

NLP tools for Intelligent Learner Utterance Evaluation

Johann HALLER, Michael CARL,
Sandrine GARNIER
IAI
Martin-Luther-Str. 14
Saarbrücken, Germany, 66111
hans@iai.uni-sb.de

Brigitte STROEDE, Lutz WIND
IDF
Plöck 55,
Heidelberg, Germany, 69117
idf@mail.idf.uni-heidelberg.de

Abstract

In this paper, work on an intelligent learners utterance evaluation program is described. After a brief introduction to the basic multimedia learning program, the motivation for the development of an NLP package is given: learners are supposed to stick closer to the program and to spend more time and energy with it if they are given intelligent diagnosis of their errors. This has been achieved by combining linguistic analysis tools with a distance measure processing.

1 Introduction

The Institute for German as a Foreign Language Philology at the University of Heidelberg offers an M.A. programme in German Linguistics and Literature which is designed to meet the needs of students who are not native speakers of German and intend to become German lecturers or teachers, or to follow careers in related fields.

A course of studies of this kind sets very high standards in language proficiency since participants have to learn how to write complex texts in an academic context. Given the fact that the level of proficiency required for admission to German universities is centralised in the form of the DSH (Language Test for Admission to Universities in Germany), this standard entails a level of competence, which is insufficient for this course of studies. Specialised teaching and learning materials which address the student's to build up the required level of written competence, are practically non-existent. Existing packages focus almost exclusively on German for beginners or German at an intermediate level. Computer-based materials which are available commercially are generally restricted to drag and drop exercises or to filling in gaps in isolated sentences. It goes without saying that tasks in which clicking takes the place of writing are of limited use in the development of writing skills.

Thanks to the financial support of the Klaus Tschira Foundation, the development of computer-

based materials for the purposes in hand is now underway and is being carried out within the context of the project "Application and Development of Multi-media Programmes for the Improvement of Teaching at University Level." The first module developed within the project deals with the structural patterns and difficulties associated with one of the marked features of specialist texts – the high frequency of complex nominalisations, the so called nominal style. The exercises, which are designed to improve both text comprehension as well as text production, are embedded in contexts which are similar in type to the student's reading agenda. The exercises are based on transformations at text level and require the student to transform complete sentences in a way which conforms to the context in which the sentence is embedded. One of the main features and challenges in designing this module is the provision of comments which anticipate possible errors made during transformation, and can thus guide the student toward an appropriate solution for the given context. Further assistance is provided by hyperlinks to reference grammars, lexica, etc.

This module is available on the internet under the name of *ProGram* (www.idf.uni-heidelberg.de/forschung/multimedia/)

The implementation of ProGram in language classes was successful in that students gave very positive ratings for the overall structure, the type of exercises, and their content. By contrast the ratings for the automatic responses provided by the programme were not as positive, given the range of errors made by the user and those anticipated by the programme. In other words the system of pattern matching has to get more intelligent. In order to meet this target a joint project has been initiated with IAI (www.iai.uni-sb.de) with the aim of developing an adequate correction programme with the required potential. Ultimately, this should enable students in the Arts Faculty to deal with the standards required when writing specialist texts.

In this recently started project ProGram-AE (<http://www.idf.uni-heidelberg.de/index.html>), the goal is to achieve a maximum of integration

between the didactic strategy, the foreseen learners way through the exercises and the possibilities of the NLP tools (called the ‘Utterance evaluator’).

These NLP tools support language learners in the following three areas:

- Intelligent spell checking
- Grammar checking
- Distance evaluation between the learners utterance and the correct solution(s)

Spelling and grammar checking are based on long experiences in NLP and have been tested and used in industrial and commercial environments, for example in the ‘DUDEN Korrektor’ (www.duden.de), a spelling and grammar checking tool for German, in CLAT (Controlled Language Authoring System) <http://iaai-sb.de/iaide/en/clat.htm> for term and style checking and in other tools for automatic indexing and classification.

There are now several possibilities to integrate these tools into the didactic process of ProGram. We omit a description of the linguistic tools, rather give an overview of how they are used within ProGram and focus more on the distance evaluation.

2 Outline of ProGram

A first group of exercises consists in transforming a German Noun Phrase with a nominalized infinitive and a genitive attribute into a simple sentence. Two steps are required in this exercise: first the learner has to type the base form of the verb (or adjective) into the empty field and in a second step he uses the verb to form the sentence. The following lines show the noun phrase repeated here as 1, which is expected to be paraphrased similar to the sentence in 2.

