Dealing with Ambiguity in a Foreign Language ITS

November 6, 2006

Amanda Nicholas
amn39@student.canterbury.ac.nz
Department of Computer Science and Software Engineering
University of Canterbury, Christchurch, New Zealand

Supervisor: Brent Martin
brent.martin@cosc.canterbury.ac.nz

Abstract

Two competing forces are at work in developing Intelligent Tutoring Systems: the desire to provide comprehensive feedback, and the desire to ensure that the knowledge learned will transfer to real life situations. One factor that can prevent comprehensive feedback being given is the problem of ambiguity, Menzel [13]. Methods for dealing with ambiguity exist. We explored one such method in this study, by constructing two systems for teaching German adjective endings. Tutor 1 requires the student to specify their complete mental model of the problem, in addition to answering the question. The system can thus correct their specific misconceptions. Tutor 2 only requires the student to answer the question.

The two systems were evaluated by first year German students at the University of Canterbury. Results of this evaluation are unclear, due in part to the small number of participating students. Tutor 1 gave greater learning overall, however the degree to which students learnt the constraints that map only to the adjective endings was approximately equal for both tutors. Further research is needed to provide a conclusive result.

Keywords: Constraint-based Modeling (CBM), Intelligent Tutoring Systems (ITS), Computer Aided Language Learning (CALL), foreign language learning

Contents

List of Figures

List of Tables

1 Introduction

Dealing with ambiguity is a serious problem in developing Intelligent Tutoring Systems for foreign languages [13]. Although the system can detect that the student has made an error, the source of this error may be difficult to determine. Menzel [13] defines four sources of ambiguity: limited observability, polysemy, alternative conceptualisations of domain knowledge, and structural uncertainty. In a domain with high ambiguity, feedback messages can be difficult to determine. Good feedback [24] should refer the student to the underlying domain principle. If it is not possible to determine which domain principle has been broken, correctly targeted feedback cannot be given.

One approach to avoid ambiguity is to require the student to specify the intermediate steps they carry out mentally. This approach is not popular as "such an interaction renders the exercise somehow unnatural." [13]. Requiring the student to specify intermediate steps also raises the issue of transference [1]. When developing an ITS, the interface is generally designed to stay as close to the real world as possible, in order to ensure that the skills learnt on the computer will transfer to the real situation. By requiring the student to specify additional information, the transference of skills may be weakened.

This research compares two tutors – one that matches the real world more closely, and one that decreases ambiguity as much as possible. Two Intelligent Tutoring Systems were developed for the domain of German adjective endings. German adjective endings is a domain with a high level of ambiguity. An error in an adjective ending could be caused by a number of factors: mistaking the case, mistaking the gender, mistaking the article, or simply not knowing the correct ending in this situation.

Tutor 1 required the student to specify the gender of the noun, the case of the noun, and the type of article. These three factors specify exactly the ending required. By considering these factors, the tutor could determine if it was in fact the ending of the adjective that the student had specified incorrectly, or if they had some misconceptions about the sentence. The tutoring system could therefore provide targeted feedback based on the student's misconceptions. Tutor 2 only asked the student to specify the adjective ending. This matches what is required in the real world, but means that the tutoring system has to make assumptions about what error the student has made.

1.1 Research Objectives

Our hypotheses are:

1. Requiring the student to specify additional information about the task will not degrade learning.
2. Requiring the student to specify additional information about the task will improve learning.

1.2 Report Structure

Chapter 2 presents background information. Section 2.1 describes Intelligent Tutoring Systems, and Constraint-based Modeling. Section 2.2 gives a brief outline of German grammar, with an emphasis on adjective endings. The difficulty of this domain is also described.

Related work is described in Chapter 3. Menzel's [13] work in determining the four causes of ambiguity is discussed. Finally, the CBM authoring tool WETAS [12] is presented.

Chapter 4 describes how the system was designed, including the interface, problem set, and choice of constraints.

The results of the evaluation study are presented in Chapter 5, including pre- and post-test results, learning curves, and anecdotal evidence collected from the students and lecturers.

A discussion of these results is given in Chapter 6, plus an outline of limitations of this research and future work.

Finally, the report is concluded in Chapter 7.

Appendix A contains all comments the students wrote as feedback during the evaluation. The pre- and post-tests are included in Appendix B. The complete set of constraints can be found in Appendix C.

2 Background

2.1 Intelligent Tutoring Systems

Computers have been used for educational purposes for more than thirty years [2]. The first systems were known as computer-aided instruction. These systems were not tailored to a specific student's needs, but followed the same path for all students. From this beginning, Intelligent Tutoring Systems (ITS) evolved. ITS aim to achieve the results of one-on-one human tutoring, without the resource requirements. They adapt to the student by giving feedback that corresponds directly to the mistakes the student has made, and by providing further problems at an appropriate level for the student to continue learning. ITS achieve this by keeping a model of each student's current knowledge, and comparing it to a domain model.

2.1.1 Constraint-based Modelling

Ohlsson [17] outlined a theory of learning from performance errors, where learning is seen as consisting of two phases. The first, error recognition, requires the learner to realise that an error has been made. The learner may be informed of this by their tutor, whether human or machine, or they may discover the error alone, by comparing the desired outcome to the actual outcome. The second phase, error correction, is when the learner improves the procedure they used, in order to perform the task more accurately in the future.

Constraint-based Modeling (CBM) is an approach to student modeling and ITS development based on the theory of learning from performance errors [16]. Earlier student modeling methods, such as Model Tracing [1], required the developer to specify the possible paths that must be followed to solve a problem. For the tutor to be able to respond appropriately, every possible error must be collected into a bug library. CBM is not concerned with the path a student has followed to reach a particular state; it concerns itself only with the current state. If the current state breaks domain principles, the solution cannot be correct.

In CBM, the domain is described using constraints, where each constraint consists of two conditions: the relevance condition and the satisfaction condition. These pairs of conditions describe the domain. Each piece of domain knowledge can be understood as

> *If <relevance condition> is true, then <satisfaction condition> must also be true, otherwise something has gone wrong.*

If the student's solution satisfies the relevance condition, it must also satisfy the satisfaction condition. If it fails to satisfy the satisfaction condition, then there is an error in the student's solution. A feedback message is associated with each constraint, and in the event of a constraint violation, this feedback is displayed.

CBM tutors have been developed for a variety of domains, including SQL, punctuation and database design [14].

2.2 German Grammar

The resources consulted for this section include the textbook for GRMN115 [21], grammar handbooks [7, 22, 20, 3], and material received from the German department.

2.2.1 Nouns

German nouns have grammatical gender, which can be masculine, feminine or neuter. This is not determined by the object the noun refers to, but is simply a property of the noun. For example, 'table' is masculine, as is 'skirt'. 'Woman' is feminine, but 'girl' is neuter. For this reason, the gender of nouns cannot be predicted, but must be memorised when vocabulary is learnt.

Nouns also have a case, which varies depending on the role the noun plays in a sentence. A noun can be in one of four cases: nominative, accusative, genitive or dative. The uses of these cases are given in Table 2.1.

Case	Use
Nominative	The subject of a sentence
Accusative	The direct object of a sentence
Genitive	Possession - the noun that does the possessing is in the genitive
Dative	Used with most prepositions, and for the indirect object

Table 2.1: Use of noun cases

For example, the sentence "The boy gave the man the woman's book" has four nouns. 'The boy' is the subject of the sentence, the object carrying out the verb, and so it is in the nominative case. 'The book' is in the accusative case, as it is the object the verb acts upon – it is the object given. 'The woman' is in the genitive case, as the woman possesses the book. Finally, 'the man' is in the dative case, as it is the indirect object – the object the verb acts on indirectly.

Nouns can also be optionally preceeded by an article, of which there are two types. The direct article corresponds to English 'the', and the indirect article corresponds to English 'a'. There may also be no article present.

2.2.2 Adjectives

In German, adjectives must agree with the nouns they modify. This means that the ending of an adjective varies based on the gender and the case of the noun, and whether the noun is preceded by a definite article, indefinite article, or no article. These endings can be seen in Table 2.2, Table 2.3 and Table 2.4.

	mas.	fem.	neu.	plural
Nominative	-e	-e	-e	-en
Accusative	-en	-e	-e	-en
Genitive	-en	-en	-en	-en
Dative	-en	-en	-en	-en

Table 2.2: Adjective ending when preceded by the direct article

	mas.	fem.	neu.	plural
Nominative	-er	-e	-es	-en
Accusative	-en	-e	-es	-en
Genitive	-en	-en	-en	-en
Dative	-en	-en	-en	-en

Table 2.3: Adjective ending when preceded by the indirect article

For example, take the sentence *Das graue Haus ist neu.* (The grey house is new). Here *Haus* is the noun, and its gender is neuter. The house is the subject of the sentence, and so it is in the nominative case. *Das* is the article, and it is the direct article. The adjective is *grau*, and it takes the ending *e* because, by

	mas.	fem.	neu.	plural
Nominative	-er	-e	-es	-e
Accusative	-en	-e	-es	-e
Genitive	-en	-er	-en	-er
Dative	-em	-er	-em	-en

Table 2.4: Adjective ending when preceded by no article

consulting Table 2.2, we can see that adjectives preceeding a neuter noun in the nominative case must end in *e*.

If we changed only the article in this sentence, so that it now read *Ein graues Haus ist neu.* (A gray house is new), the ending on the adjective changes also, from *e* to *es*.

It is important to note that the endings are not unique. The ending *e* appears in a number of situations, as does the ending *en*. This is one reason why these endings are ambiguous.

