

Computer support for second language learners' free text production – Initial studies

Ola Knutsson, Teresa Cerratto Pargman and Kerstin Severinson Eklundh

IPLab, Department of Numerical Analysis and Computer Science

Royal Institute of Technology

Stockholm, Sweden

Keywords: second language learning, learner-centered design, language technology, writing documents, reviewing process.

Abstract

This paper discusses the role of computer-based language tools for adult writers in the context of second language learning. It presents initial studies of second-language learners who write with the aid of a language tool developed at our department. The studies focus on the revision process in authentic free text production. A method for collecting data from free text production has been developed and explored. In the study, the users seemed to be supported by the language tool, and they followed the advice from program when appropriate feedback was given. The method used seems to be adequate both in pointing out the directions for an extended study, and also in pointing on which parts of the language tool that must be further developed and improved. As a complement to the study of the writers, interviews with teachers have been conducted. There is an overlap between the errors the teachers said were important and the errors the language tools searches for.

1 Introduction

The interest in the acquisition and development of a second language is not a new subject of research. In Sweden, for instance, the study of acquisition and development of Swedish as a second language is a vast area of research that develops from the early 70's [1, 2, 3]. The studies have often been conducted from linguistic, pedagogical and socio-linguistic perspectives. They agree on viewing the acquisition and development of Swedish as a second language as a multifaceted process in which it is necessary to combine different foci [4]. They have, however, most often focused on the study of speech and the development of second language learners communicative competence [5, 6, 7]. Questions regarding the role of writing during the acquisition and development of a second language have usually been overlooked. One exception is the work conducted by Gunilla Jansson [8] on writing strategies developed by writers who have Swedish as a second language. Our interest in writing relies on the central place that writing occupies in the development of language and thinking processes ([9, 10], Luria, 1946 cited by Downing [11]). *“Cognitive processes and structures are transformed significantly by the acquisition of our best recognized cultural (and intellectual) tool, namely, writing”* ([12] p. 96.). Both Vygotsky and Luria suggested that writing not only allowed one to do new things but more importantly, turned speech and language into objects of reflection and analysis (cf. [12]). From this perspective, writing is of utmost importance as it affects consciousness and cognition through providing a model for speech and a theory for thinking about what is said. It is in fact this new consciousness of language that is central to the conceptual implications of writing. *“Far from transcribing speech, writing creates the categories in terms of which we become conscious of speech”* ([12] p. 119).

Another common characteristic that we find in prior studies conducted in that area is that they have neglected the question of the role of language tools in supporting learning, and more in particular, in helping learners to reflect on and develop awareness of the language they produce.

Moreover, most of the computer language programs supporting free text production available in Swedish have been developed for native speakers such as the grammar checker in Microsoft Word developed by Lingosoft [13, 14], the research prototype Scarrie developed at Uppsala University [15] and the prototype Granska developed at our department (see [16]). Some others have been designed for dyslectics. In comparison, the development of computer assisted second language learning has received very little attention. For example, most of the CALL applications available on the market are rarely able to analyze learners' written or spoken productions [4]. These factors together with the necessity to help second-language learners to study and work in Swedish, have motivated this research.

The question that we address regards the use and the design of language technology tools for the writing process of second language adult learners. Preliminary findings point to the importance of taking human-computer interaction aspects into account for the development of language tools for second language writers. In the following sections, we will give a background, and present a developmental perspective on the use of language tools. Subsequently, we describe interviews with teachers and a pilot study on second language writers using Granska.

1.1 Background

Computers have been used in second language teaching for fifteen years (cf. [17]). However, they have never been recognized by the majority of language teachers as exemplifying good teaching, and still remain peripheral to the core of classroom teaching [17]. These programs have often been based on behaviorist and cognitive models of learning and support mechanical aspects of the language through text-construction software, concordancing software and multimedia simulation software (see [18]).

In contrast to cognitive approaches, communicational approaches emphasize the social aspect of the language acquisition and view learning language as a process of apprenticeship or socialization into particular discourse communities (cf. [18]). Furthermore, communicational approaches have found with the advent of the Internet, a powerful tool for assisting language learning. According to Warschauer and Meskill [18], there are today many ways in which students and teachers can use the Internet to facilitate interaction within and across discourse (cf. computer mediated communication in a classroom; computer-mediated interaction for long distance exchange; accessing resources and publishing on the World Wide Web).

However, as Laurillard and Marullo [17] argued, communicational approaches are not a panacea, for instance “*some teachers focus on entirely communicative functions and offer no systematic treatment of grammatical forms*”; “*when problems of communication did arise, it was often shaky grammar that caused them, because the use of the wrong tense or the wrong person of a verb made what the candidate was trying to convey unclear through ambiguity*” (p.147). Probably the main disadvantage with the communicational approach is that the learner should learn the rules of syntax and morphology by *induction*, and learning in this form puts a significant cognitive load on the learner. This holds specifically, when the learners cannot get the appropriate feedback explaining the way in which the target language works.

