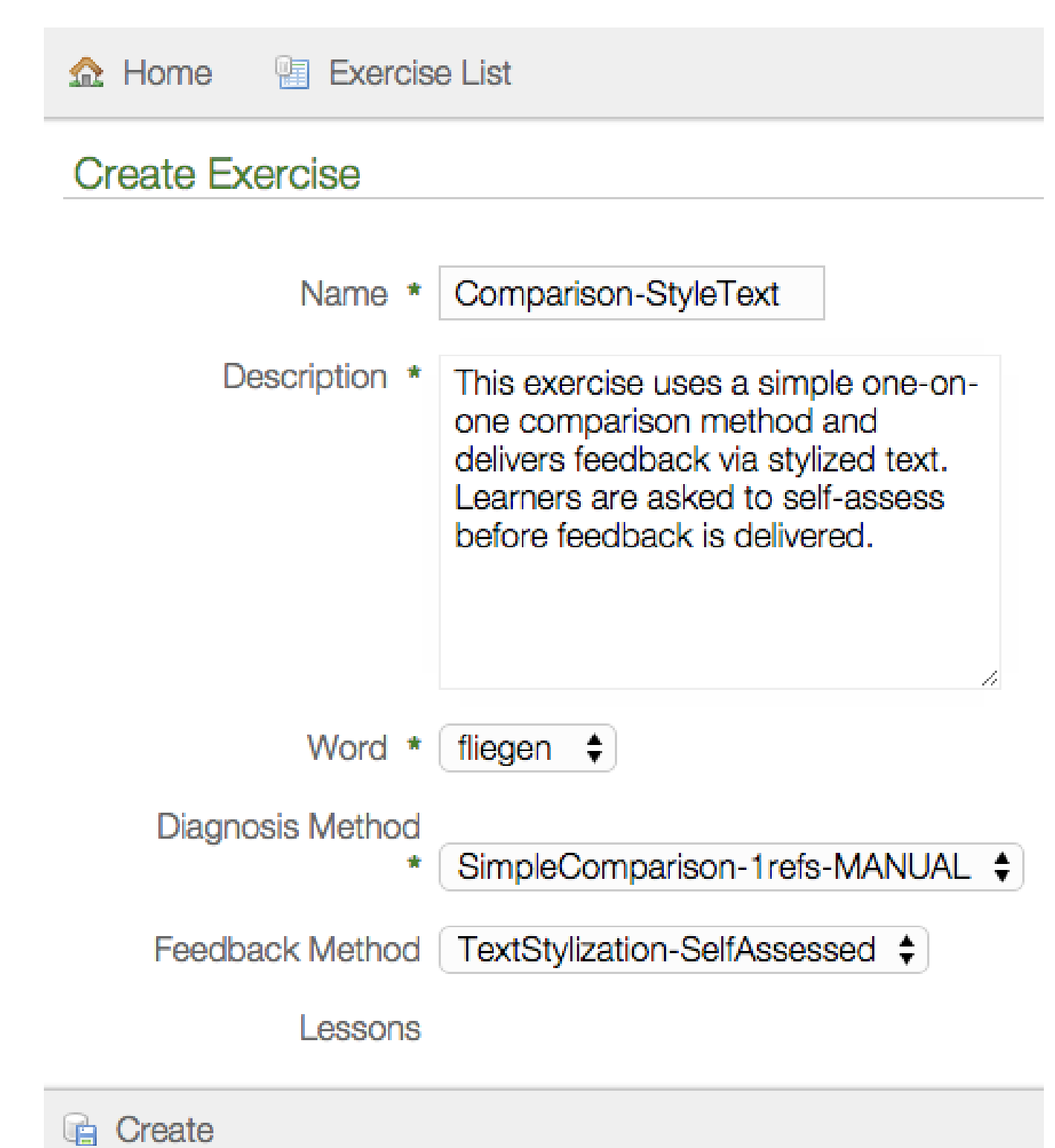


Figure 4: **[TODO caption]**

References

- [1] A. S. Vakil, "de-stress," <http://github.com/vakila/de-stress>.
- [2] LORIA Speech Team, "JSnoori," <http://jsnoori.loria.fr>.
- [3] A. S. Vakil and J. Trouvain, "Automatic classification of lexical stress errors for German CAPT," in *SLaTE*, 2015.