2.2.3 Difficulty of Domain

Adjective endings are a difficult topic for students to master. This is due to the number of endings that must be memorised, and the amount of knowledge required of the sentence to get the ending correct. During initial discussions with the German department, the topic of adjective endings was suggested as a good trial domain, as students in later years of study often still struggle with this topic.

Rogers [19] studied the main areas of weakness in students with more than four years of experience learning German. She states "... much anecdotal 'evidence' from teachers of German as a foreign language emphasises morphology as a major areas of weakness (e.g. adjective endings...". Her study showed that approximately 5% of errors made by advanced learners of German were errors in adjective endings. The number one cause of problems was in selecting gender, which could also affect the choice of adjective ending. Each error was only classified once, so if the cause of the error was due to mistaking the gender, it would not also appear as a mistaken adjective ending. The number of errors in adjective endings must be much higher than 5% when all reasons are considered.

Juozulynas[6] studied students with two years of experience learning German and found that

> "The biggest problem in the students' writing seems to be syntax ... Inflectional morphology with its much-feared endings takes second place. Syntax and morphology together make up 53% of the errors in the corpus."

Note that adjective endings are contained in inflectional morphology.

3 Related Work

This chapter outlines previous research of interest. Section 3.1 discusses Menzel's work in determining the effect of ambiguity on CBM tutors. The authoring shell WETAS is described in Section 3.2.

3.1 The Problem of Ambiguity

Menzel [13] identified four major sources of ambiguity that should be considered when creating CBM tutors, particularly for foreign languages. He suggests that constraints alone are not sufficient to provide enough information to respond to students appropriately. His primary focus is applying CBM to natural language parsing.

The four sources of ambiguity he defines are:

- a limited observability of internal variables of the problem domain
- polysemy of symbols used in the problem domain
- alternative conceptualisations of domain knowledgeInteractiveGerman, Kontakte, SRG, GermanIn-Review, PracticeGrammar
- uncertainty about the intended structure of the students solution

3.1.1 Limited observability

In certain domains, variables which relate to the solution are not explicitly stated, and so cannot be observed by the tutor. This may mean that the tutor is unable to tell which constraint has been violated, or possibly even whether or not a constraint has been violated at all. Menzel notes that "Although sometimes ... students can be asked to provide the missing information, in many cases such an interaction renders the exercise task somehow unnatural." An example of this is when adding large numbers, the carry between columns can be held inside the student's head. This value can directly affect the solution, but the tutoring system is not aware of the value the student has allocated it.

3.1.2 Polysemy

Polysemy is defined as a symbol which has multiple meanings. The correct meaning must be selected according to the context. Menzel gives as an example the word 'fish'. 'Fish' can be either single or plural – 'This fish stinks' or 'These fish stink'. If the student writes 'These fish stinks', it is not possible to determine from this information only what the student intended to say.

Polysemy can be considered a special case of limited observability. The form of the word is visible, but the value the word is given is not. Asking the student to fill in the missing information is not considered an appropriate solution: "the intended selection of a grammatical feature cannot easily be elicited from the student without heavily disturbing the natural conditions of human language production."

3.1.3 Alternative conceptualisations of domain knowledge

If a constraint is violated, one assumption is that the student did not know the underlying concept. Such an error is known as a *rule error*. However, a second possible view is that the student was aware of the underlying grammatical concept, but made wrong assumptions about the lexical value of words. This is known as a *fact error*, as the student was unaware of specific information about the word.

These two mistakes are caused by different misunderstandings by the student, and require different remediation.

3.1.4 Structural uncertainty

Parsing a sentence into the correct components is not a trivial task. The same word can take a different role in a different sentence. A well known example is the pair of sentences "Time flies like an arrow" and "Fruit flies like a banana", where 'flies' is a verb in the first sentence, and a noun in the second sentence.

Structural uncertainty is a major challenge in natural language parsing applications.

3.1.5 Ambiguity and Adjective Endings

German adjective endings suffer from three of the four defined sources of ambiguity. Limited observability and polysemy are both present in the multiple possible meanings of a single ending. For example, a student could correctly give an adjective requiring a nominative, masculine, definite article ending, the ending *e*. However, it is also possible that the student thought that the adjective required a nominative, feminine, definite article ending, for which the ending is also *e*. The student might even believe that the adjective requires a dative, masculine, definite article ending, which should be *en*, and might have given it the ending *e* incorrectly, based on their knowledge. Without awareness of the student's thought processes, the tutor is unable to determine if the student has answered the question correctly on purpose, or by mistake.

This problem also encompasses that of alternative conceptualisations of domain knowledge. When the student incorrectly gives an adjective ending, it could be due to either a rule error, or a fact error. If the student does not know the gender of the noun, or the case of the noun, they have made a fact error. If the student has correctly determined the case, gender and article, and still gives the adjective ending incorrectly, they have made a rule error. They do not know the underlying grammatical principle that determines the adjective ending. It is also possible for a student to make a rule error and a fact error simultaneously.

Because the domain selected is highly constrained, we do not make use of natural language parsing, and therefore structural uncertainty is not an issue.

3.2 WETAS

WETAS is an authoring tool for CBM tutors [12]. It is web-based, and provides all of the functionality that is not domain dependent. It is written in Allegro Lisp [4], and implemented as a web server [5].

WETAS provides problem selection, answer evaluation, student modelling, feedback, and the user interface. The author only needs to provide the problem and solution set, the constraints, and, if desired, a different interface.

A number of tutors have previously been implemented in WETAS, including a tutor for English language skills, and a tutor for SQL [12, 9, 11, 10].

WETAS was used to develop the ITS used in this research, as the purpose was not to create a tutor, but to determine what effect different constraints and interfaces can have on learning.

4 Design and Implementation

The tutors were developed using WETAS [12, 9]. WETAS was developed by Brent Martin, and is a shell that can be quickly adapted to provide basic functionality for an ITS. It provides student modelling, student management, and other features. The developer must supplement this with the problem set, the necessary constraints and, if desired, an interface.

This chapter outlines how the problem set, interface, and constraints were developed.

4.1 Problem Set

The problem set comprised of 55 problems. This set of problems was identical for both tutors. Some were obtained from a number of sources [23, 7], however, most problems were written especially for this ITS.

It was decided that, for the purpose of this study, it was not necessary to have the next problem selected adaptively. The problems were all of approximately equivalent difficulty, and the constraints were spread throughout the problems. This meant that there was one less level of complexity to consider in the results.

An example of one of the problems in the tutor is

> Die ? Blumen gefallen mir. (bunt) *[I like the colourful flowers]*

4.2 Interface

The two tutors shared a very similar interface. In the centre of the screen was an area for the student to answer the question. Below the problem, a selection box allowed the student to choose the desired feedback level, and then request feedback. Feedback messages appeared at the bottom of the screen.

The problem was displayed in the form of a sentence. A gap was left where the adjective should be, and adjective to be inserted was given in brackets at the end of the sentence. This was a format the students were familiar with, because it had been used during class and quizzes.

Students using Tutor 1 were asked to fill in the gender and case of the noun, the article type, and the adjective with its ending. The possible answers for gender, case and article were all given in combo boxes. This ensured that there would not be problems with students referring to the same concept by a different name, or misspelling names. Below the combo boxes, there was a text field for the student to fill in the adjective. (Figure 4.1)

Students using Tutor 2 were only asked to fill in the correct adjective and ending. A textbox for the student to fill in was placed in the correct location in the sentence. (Figure 4.2)

4.3 Constraints

Ohlsson [16] stated that constraints should express domain principles. Zakharov et al. [24] found that feedback is more effective when it tells the learner

> (a) where exactly the error is, (b) what constitutes the error (perform blame allocation), and (c) refer the user to the underlying concept of the correct solution (revise underlying knowledge)

Therefore, a good constraint should be express a domain principle, and have an attached feedback message that states the underlying principle as well as explaining the error.

Determining appropriate constraints for this domain was a challenging task, as the underlying principles were not immediately apparent. The principles are either very low-level: “the adjective ending for neuter,