1. *das Lernen einer fremden Sprache*
The learning of a foreign language
2. *Man lernt eine fremde Sprache*
One learns a foreign language

As the learner clicks on the button “Bewerten” (evaluate) in the lower part of the screen, ProGram starts a number of processing steps which we shall describe in more detail in the next sections.

A sample screen should just illustrate the integration of ProGram and the evaluation package:

ProGram is implemented in Java and each exercise comprises a number of anticipated correct answers and a number of anticipated wrong answers. In case the learner’s answer corresponds to an anticipated correct answer the sentence is highlighted in green and the learner can go on with

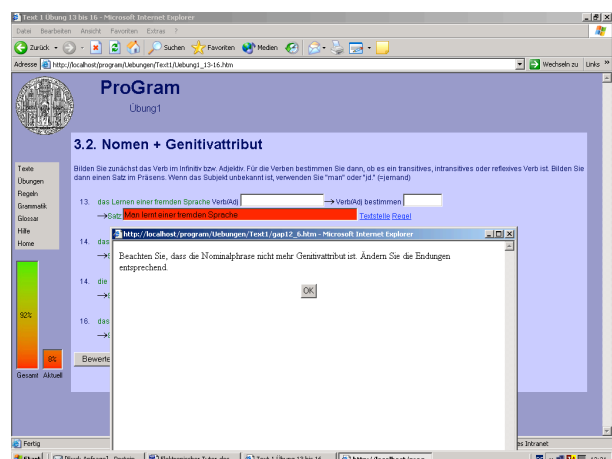
the next exercise. Otherwise, the sentence is highlighted in red and a commenting box appears which provides hints as to how the sentence might be corrected. While the sentences 3 through 9 show possible correct answers, 10 through 12 show some wrong anticipated answers.

3. *Man lernt eine fremde Sprache*
4. *Jd. lernt eine fremde Sprache*
5. *Eine fremde Sprache wird erlernt*
6. *Eine fremde Sprache wird gelernt*
7. *Man lernt eine Fremdsprache*
8. *Jd. lernt eine Fremdsprache*
9. *Eine Fremdsprache wird gelernt*
10. **Man lernt einer fremden Sprache*
11. **Jd. lernt einer fremden Sprache*
12. **Einer fremden Sprache wird gelernt*

Obviously, there are many more wrong answers possible than those listed above. There are also many more correct answers possible, but it’s particularly the wrong answers that merit a more detailed analysis and well-reflected comments to guide a language learner. Thus, in figure 2, the learner’s utterance corresponds to an anticipated error 10: The learner simply copied the genitive form of the attribute *einer fremden Sprache*. However it should be the accusative object here and endings have to be adjusted accordingly. Since this is a frequent error, the author of the unit anticipated this solution and presents an exact comment to the learner:

Beachten Sie, dass die Nominalphrase nicht mehr Genitivattribut ist. Ändern Sie die Endungen entsprechend!

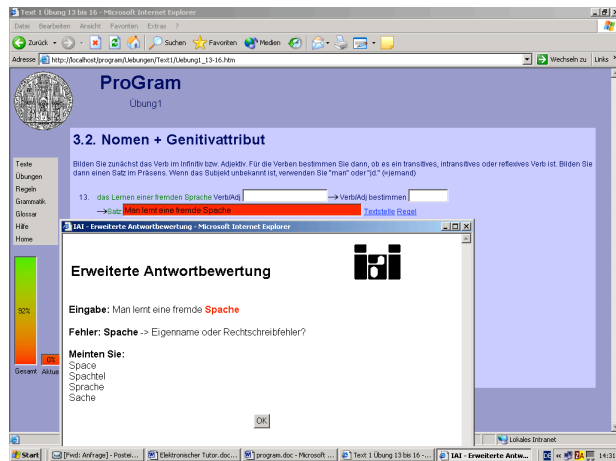
Pay attention to the fact that the noun phrase is not more genitive. Modify the endings!



3 Intelligent spell and grammar checking

In case the learner’s answer does not correspond exactly to one of the anticipated correct or wrong

answers, more sophisticated processing devices are required. In order to guide a learner and generate helpful hints, we distinguish between three processes: first it is checked whether the learner's answer contains spelling errors. In this case, the error is presented on the screen, together with several alternatives containing the right solution:



If no spelling error is detected, the answer is checked for grammatical correctness. Last a commenting distance is computed between the learner’s utterance and the anticipated correct answers.