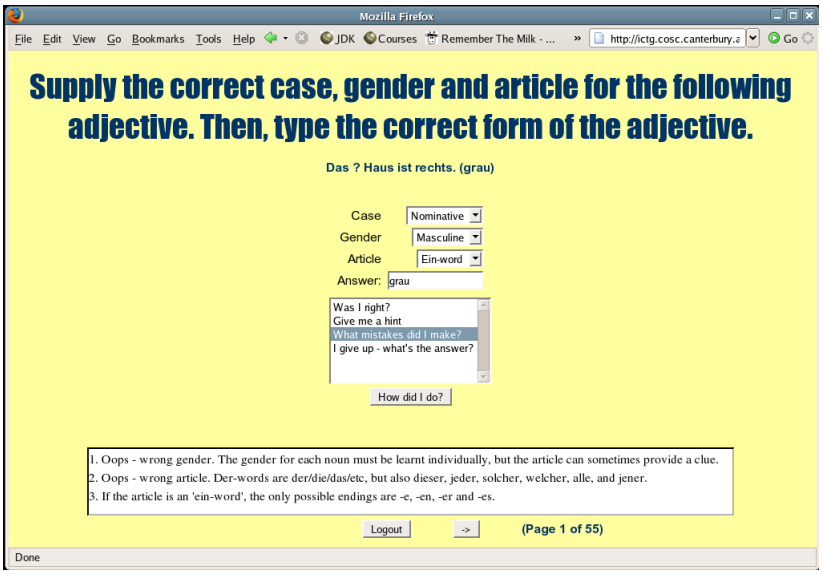

Figure 4.1: Interface for Tutor 1

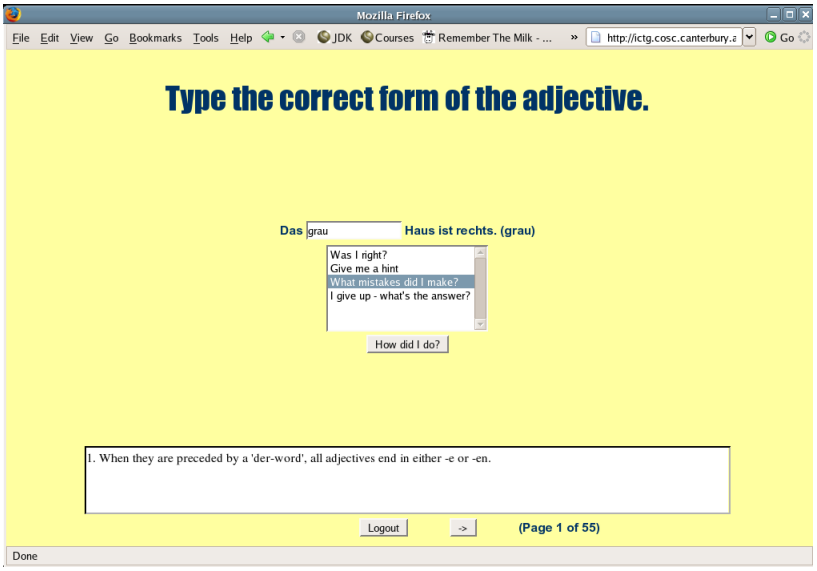

Figure 4.2: Interface for Tutor 2

nominative, indefinite article should be *e*"; or very high-level: "In many instances, the adjective ending is the same as the ending of the definite article." [21]. The first example is too specific. Providing the student with feedback at this level would be equivalent to providing them with the answer. However, the second example is too general, and does not apply in all cases. Feedback provided at this level could frustrate the student.

In order to provide appropriate feedback to the student, artificial domain constraints were determined. These were sourced from a number of German textbooks [7, 21, 22, 20, 3], which contain advice on how students can remember the endings more easily. They typically explain a pattern in the endings, for example that every adjective after the indirect article endings in either *e* or *en* (see Table 2.2).

The constraints authored can be divided into three groups. The first set of constraints are used for error checking: ensuring that the student has answered the question and used the appropriate adjective. The second set occurs only in Tutor 1, and check whether the student has specified the gender, case and article correctly. The third set of constraints is the group that checks the validity of the adjective ending.

The constraints are written in AllegroLisp [4], and integrated into the WETAS Authoring Shell [9].

Both sets of constraints can be found in Appendix C.

4.3.1 Tutor 1

Tutor 1 has 33 constraints. Six are for error checking; ten are for checking that the student has specified the case, gender and article correctly; the remaining constraints check the adjective ending. The adjective ending is checked for validity with respect to the case, gender and article the student considers to be correct. Incorrect values for case, gender and article will trigger other feedback messages. In this manner, the system determines whether the student has made an error because they have inaccurate knowledge about the sentence, or because they do not know their adjective endings, ie whether they have made a rule error or a fact error.

Although Tutor 1 has more constraints for checking the adjective ending than Tutor 2, they check the same principles. In Tutor 2, many endings can be checked for validity by comparing them against the ideal solution. In Tutor 1 this is not possible, as the ending is checked for validity against the student's assumptions, not the actual state of the sentence. Because of this, there is often a number of parallel constraints that produce the same error message in different situations.

For example

```
(30
        "If there is no ein-word or der-word, the adjective takes the ending that the der-word would have if
           it was present."
        (and
         (match SS ARTICLE ("N"))
         (match SS ANSWER (?something ?*))
         (or-p
          (match SS GENDER ("F"))
          (match SS GENDER ("P"))
          )
         (or-p
          (match SS CASE ("N"))
          (match SS CASE ("A"))
          )

         )
        (match SS ANSWER (?* "e"))
        "ANSWER")

        (31
        "If there is no ein-word or der-word, the adjective takes the ending that the der-word would have if
           it was present."
        (and
         (match SS ARTICLE ("N"))
         (match SS ANSWER (?something ?*))
         (match SS GENDER ("N"))
         (or-p
          (match SS CASE ("N"))
          (match SS CASE ("A"))
          )
         )
        (match SS ANSWER (?* "e" "s"))
         "ANSWER")
```

For contrast, the same constraint is represented in Tutor 2 by

```
(30
        "When there is no ein-word or der-word, the adjective takes the ending that the der-word would have if
           it was present."
        (and
         (match IS ARTICLE ("N"))
         (match SS ANSWER (?something ?*))
         (match IS ANSWER (?*w1 ?second ?last))
         )
        (match SS ANSWER (?*w2 ?second ?last))
        "ANSWER")
```

4.3.2 Tutor 2

This tutor has twelve constraints. Three are for error checking, and the remaining nine check the ending the adjective has been given.

Because the only information available to the tutor is the ending the student has given the adjective, the tutor provides feedback relative to the correct gender, case and article. It is assumed that the student knows this information, but may be unaware of the ending that matches correctly. This means that the tutor considers all mistakes to be rule errors, not fact errors. An example of one such constraint is given below.

```
(10
        "When they are preceded by a 'der-word', all adjectives end in either -e or -en."
        (and
         (match IS ARTICLE ("D"))
         (match SS ANSWER (?something ?*))
         )
        (or-p
         (match SS ANSWER (?*w2 "e" "n"))
         (match SS ANSWER (?*w1 "e"))
         )
        "ANSWER")
```

The relevance clause of this constraint checks that the sentence containts a definite article (“D”), and that the student has answered the question. If this is true, the student’s answer must end in -e or -en, as all adjectives end in -e or -en after the definite article. If the student’s answer does not end in -e or -en, the system assumes that they have forgotten the rule, not that they have not realised that the sentence contains a definite article.

5 Evaluation

An evaluation study of the two tutors was conducted on the 6th of September 2006 at the University of Canterbury, Christchurch. Students enrolled in GRMN115, a beginning German course, used one of the two systems over one class period. The students had been taught adjective endings previously in class, however there was a two week holiday period between when the topic was taught and when the study was carried out.

Section 5.1 describes how the evaluation was designed and carried out; Section 5.2 gives the results of the study. Three types of results are considered. First of all, the results of the pre- and post-test are given. This is followed by a description of the learning curves achieved, and how well the constraints were learned. Finally, feedback received from the participating students is discussed.

5.1 Experiment Design

The evaluation study was carried out on GRMN115 students. There were 32 students enrolled in this course, in two streams, taught by different lecturers. Students were not required to attend the same stream consistently.

Students taking GRMN115 spoke no German at the beginning of the year: it is described as an introductory course. A pass in GRMN115 means the student is eligible to take GRMN108, the first year German course, in the future.

The purpose of the evaluation study was to compare the two systems. The class was divided into two even groups. This was done alphabetically by last name.

The evaluation took place during one lecture period, a time span of 50 minutes. The students were first asked to complete a pre-test. They then used the tutoring system for as long as time permitted, or until they finished all 55 questions. Afterwards they completed a post-test.

The pre- and post-tests can be found in Appendix B. Two tests were written, each with six questions. All questions contained sentences of the form

> "Die _____ Jacke ist preiswert. (gelb)" *[The yellow jacket is affordable]*

The student was expected to transfer the adjective (here 'gelb') into the gap in the sentence, and give it the appropriate ending. The final three questions also asked the student to specify the gender and case of the noun present in the sentence, and the type of article preceding the noun.

To allow for any difference in the difficulty of the pre- and post-tests, Test 1 was used as the pre-test for Stream A, and the post-test for Stream B; Test 2 was used as the post-test for Stream A and the pre-test for Stream B.

5.2 Results

23 students took part in the evaluation. 12 students used Tutor 1; 11 students used Tutor 2. The spread of students between Tutor 1 and Tutor 2, and Stream A and Stream B can be seen in Table 5.1. Statistics about the system usage can be seen in Table 5.2.

From Table 5.2, we can see that students using Tutor 2 solved more problems with fewer attempts than those using Tutor 1. This result is unsurprising, as students using Tutor 2 only had fill in one value correctly, whereas students using Tutor 1 had to fill in four values. Students using Tutor 1 also saw more messages. This is also unsurprising, as there were more constraints for Tutor 1 and thus more messages to be seen.

	Tutor 1	Tutor 2	Total
Stream A	9	5	14
Stream B	3	6	9
Total	12	11	23

Table 5.1: Allocation of students to Tutors and Streams

	Tutor 1	Tutor 2
Attempted Problems	22.33	51.82
Solved Problems	20.58	49.45
Attempts per Problem	4.03	1.97
Seen Messages per Problem	4.97	1.45

Table 5.2: System usage information

5.2.1 Pre- and Post-test Results

Unfortunately, the pre- and post-test were not of comparable difficulty. Over all students, irrespective of which tutor the student used or whether the test was taken as a pre- or post-test, the average score for Test 1 was 83%, and the average score for Test 2 was 65%. This means that the scores for the pre-and post test are not directly comparable.

	Pre-test (Test 1)	s.d.	Post-test (Test 2)	s.d.
Tutor 1	76%	26.50	61%	16.67
Tutor 2	80%	13.94	70%	13.94

Table 5.3: Pre- and Post-test results for Stream A

Stream A was given Test 1 as a pre-test, and Test 2 as a post-test. Table 5.3 shows that results on the post-test were poorer than those for the pre-test. This is due to the increased difficulty of Test 2. Students using Tutor 1 unimproved more than those using Tutor 2.

	Pre-test (Test 2)	s.d.	Post-test (Test 1)	s.d.
Tutor 1	67%	16.67	94%	9.62
Tutor 2	64%	24.53	83%	10.54

Table 5.4: Pre- and Post-test results for Stream B

Stream B was given Test 2 as a pre-test, and Test 1 as a post-test. The results are shown in Table 5.4. The results for the pre-test are comparable, but in Test 1, students using Tutor 1 outperformed those using Tutor 2. This result is not convincing, however, as the group using Tutor 1 consisted of only three students.

We are unable to conclude anything from these results, as the difference between the difficulty of the two tests was too great. What small results we did find are inconclusive, as students using Tutor 1 improved more than those using Tutor 2 in Stream B, while in Stream A those using Tutor 2 unimproved less.

Thus far we have only compared results of the six adjective ending questions in Test 1 and Test 2. Let us now consider the performance of students in specifying the gender and case of the noun in the sentence. The students were also asked to specify the article present in the sentence, but this question was primarily misunderstood or ignored, so we will not consider it.

The results can be seen in Table 5.5 and Table 5.6. Although there is still some difference in difficulty between the two tests, the results are much clearer. From comparable levels in the pre-test, students using Tutor 1 achieved better results in the post-test. This is expected, as students using Tutor 1 were asked to specify the gender and case of the noun while using the system, while students using Tutor 2 were not.

	Pre-test (Test 1)	s.d.	Post-test (Test 2)	s.d.
Tutor 1	81%	15.47	80%	16.20
Tutor 2	83%	20.41	70%	21.73

Table 5.5: Results for specifying gender and case for Stream A

	Pre-test (Test 2)	s.d.	Post-test (Test 1)	s.d.
Tutor 1	72%	9.62	94%	9.62
Tutor 2	75%	17.48	81%	6.80

Table 5.6: Results for specifying gender and case for Stream B

5.2.2 Learning Curves

A learner is described as getting 'better' at doing a task when they perform the task with fewer errors. Plotting the number of errors against the number of attempts leads to a negatively accelerated curve, known as a learning curve. This can be modeled as a power curve, and is known as the power law of practice [18, 15].

Learning curves are often used to evaluate ITS, as other methods such as pre- and post-tests are often ineffective. The slope and fit of the learning curve can be used to determine how well the student is learning [8].

The formula for the power law is:

$$y = Ax^{-B}$$

A represents the intercept of the y axis, which is the error rate before the student has had any practise. The slope of the power curve is given by B, which indicates the speed at which the student is learning the material. Finally, the fit of the data to the power curve is also calculated. A poor fit may indicate that the domain model is not optimal.

Based on the student models created during the evaluation, we calculated the probability of each constraint being violated on the first occasion that it was relevant, then the second, third, and so on. These values were averaged over all students and all constraints to give learning curves. The x axis indicates the occasion on which a constraint was applicable, while the y axis shows the likelihood that the constraint was broken.

The learning curve for all constraints in Tutor 1 is given in Figure 5.1. This curve has a learning rate of 0.4284. The fit is $R^2 = 0.7623$, which is acceptable. The students using Tutor 1 learnt the constraints well.

Figure 5.2 gives the learning curve for Tutor 2. Both the learning rate of this graph, at 0.26, and the fit of this graph, at $R^2 = 0.6045$, are much lower than that for Tutor 1.

If we compare the two graphs directly (Figure 5.3), it is apparent that the students using Tutor 1 both learnt at a higher rate of 0.4284 vs. 0.2099, and that this rate more accurately fits the data, with a fit of $R^2 = 0.7623$ vs $R^2 = 0.4409$. Note that the values for Tutor 2 are different from those mentioned above because the graph is plotted over fewer occasions. This is because the data for Tutor 1 extends over fewer occasions, and so the data for Tutor 2 had to be cut off to match Tutor 1.

This data would appear to imply that Tutor 1 is superior to Tutor 2. However, the set of constraints must be considered. Tutor 1 has constraints pertaining to gender, case, and article, as well as to adjective endings, while Tutor 2 only has constraints for error checking and adjective endings. If we compare only those contraints which apply to adjective endings, we produce Figure 5.4.

The curves are now much more similar. The fits are both poor, although the fit for Tutor 1 is somewhat better at $R^2 = 0.6303$ vs $R^2 = 0.4498$. The slopes are also very similar, with Tutor 1 having a learning rate of 0.2347, and Tutor 2 having a learning rate of 0.214.

To try to determine what the students were learning from Tutor 1, if they were not learning the constraints for adjective endings, we also calculated the learning curve of the supplementary constraints for Tutor 1 - those that check the gender, case and article. See Figure 5.5 for this graph. The fit is quite good,