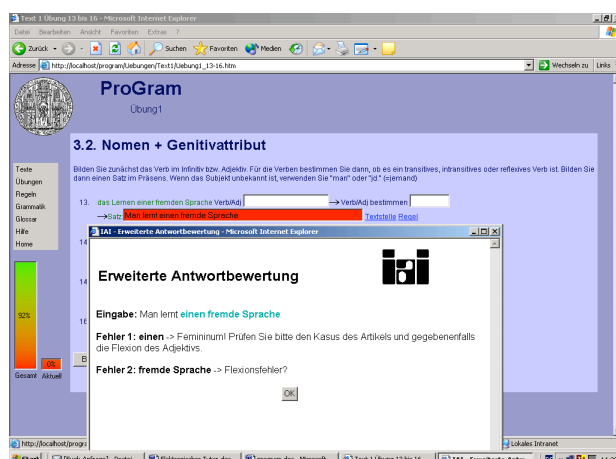
4 Grammatical errors within a learners utterance

Most error prone in grammar checking is the agreement within a noun phrase. In example 14, the grammar checker highlights the error group (*einen fremde Sprache*) and gives the following hint:

Prüfen Sie bitte den Kasus des Artikels und ggf. die Flexion des Adjektivs!

Verify the case of the article and the inflected adjective!

13. *Man lernt eine fremde Sprache*
One learns a (masc.) foreign language (fem).



5 Distance measure

In case the learners answer contains one or more unanticipated words or an error has not been detected by the grammar or the style checker, the learner's utterance is matched against the set of anticipated correct answers in a more subtle way than merely by matching their word forms. That is, as the learner's utterance, also the correct answers are linguistically analysed. A layered matching computes the similarity between a learner's incorrect answer and the anticipated analysed correct answers. The differences between the learner's wrong answer and the closest correct answer are marked and commented such that the learner can do appropriate modifications. 14 is very similar to 10, which was one of the wrong responses coded by the unit author, but this time the diagnosis is generated automatically (here shown directly in the XML notation which is used as interface between ProGram and the evaluation package):

Learner response

14. *Hans lernt einer fremden Sprache*
John learns a foreign language

`<style code=Bitte ein anderes Wort benutzen. Auf Inflektion achten.>Hans<\style>`

lernt

`<style code=Auf Kasus achten.>einer<\style>`

`<style code=Auf Kasus
achten.>fremden<\style>`

Sprache

This process could still be improved in comparing the learners utterance with the coded wrong answers, but here a more detailed and sophisticated analysis will be needed. It is no always trivial to decide which wrong answer was intended by the learner.

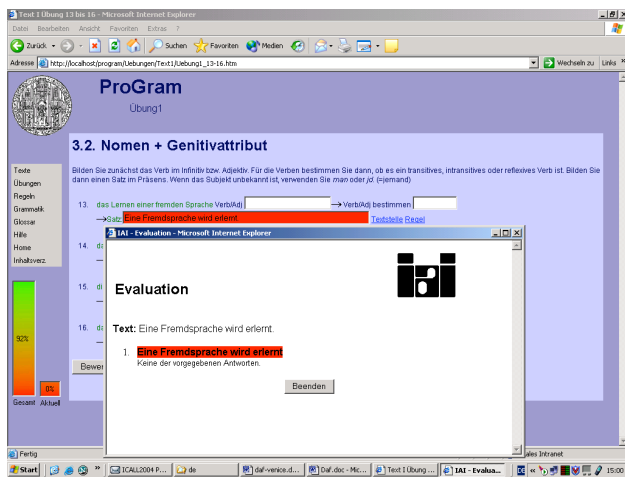
If, however, the distance between the learners response and the correct answers exceeds a certain degree, a more general hint is issued:

Learner response

15. *Eine Fremdsprache wird gelernt.*

Keine der vorgegebenen Antworten!

Your answer does not correspond to the task.



6 Conclusion

In this paper we have described an integrated system for the intelligent evaluation of learner utterances. The techniques used are a combination of industrially tested NLP analyses, checking tools and a distance-measure process taking into account the NPL analysis results both of given learner responses and actual utterances.

In a second step, several groups of learners will be evaluated with respect to their behaviour in front of such techniques; it will be registered if the number of abandonments can be significantly reduced by using the described AI techniques.