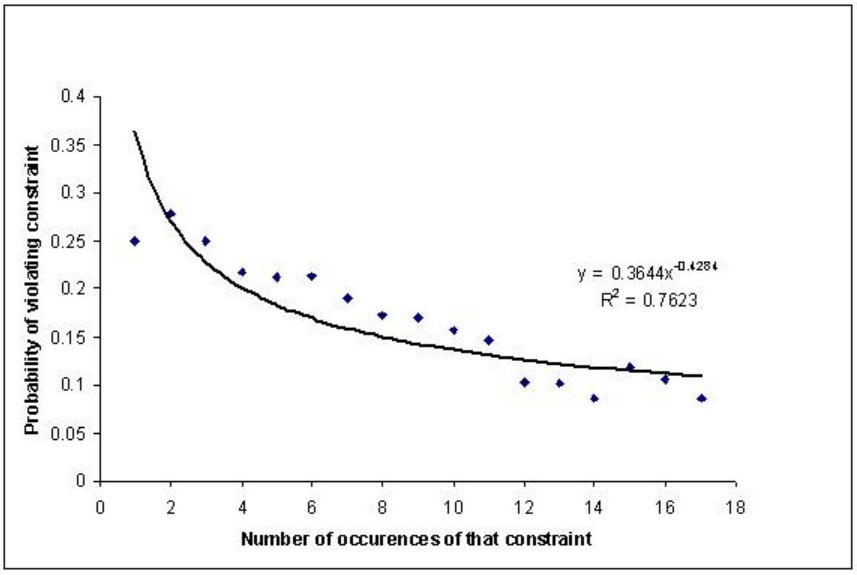

Figure 5.1: Tutor 1, all constraints

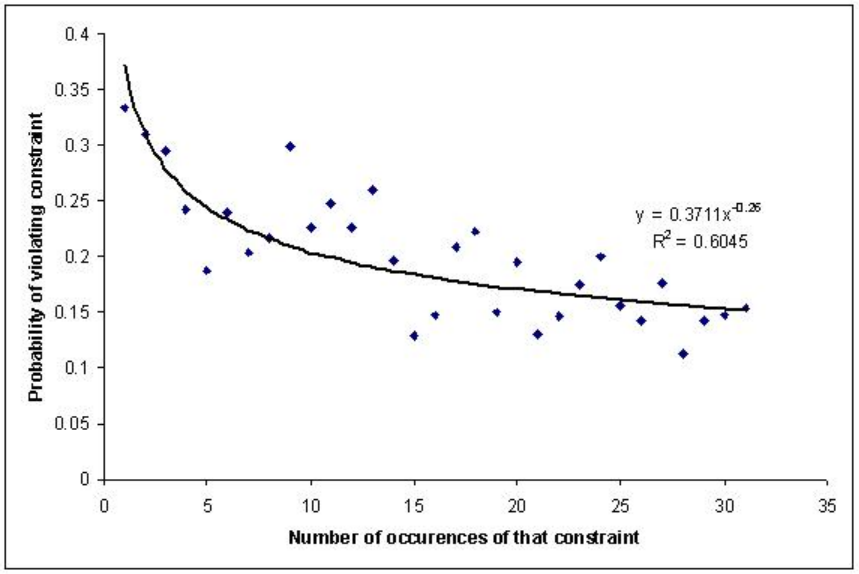

Figure 5.2: Tutor 2, all constraints

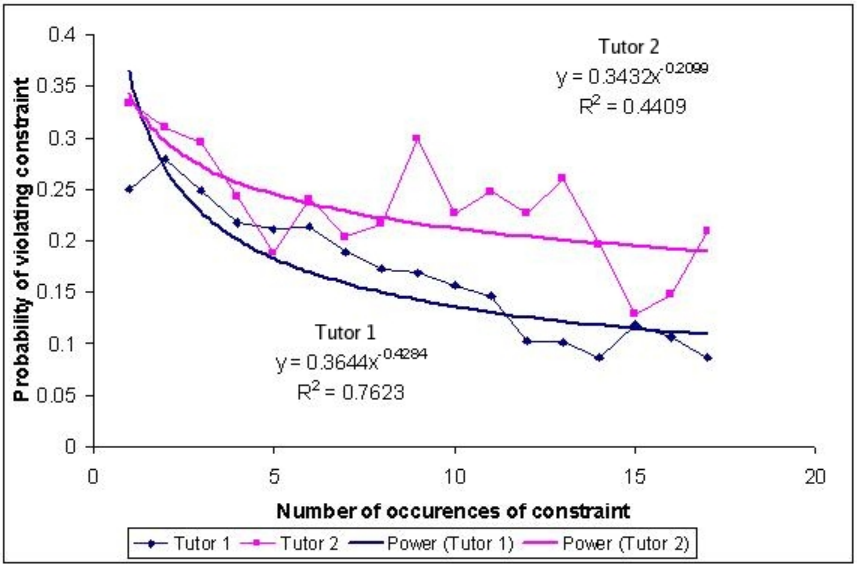

Figure 5.3: Tutor 1 and Tutor 2, all constraints

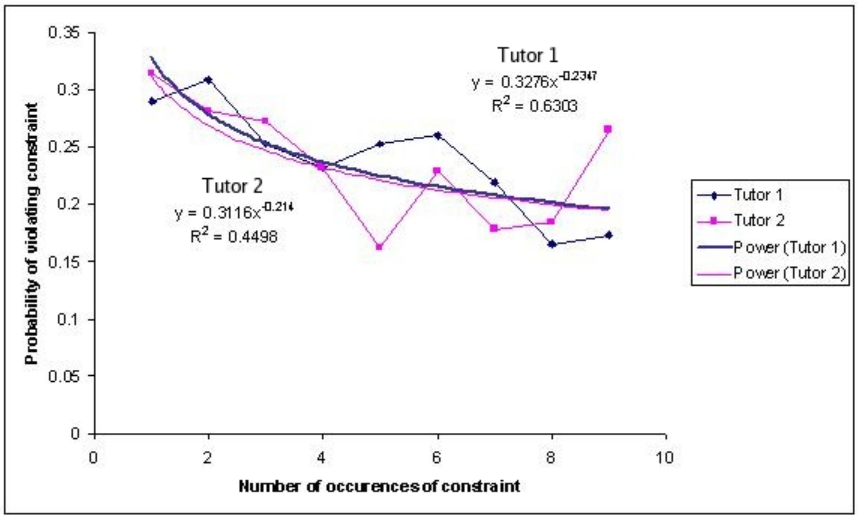

Figure 5.4: Tutor 1 and Tutor 2, all shared constraints

with $R^2 = 0.7102$, and the learning rate is also good at 0.3171. Clearly, students were learning to identify the correct gender, case and article.

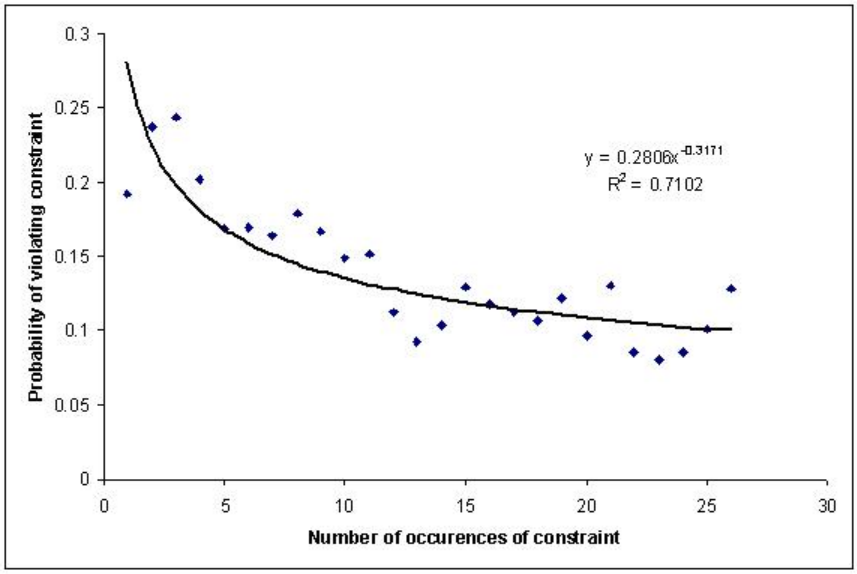

Figure 5.5: Tutor 1, all extra constraints

An unexpected result can be seen in Figure 5.6. This is the learning curve for all constraints related to gender. While the fit is not good, with $R^2 = 0.5$, the slope is better, with a value of 0.444. This result is interesting, as we had expected that students would not learn the gender of the nouns. One possible explanation is that students became better at reading the clues to the gender of the noun (ie from the article).

When students specified the gender incorrectly, the feedback message they were given was "Oops - wrong gender. The gender for each noun must be learnt individually, but the article can sometimes provide a clue." It is interesting that this actually helped them to learn the concept.

5.2.3 Anecdotal Results

Twelve of the twenty-three students who took part in the study took the opportunity of providing feedback. The feedback was generally positive.

> "It was good that the mistakes were explained + the grammer rules were also explained."
>
> "I liked it and found very useful"
>
> "I think it was a good/useful learning tool."
>
> "I like exercises on the computer more than written ones as they have the answers there so you can do it on your own + I felt my confident getting better as I progressed through them"

The main complaint from students was that some of the feedback messages were unclear.

> "Comments were not as clear as they could be."
>
> "...the explanations make sense (except the = ... until ein word change = bit, but that may be because I didn't listen in class)."
>
> "Even if I use 'What mistakes did I make' I can't understand explanation and can't understand my mistakes."

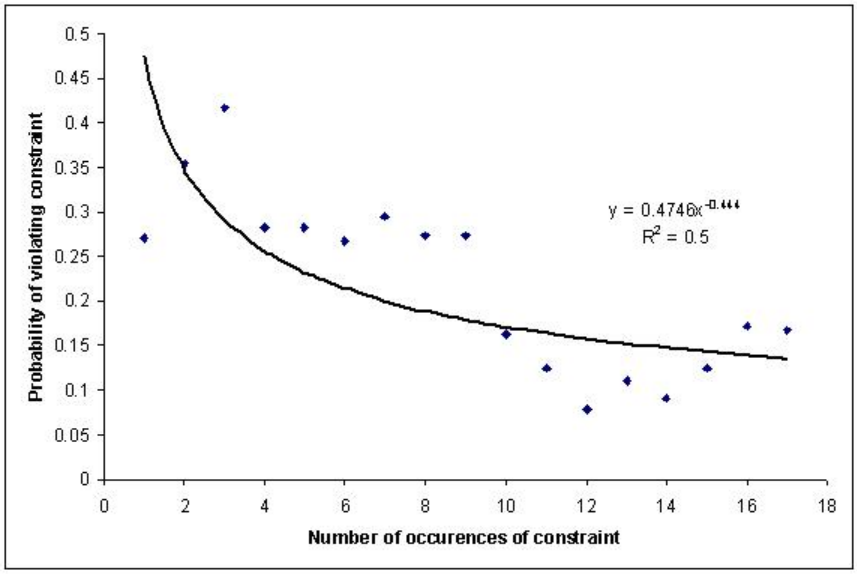

Figure 5.6: Tutor 1, constraints for gender

Many students also provided advice on how to improve the system.

> "I think that incorrect letter can be in red colour."

> "The ? mark confused me at first. Maybe put it like this [?] or just a blank square or a line _____ ."

> "Maybe a button to press to show just the gender of the noun."

> "When you get it right you should see gender/case/article."

Only two students mentioned the differences between the two systems. The student who wrote the first comment used Tutor 1, while the second student used Tutor 2. Interestingly, they both preferred Tutor 1.

> "The logical process of establishing gender, case and article before doing the answer is helpful"

> "I think at this point the first tutor (with cases and such) looks great (and better for me personally lol)."

Feedback from the lecturers was positive. The German department has expressed a desire to have the system permanently available for their students.

> "Having marked our end of week quiz, I really feel that the marks on adjective endings have improved - thanks to your programme!"

> "The session was a great help for the students and i can see that they have done really well in the quiz we had the following day. I am amazed how well it works for adjective endings!!!"

6 Discussion and Future Work

This chapter discusses the results presented in the previous chapter, and what these results imply. Limitations of the research are considered, and future work is suggested.

6.1 Discussion of Results

The original hypotheses were that:

1. Requiring the student to specify additional information about the task will not degrade learning.

2. Requiring the student to specify additional information about the task will improve learning.

Hypothesis 1 has been borne out. On the post-test, students who used Tutor 1 performed at approximately the same level as those who used Tutor 2. There is a possibility that they performed better, however because of the difference in difficulty of the two tests, this cannot be stated for certain. Students who used Tutor 1 also learned the constraints pertaining to adjective endings at a similar rate to the students using Tutor 2. This can been seen in Figure 5.4.

Hypothesis 2 has not been proven; it has also not been disproven. Students learning from Tutor 1 did learn more than students learning from Tutor 2, as can be seen from Figure 5.3. However, the improved learning was not in the area the tutor was intended to teach. The tutor was intended to teach students to specify adjective endings correctly. Although they can perform that task equally well as students who used Tutor 2, they cannot perform it better, so specifying additional information has not improved their ability with adjective endings.

Students using Tutor 2 have improved in other areas. They became more skilled at determining the gender and case of a noun, and the type of article. This is illustrated both by the pre- and post-test results, as seen in Tables 5.5 and 5.6, and also by the learning curves for all constraints (Figure 5.1) and for the extra constraints (Figure 5.5).

It is open to debate whether teaching students something unintended is a positive or a negative occurence. The extra skills the students learn do relate to the domain at hand, and perhaps with more instruction, they could combine their news skills together with the correct information, and master adjective endings. Conversely, learning these constraints while using the system could have distracted the students from learning what was really important.

6.2 Limitations

There were a number of limitations in the evaluation study. The number of students was very small, with only 23 students total. The time the students spent using the system was also far from ideal. The average length of time in the tutor was half an hour, as the students only had one class period in which to use the tutor, and complete the pre- and post-tests.

The difference in the level of the pre- and post-test was very harmful. A possible way to avoid having tests of different difficulties would be to determine what skill each question should test, and then fill in the questions with vocabulary of approximately equal difficulty. This would hopefully lead to a more even pair of tests.

Results might have been clearer if the system was set up to allow only one feedback level, which displays all the current messages. This would avoid possible differences caused by some students not using the feedback.

The different number of constraints in the two systems made comparing the learning curves of the two systems difficult. The smaller number of constraints for Tutor 2 affected the curve by making it rougher. The different number of constraints also made direct comparisons difficult, especially because one constraint in Tutor 2 was represented by many in Tutor 1.

6.3 Future Work

A number of improvements could be made to the software. The ideas suggested by the students are valid, such as having a button to display the gender of the noun, and improving the wording of some of the feedback messages. More problems could also be added, and problem selection could be made adaptive.

A second study would be interesting, with more students, or carried out over more time. It would also be interesting to see if the level of the student affected which tutor was more appropriate. It is conceivable that more advanced students would find the scaffolding of Tutor 1 restrictive and pointless.

This tutor could be extended to cover different areas of German grammar, and then the new implementations could be tested. Further systems in other domains could also be constructed to compare the effect of ambiguity vs transference.

Further analysis of the results produced from this evaluation is also possible. It would be worth exploring whether students using Tutor 1 learnt more per problem. As student using Tutor 1 completed on average only 21 problems vs 49 for Tutor 2, it is plausible that although both group learnt the same amount about adjective endings, students using Tutor 1 learnt more on average per problem.

7 Conclusion

The problem of ambiguity in diagnosing errors in student answers is a difficult one. Decreasing ambiguity leads to a less realistic problem solving environment.

In this research, we presented two tutors for the same domain. Tutor 1 required the student to specify all variables, while Tutor 2 only required the student to specify the adjective ending. These tutors were tested in a first-year German class, during a fifty minute period.

Students using Tutor 1 had increased learning, however the domain principles that they learnt were not the ones the system intended to teach them. Students learnt to identify the correct case, gender and article, but their ability to provide the correct adjective ending remained constant with that of the students using Tutor 2.

Further research is needed to determine whether the improved feedback gained by reducing ambiguity does increase learning. This study suffered from a small number of students, and a small length of exposure to the system. It would be interesting to see if a difference does appear between the two styles of system after a longer period of exposure.

Acknowledgments

First of all I would like to thank those most closely involved in this research - my supervisor Brent, and the lecturers and students in the German department for all their time and support, especially Linda, Gudrun and Susanne. I would also like to thank Tanja for being my supervisor for the first half of the year.

Thanks must go to Kon for his amazing program that generates learning curves out of anything. Moffat for helping me out whenever I needed it. Nancy, Jay and everyone in the ICTG lab for their entertainment and support.

Thanks to Ollie and Sascha for checking my German for me. Danke!

Jason for this style sheet, all of the postgrads for keeping me amused with their lunchtime conversation, and fed with their cake.

Nick and Ben for not kicking me out of the flat when I didn't pull my weight, Graham for looking after me when I'm too tired to look after myself, my family for telling me to 'just keep swimming'.

Finally, Marcus, Kiyo, Sarah, Moffat, Brendon, Steve, Leighton, Cong.... it's been a crazy year, but at least it's never been boring!

Bibliography

[1] ANDERSON, J. R., CORBETT, A. T., KOEDINGER, K. R., AND PELLETIER, R. Cognitive Tutors: Lessons Learned. *The Journal of the Learning Sciences 4*, 2 (1995), 167–209.

[2] BECK, J., STERN, M., AND HAUGSJAA, E. Applications of AI in education. *ACM Crossroads 3*, 1 (1996).

[3] DREYER, H., AND SCHMITT, R. *A practice grammar of German*. Max Hueber Verlag, 2001.

[4] FRANZ INC. Franz Inc: Allegro Common Lisp and Common Lisp Products. `http://www.franz.com/`, 2004.

[5] JOHN FODERARO, FRANZ INC. AllegroServe—a Web Application Server. `http://opensource.franz.com/aserve/index.html`, 2004.

[6] JUOZULYNAS, V. Errors in the compositions of second-year german students: An empirical study for parser-based icali. *CALICO Journal 12*, 1 (1994), 5–17.

[7] KAHLEN, L. *Interactive German Grammar Made Easy*. McGraw-Hill, 2006.

[8] MARTIN, B., KOEDINGER, K. R., MITROVIĆ, A., AND MATHAN, S. On Using Learning Curves to Evaluate ITS. In *Proceedings of the 12th International Conference on Artificial Intelligence in Education* (2005), C.-K. Looi, G. McCalla, B. Bredeweg, and J. Breuker, Eds., IOS Press, pp. 419–426.

[9] MARTIN, B., AND MITROVIC, A. Authoring web-based tutoring system with WETAS. In *ICCE* (2002), pp. 183–187.

[10] MARTIN, B., AND MITROVIC, A. WETAS: A web-based authoring system for constraint-based ITS. In *Adaptive Hypermedia and Adaptive Web-Based Systems, Second International Conference, AH 2002, Malaga, Spain, May 29-31, 2002, Proceedings* (2002), P. D. Bra, P. Brusilovsky, and R. Conejo, Eds., vol. 2347 of *Lecture Notes in Computer Science*, Springer, pp. 543–546. Available from: `http://link.springer.de/link/service/series/0558/bibs/2347/23470543.htm`.

[11] MARTIN, B., AND MITROVIĆ, A. A. Domain Modeling: Art or Science? In *Proceedings of the 11th International Conference on Artificial Intelligence in Education* (2003), U. Hoppe, F. Verdejo, and J. Kay, Eds., IOS Press, pp. 183–190.

[12] MARTIN, B. I. *Intelligent Tutoring Systems: The Practical Implementation of Constraint-based Modelling*. PhD thesis, University of Canterbury, 2001.

[13] MENZEL, W. Constraint-based modeling and ambiguity. *International Journal of Artificial Intelligence in Education 16*, 1 (2006), 29–63.

[14] MITROVIĆ, A., MAYO, M., SURAWEERA, P., AND MARTIN, B. Constraint-based Tutors: a Success Story. In *Proceedings of the 14thInternational Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems* (Berlin, 2001), L. Monostori, J. Váncza, and M. Ali, Eds., vol. 2070, Springer-Verlag GmbH, pp. 931–940.

[15] NEWELL, A., AND ROSENBLOOM, P. S. Mechanisms of Skill Acquisition and the Law of Practice. In *The Soar Papers: Research on Integrated Intelligence*, vol. 1. MIT Press, 1993, pp. 81–135.

[16] OHLSSON, S. Constraint-Based Student Modelling. In *Student Modelling: The Key to Individualized Knowledge-based Instruction* (Berlin, 1994), J. E. Greer and G. I. McCalla, Eds., vol. 125 of *Computer Systems and Sciences*, Springer-Verlag GmbH, pp. 167–189.

[17] OHLSSON, S. Learning From Performance Errors. *Psychological Review 103*, 2 (Apr. 1996), 241–262. Available from: http://www.sciencedirect.com/science/article/B6X04-46SGF70-2/2/3da274890b14074e25bab866d40b7f40.

[18] OHLSSON, S. Learning and Instruction: An Introduction. COSC420 Course Reading Material, July 2004.

[19] ROGERS, M. On major types of written error in advanced students of german. *International Review of Applied Linguistics in Language Teaching 22*, 1 (Feb. 1984), 1–39.

[20] SPARKS, K., AND VAIL, V. H. *German in review*. Harcourt Brace Jovanovich, 1986.

[21] TERRELL, T. D., TSCHIRNER, E., AND NIKOLAI, B. *Kontakte : a communicative approach*. McGraw-Hill, 2004.

[22] WEBSTER, P. *Schwarz, rot, gold : the German handbook : a practical grammar guide*. Cambridge University Press, 1987.

[23] WERNER, G. *Langenscheidts Grammatik-training Deutsch*. Langenscheidt KG, 2001.

[24] ZAKHAROV, K., MITROVIĆ, A., AND OHLSSON, S. Feedback Micro-engineering in EER-Tutor. In *Proceedings of the 12th International Conference on Artificial Intelligence in Education* (July 2005), C.-K. Looi, G. McCalla, B. Bredeweg, and J. Breuker, Eds., vol. 125 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, pp. 718–725.

A Student Feedback

All feedback has been transcribed exactly as the student wrote it.

A.1 Tutor 1

”Comments were not as clear as they could be. Pretty decent.” - mje72

”The system is good, it tells you what mistakes you’ve made and the explanations make sense (except the = ... until ein word change = bit, but that may be because I didn’t listen in class).” - hkh15

”The ? mark confused me at first. Maybe put it like this [?] or just a blank square or a line _____ . ”Maybe a button to press to show just the gender of the noun.” - kha50

”The logical process of establishing gender, case and article before doing the answer is helpful, although I don’t usually have problems with this personally. It took me a while to realise I should use the ’hint’ option when I was stuck, but it was really useful once I remembered to use it.” - hse22

”All good =) Thanks for helping me practise adjectival endings!!” - lrg21

A.2 Tutor 2

”It was good that the mistakes were explained + the grammer rules were also explained.” - hlm36

”I like exercises on the computer more than written ones as they have the answers there so you can do it on your own + I felt my confident getting better as I progressed through them” - jmw225

”I liked it and found very useful” - aya21

”The coding gave rise to a couple of interesting results - a mispelt adjective led to several mistakes, wih the omitted letter as the root - even though the ending itself was right. :) ”I liked the ’what did I do wrong’ section - perhaps if I try this some more the rules will eventually stick in my head! ”However after one use of the tutor I think I forgot which case is which again... ”I think at this point the first tutor (with cases and such) looks great (and better for me personally lol).” - mdo38

”Even if I use ’What mistakes did I make’ I can’t understand explanation and can’t understand my mistakes. ”I think that incorrect letter can be in red colour.” - omu10

”When you get it right you should see gender/case/article.” - hmz10

”I think it was a good/useful learning tool.” - mrm67 (same as mrm72?)

B Pre- and Post-tests

B.1 Test 1

Complete the sentence with the given adjective.

Das _____ Kleid ist zu kurz. (weiss)

Ich gehe in einen _____ Park. (gross)

Ich kaufe mir eine _____ Wohnung. (neu)

Complete the sentence with the given adjective, and fill in the information about the noun.

Ich mag _____ Sonne. (warm)

Case _____

Gender _____

Article _____

Ich fahre gern mit diesem _____ Auto. (schnell)

Case _____

Gender _____

Article _____

Die _____ Frau ist meine Tante. (klein)

Case _____

Gender _____

Article _____

B.2 Test 2

```
Complete the sentence with the given adjective.

Das _____________ Auto ist schnell.  (rot)

In diesem _____________ Haus wohnt meine Freundin.  (alt)

Wir lesen ein _____________ Buch.  (spannend)


Complete the sentence with the given adjective, and fill in the information
about the noun.

Ich mache mir __________ Schokolade. (heiss)

Case _____________

Gender _____________

Article  _____________


Ich sehe den _________ Mann.  (alt)

Case _____________

Gender _____________

Article  _____________


Ein _____________ Tee schmeckt wunderbar. (gut)

Case _____________

Gender _____________

Article  _____________
```

B.3 Post-test

The post-test also contained a question asking for feedback.

> Feel free to write any comments or criticisms about the system on the reverse of this page what did you like? What did you dislike? What did you find was useful, or not useful?

C Constraints

C.1 Tutor 1

```
(50
 "Oops - wrong case. The nominative case is used for the subject of
    a sentence, the noun that carries out the verb."
 (and
  (match IS CASE ("N"))
  (match IS CASE (?case))
  (match SS CASE (?case2))
  )
 (test SS ((?case) ?case2))
"CASE")

(51
 "Oops - wrong case. The accusative case is used for the object of a
    sentence, and after some prepositions."
 (and
  (match IS CASE ("A"))
  (match IS CASE (?case))
  (match SS CASE (?case2))
  )
 (test SS ((?case) ?case2))
 "CASE")

(52
 "Oops - wrong case. The dative case is used for the indirect object,
    and after some prepositions."
 (and
  (match IS CASE ("D"))
  (match IS CASE (?case))
  (match SS CASE (?case2))
  )
 (test SS ((?case) ?case2))
"CASE")

(53
 "You haven't specified a case."
 (match IS CASE (?case2))
 (match SS CASE (?case))
 "CASE")

(54
 "Oops - wrong gender. The gender for each noun must be learnt
```

```
    individually, but the article can sometimes provide a clue."
 (and
  (match IS GENDER ("M"))
  (match IS GENDER (?gender))
  (match SS GENDER (?gender2))
  )
 (test SS ((?gender) ?gender2))
 "GENDER")

(55
 "Oops - wrong gender. The gender for each noun must be learnt
    individually, but the article can sometimes provide a clue."
 (and
  (match IS GENDER ("F"))
  (match IS GENDER (?gender))
  (match SS GENDER (?gender2))
  )
 (test SS ((?gender) ?gender2))
 "GENDER")

(56
 "Oops - wrong gender. The gender for each noun must be learnt
    individually, but the article can sometimes provide a clue."
 (and
  (match IS GENDER ("N"))
  (match IS GENDER (?gender))
  (match SS GENDER (?gender2))
  )
 (test SS ((?gender) ?gender2))
 "GENDER")

(57
 "Oops - wrong gender. The gender for each noun must be learnt
    individually, but the article can sometimes provide a clue."
 (and
  (match IS GENDER ("P"))
  (match IS GENDER (?gender))
  (match SS GENDER (?gender2))
  )
 (test SS ((?gender) ?gender2))
 "GENDER")

(58
 "You haven't specified a gender."
 (match IS GENDER (?gender))
 (match SS GENDER (?gender2))
 "GENDER")

(59
 "Oops - wrong article. Ein-words are ein/eine/etc, but also kein,
    mein, dein, sein, ihr, unser, euer, ihr and Ihr."
 (and
  (match IS ARTICLE ("I"))
  (match IS ARTICLE (?a))
```

```
  (match SS ARTICLE (?a2))
  )
 (test SS ((?a) ?a2))
 "ARTICLE")

(60
 "Oops - wrong article. Der-words are der/die/das/etc, but also dieser,
    jeder, solcher, welcher, alle, and jener."
 (and
  (match IS ARTICLE ("D"))
  (match IS ARTICLE (?a))
  (match SS ARTICLE (?a2))
  )
 (test SS ((?a) ?a2))
 "ARTICLE")

(61
 "Oops - wrong article. There is no article when the adjective stands
    alone before the noun.  There is an article here."

 (and
  (match IS ARTICLE ("N"))
  (match IS ARTICLE (?a))
  (match SS ARTICLE (?a2))
  )
 (test SS ((?a) ?a2))
 "ARTICLE")

(70
 "You haven't specified an article."
 (match IS ARTICLE (?a))
 (match SS ARTICLE (?a1))
 "ARTICLE")

(4
 "You have spelt the adjective incorrectly."
 (and
  (match IS ANSWER (?*w1 ?letter ?letter2))
  (match SS ANSWER (?l ?*))
  )
 (match SS ANSWER (?*w1 ?*))
 "ANSWER")

(5
 "You have unnecessary letters in your answer."
 (and
  (match IS ANSWER (?*w1 ?letter ?letter2))
  (match SS ANSWER (?*w1 ?*))
  (match SS ANSWER (?* ?letter2))
  )
 (match SS ANSWER (?*w1 ?letter ?letter2))
 "ANSWER")

(10
```

```
 "If an adjective is preceded by a 'der-word', it must end in either
    -e or -en."
 (and
  (match SS ARTICLE ("D"))
  (match SS ANSWER (?something ?*))
  )
 (or-p
  (match SS ANSWER (?*w2 "e" "n"))
  (match SS ANSWER (?*w1 "e"))
  )
 "ANSWER")

(11
 "If an adjective is plural, and preceded by a der or ein word, it must
    take the ending -en."
 (and
  (or-p
   (and
    (or-p
     (match SS ANSWER (?*w2 "e" "n"))
     (match SS ANSWER (?*w1 "e"))
     )
    (match SS ARTICLE ("D"))
    )
   (and
    (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
     )
    (match SS ARTICLE ("I"))
    )
   )
  (match SS GENDER ("P"))
  (match SS ANSWER (?something ?*))

  )
 (match SS ANSWER (?*w1 "e" "n"))
 "ANSWER")

(12
 "If a noun is nominative and singular, and preceded by a 'der-word',
    it must end in -e, no matter what the gender."
 (and
  (match SS ARTICLE ("D"))
  (not-p (match SS GENDER ("P")))
  (match SS CASE ("N"))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?*w2 "e" "n"))
  (match SS ANSWER (?*w1 "e"))
  )
  )
```

```
 (match SS ANSWER (?*w1 "e"))
 "ANSWER")

(13
 "If a noun is in the dative case, and preceded by 'der-' or
    'ein-word', the ending on the adjective is -en."
 (and
  (or-p
   (and
    (or-p
     (match SS ANSWER (?*w2 "e" "n"))
     (match SS ANSWER (?*w1 "e"))
     )
    (match SS ARTICLE ("D"))
    )
   (and
    (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
     )
    (match SS ARTICLE ("I"))
    )
   )
  (match SS CASE ("D"))
  (match SS ANSWER (?something ?*))
  )
 (match SS ANSWER (?* "e" "n"))
 "ANSWER")

(14
 "If a noun is singular and preceded by a 'der-word', the adjective
    ends in -e until the 'der-word' changes.  After this the
    adjective ends in -en."
 (and
  (not-p (match SS GENDER ("P")))
  (match SS ARTICLE ("D"))
  (match SS ANSWER (?something ?*))
  (or-p
   (match SS ANSWER (?*w2 "e" "n"))
   (match SS ANSWER (?*w3 "e"))
   )
  (or-p
   (match SS CASE ("N"))
   (and
    (match SS CASE ("A"))
    (or-p
     (match SS GENDER ("F"))
     (match SS GENDER ("N"))
     )
    )
   )
  )
```

```
 (match SS ANSWER (?* "e"))
 "ANSWER")

(15
 "If a noun is singular and preceded by a 'der-word', the adjective
    ends in -e until the 'der-word' changes.  After this the
    adjective ends in -en."
 (and
  (not-p (match SS GENDER ("P")))
  (match SS ARTICLE ("D"))
  (match SS ANSWER (?something ?*))
  (or-p
   (match SS ANSWER (?*w2 "e" "n"))
   (match SS ANSWER (?*w3 "e"))
   )
  (or-p
   (match SS CASE ("D"))
   (and
    (match SS CASE ("A"))
    (match SS GENDER ("M"))
    )
   )
  )
 (match SS ANSWER (?* "e" "n"))
 "ANSWER")

(21
 "If the article is an 'ein-word', the only possible endings are
    -e, -en, -er and -es."
 (and
  (match SS ARTICLE ("I"))
  (match SS ANSWER (?something ?*))
  )
 (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
 "ANSWER")

(22
 "If it follows 'ein-word', before a singular noun, the adjective
    takes the ending the 'der-word' would have had, until the ein-word
    changes.  After this the adjective ends in -en."
 (and
  (match SS ARTICLE ("I"))
  (not-p (match SS GENDER ("P")))
  (match SS CASE ("D"))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
```

```
  (match SS ANSWER (?* "e" "s"))
  )
  )
 (match SS ANSWER (?* "e" "n"))
 "ANSWER")

(23
 "If it follows 'ein-word', before a singular noun, the adjective takes
    the ending the 'der-word' would have had, until the ein-word changes.
    After this the adjective ends in -en."
 (and
  (match SS ARTICLE ("I"))
  (or-p
   (match SS CASE ("N"))
   (match SS CASE ("A"))
   )
  (match SS GENDER ("F"))
  (match SS ANSWER (?something ?*))
  (or-p
   (match SS ANSWER (?* "e"))
   (match SS ANSWER (?* "e" "n"))
   (match SS ANSWER (?* "e" "r"))
   (match SS ANSWER (?* "e" "s"))
  )
  )
 (match SS ANSWER (?* "e"))
 "ANSWER")

(24
 "If it follows 'ein-word', before a singular noun, the adjective takes
    the ending the 'der-word' would have had, until the ein-word changes.
    After this the adjective ends in -en."
 (and
  (match SS ARTICLE ("I"))
  (or-p
   (match SS CASE ("N"))
   (match SS CASE ("A"))
   )
  (match SS GENDER ("N"))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
  )
 (match SS ANSWER (?* "e" "s"))
 "ANSWER")

(25
 "If it follows 'ein-word', before a singular noun, the adjective takes the
    ending the 'der-word' would have had, until the ein-word changes.  After
    this the adjective ends in -en."
```

```
 (and
  (match SS ARTICLE ("I"))
  (match SS CASE ("N"))
  (match SS GENDER ("M"))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
  )
 (match SS ANSWER (?* "e" "r"))
 "ANSWER")

(26
 "If it follows 'ein-word', before a singular noun, the adjective takes the
    ending the 'der-word' would have had, until the ein-word changes.  After
    this the adjective ends in -en."
 (and
  (match SS ARTICLE ("I"))
  (match SS CASE ("A"))
  (match SS GENDER ("M"))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
  )
 (match SS ANSWER (?* "e" "n"))
 "ANSWER")

(30
 "If there is no ein-word or der-word, the adjective takes the ending that
    the der-word would have if it was present."
 (and
  (match SS ARTICLE ("N"))
  (match SS ANSWER (?something ?*))
  (or-p
   (match SS GENDER ("F"))
   (match SS GENDER ("P"))
   )
  (or-p
   (match SS CASE ("N"))
   (match SS CASE ("A"))
   )

  )
 (match SS ANSWER (?* "e"))
 "ANSWER")

 (31
```

```
"If there is no ein-word or der-word, the adjective takes the ending that
   the der-word would have if it was present."
(and
 (match SS ARTICLE ("N"))
 (match SS ANSWER (?something ?*))
 (match SS GENDER ("N"))
 (or-p
  (match SS CASE ("N"))
  (match SS CASE ("A"))
  )
 )
(match SS ANSWER (?* "e" "s"))
 "ANSWER")

(32
"If there is no ein-word or der-word, the adjective takes the ending that
   the der-word would have if it was present."
(and
 (match SS ARTICLE ("N"))
 (match SS ANSWER (?something ?*))
 (match SS CASE ("D"))
 (or-p
  (match SS GENDER ("M"))
  (match SS GENDER ("N"))
  )
 )
(match SS ANSWER (?* "e" "m"))
 "ANSWER")

(33
"If there is no ein-word or der-word, the adjective takes the ending that
   the der-word would have if it was present."
(and
 (match SS ARTICLE ("N"))
 (match SS ANSWER (?something ?*))
 (or-p
  (and
   (match SS CASE ("N"))
   (match SS GENDER ("M"))
   )
  (and
   (match SS CASE ("D"))
   (match SS GENDER ("F"))
   )
  )
 )
(match SS ANSWER (?* "e" "r"))
 "ANSWER")

(34
"If there is no ein-word or der-word, the adjective takes the ending that
   the der-word would have if it was present."
(and
 (match SS ARTICLE ("N"))
```

```
 (match SS ANSWER (?something ?*))
 (or-p
  (and
   (match SS CASE ("A"))
   (match SS GENDER ("M"))
   )
  (and
   (match SS CASE ("D"))
   (match SS GENDER ("P"))
   )
  )
 )
(match SS ANSWER (?* "e" "n"))
 "ANSWER")

 (99
  "You haven't answered the question!"
  (match IS ANSWER (?what ?*))
  (match SS ANSWER (?what2 ?*))
  "ANSWER")
```

C.2 Tutor 2

```
(4
         "You have spelt the adjective incorrectly."
         (and
          (match IS ANSWER (?*w1 ?letter ?letter2))
          (match SS ANSWER (?l ?*))
          )
         (match SS ANSWER (?*w1 ?*))
         "ANSWER")

        (5
         "You have unnecessary letters in your answer."
         (and
          (match IS ANSWER (?*w1 ?letter ?letter2))
          (match SS ANSWER (?*w1 ?*))
          (match SS ANSWER (?* ?letter2))
          )
         (match SS ANSWER (?*w1 ?letter ?letter2))
         "ANSWER")

        (10
         "When they are preceded by a 'der-word', all adjectives end in either
            -e or -en."
         (and
          (match IS ARTICLE ("D"))
          (match SS ANSWER (?something ?*))
          )
         (or-p
          (match SS ANSWER (?*w2 "e" "n"))
          (match SS ANSWER (?*w1 "e"))
          )
```

```
 "ANSWER")

(11
 "All adjectives take the ending -en when the noun is plural, and there
    is an ein-word or der-word."
 (and
  (or-p
   (and
    (or-p
     (match SS ANSWER (?*w2 "e" "n"))
     (match SS ANSWER (?*w1 "e"))
     )
    (match IS ARTICLE ("D"))
    )
   (and
    (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
     )
    (match IS ARTICLE ("I"))
    )
   )
  (match IS GENDER ("P"))
  (match SS ANSWER (?something ?*))

  )
 (match SS ANSWER (?*w1 "e" "n"))
 "ANSWER")

(12
 "When the noun is nominative and singular, and preceded by a 'der-word',
    all adjectives end in -e, no matter what the gender."
 (and
  (match IS ARTICLE ("D"))
  (not-p (match IS GENDER ("P")))
  (match IS CASE ("N"))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?*w2 "e" "n"))
  (match SS ANSWER (?*w1 "e"))
  )
  )
 (match SS ANSWER (?*w1 "e"))
 "ANSWER")

(13
 "When the noun is in the dative case, and preceded by 'der-' or
    'ein-word', the ending on the adjective is always -en."
 (and
  (or-p
   (and
    (or-p
```

```
     (match SS ANSWER (?*w2 "e" "n"))
     (match SS ANSWER (?*w1 "e"))
     )
    (match IS ARTICLE ("D"))
    )
   (and
    (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
     )
    (match IS ARTICLE ("I"))
    )
   )
  (match IS CASE ("D"))
  (match SS ANSWER (?something ?*))
  )
 (match SS ANSWER (?* "e" "n"))
 "ANSWER")

(14
 "If the noun is singular and preceded by a 'der-word', the adjective
    ends in -e until the 'der-word' changes.  After this the adjective
    ends in -en."
 (and
  (not-p (match IS GENDER ("P")))
  (match IS ARTICLE ("D"))
  (match IS ANSWER (?*w1))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?*w2 "e" "n"))
  (match SS ANSWER (?*w1 "e"))
  )
  )
 (match SS ANSWER (?*w1))
 "ANSWER")

(21
 "When the article is an 'ein-word', the only possible endings are
    -e, -en, -er and -es."
 (and
  (match IS ARTICLE ("I"))
  (match SS ANSWER (?something ?*))
  )
 (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
 "ANSWER")

(22
```

```
 "When following an 'ein-word', before a singular noun, the adjective
    takes the ending the 'der-word' would have had, until the ein-word
    changes.  After this the adjective ends in -en."
 (and
  (match IS ARTICLE ("I"))
  (not-p (match IS GENDER ("P")))
  (match IS ANSWER (?* ?second ?last))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
  )
 (or-p
  (and
   (match SS ANSWER (?* "e"))
   (match SS ANSWER (?* ?last))
   )
  (match SS ANSWER (?* ?second ?last))
  )
 "ANSWER")

(23
 "When the noun is nominative or accusative and singular, and the article
    is an 'ein-word', the adjective takes the ending the der-word
    would have had."
 (and
  (match IS ARTICLE ("I"))
  (or-p
   (match IS CASE ("N"))
   (match IS CASE ("A"))
   )
  (not-p (match IS GENDER ("P")))
  (match IS ANSWER (?* ?second ?last))
  (match SS ANSWER (?something ?*))
  (or-p
  (match SS ANSWER (?* "e"))
  (match SS ANSWER (?* "e" "n"))
  (match SS ANSWER (?* "e" "r"))
  (match SS ANSWER (?* "e" "s"))
  )
  )
 (or-p
  (and
   (match SS ANSWER (?* "e"))
   (match SS ANSWER (?* ?last))
   )
  (match SS ANSWER (?* ?second ?last))
  )
 "ANSWER")

(30
```

```
"When there is no ein-word or der-word, the adjective takes the ending
   that the der-word would have if it was present."
(and
 (match IS ARTICLE ("N"))
 (match SS ANSWER (?something ?*))
 (match IS ANSWER (?*w1 ?second ?last))
 )
(match SS ANSWER (?*w2 ?second ?last))
"ANSWER")

 (99
  "You haven't answered the question!"
  (match IS ANSWER (?what ?*))
  (match SS ANSWER (?what2 ?*))
  "ANSWER")
```