From our experience, we believe that it is important to stimulate communicational competence during a learning language process making *explicit* the grammatical rules that underlie the language. As our interest lies on the problem of supporting second language writing processes, we see it as fundamental to provide adult learners/writers/users with tools for helping them to develop understanding, progress and enjoyment when using the target language. We regard writing a second language as a complex process requiring the interplay of diverse aspects such as personal motivation, the individual’s identity, context, culture, intellectual competence, will, etc. (cf. [19]). And more in particular, we view the process as a combination of spontaneous, inductive learning with systematic, deductive learning strategies. The role that a computer language checking program can play in this developmental process is still limited. Computer programs can support some aspects of the language learning process; often the mechanical and rational ones. However, the integration of a language checking computer program into real writing environments providing writers/learners with feedback during the drafting process, would make language technology a promising help.

1.2 A developmental perspective on the use of language tools

The perspective on written language as a tool is grounded in the Socio-cultural perspective (cf. [9, 10]). According to this perspective, the most important psychological tool is language, understood as a semiotic resource providing signs that can be flexible and creatively used in social practices. The role of tools – psychological as well as technical – is viewed fundamental in the understanding of human thinking and learning. This perspective has implications in the way we regard language errors and feedback.

1.2.1 Our view of errors

At first view, it may seem strange to focus on what learners get wrong rather than on what they get right. However, there are good reasons for focusing on errors. >From a developmental perspective, learners’ errors are a rich source for the understanding of how they make sense of and construct a new symbolic system. Errors can also be viewed as the expression of a conflict between the learner’s conceptions of what is correct a use of the target language and what is really correct use in the target language. Reflecting on learners’ misconceptions is a way to understand their errors as steps in a developmental process. As Scott [20] mentioned, “*the focus on development does not, however, imply a model of pedagogy in which the teacher simply transmits her knowledge to the student – a model which [...] could refer to learners acquiring language ready-made*” (p. 164). Scott thus places the individual learner and her meanings center stage and thus perceive a synonymy between cognition and the (re) making of written English as a system of representation. In Kress’ terms “*the individual’s semiotic work is cognitive work*” (1996 p. 237 cited by Scott [20]). From this perspective, “*the pedagogic emphasis is placed instead on enlarging the individual learner’s awareness of written English as a resource for making meaning in the student’s particular field of study*” ([20], p.164). A developmental perspective on learners’ grammatical errors has important implications on the way in which adults learn the target language. >From this perspective, learners acquire the target language – the new symbolic system – reconstructing it from their understanding, needs and already acquired language resources.

1.2.2 Our view on written feedback

Providing written feedback to errors is considered as an important tool for the learners’ understanding although Leki [21] argues that “*written commentary on learner papers is, of course, intended to produce*

improvement, but what constitutes improvement is not so clear” (p. 58). And she points thus to “the necessity to look not at responses written on final drafts but rather at responses written on intermediate drafts and how those drafts were reshaped as a result of the teacher’s comments and, to look at the ongoing dialogue between students and teachers (p. 63).

According to Laurillard and Marullo [17] explicit feedback plays a fundamental role in the learner’s language understanding. Feedback in their terms “*must deal explicitly with any diagnosed misconceptions or mal-rules, must provide help with analyzing the gap between learner performance and target form in terms of the information supplied, or in terms of previously mastered items*” (p. 157).

The view of errors as steps and feedback as a tool within second-language processes contributes much to our understanding of the role that computer programs could play in second-language learning environments. Computer’s features such as analyzing the writer/learner/user’s free text, detecting errors, diagnosing them and providing feedback to them might become a great help for second-language writers. Moreover, the development of such features should be designed for contexts where second language writers have a chance to discuss and reflect on their errors and feedback.

2 Questionnaires and interviews with second language teachers

A questionnaire was sent by e-mail to six second-language teachers from two different teaching language institutions. The questionnaire consisted of 21 open questions. In addition, to the questionnaires, one of us conducted one-hour interviews with each one of the teachers in order to discuss their written answers to the questionnaire. The questions were quite general and covered the following topics: learners’ errors, teachers’ feedback, learners and teachers’ computer use, learners’ language level, the learners’ writing process, and the goals of second language teaching. The primary goal with the questionnaire and interviews was to know second language teachers’ opinions on common and important language errors and the appropriate feedback to them.

2.1 Preliminary findings

2.1.1 Second language teachers’ opinions

Second language teachers mentioned that the type of errors made by the learners depends much on the level of the language that they have reached. In general they point to syntactical errors and in particular word order, verb inflection, agreement and use of prepositions as errors often committed by second language learners.

The ways in which they commented on errors varied although there seem, to be at least two approaches. Those who don’t provide any written explanation, they only underline the phrase or the word expecting that the student will actively find the appropriate explanation and, the others who comment in the margin the type of the error made and may even suggest an alternative word or sentence.

Among the teachers providing explanations of the errors committed, there are some who also discuss orally the corrections suggested, together with the students (in general teachers do that with the whole class, writing the most common errors on the blackboard). Five out of six teachers confessed that they do not have enough time to provide individual and written feedback to each one of the learners. Only one teacher mentioned to be strict with correcting at least two versions of the same text.

When we asked the teachers about the appropriate written commentary or about the impact of their comments on the teachers’ texts, their answers are unclear and some teachers seem not to have reflected on the role that written feedback plays for the students’ improvement. According to Leki [21], the available research on second language writing suggests that “teaching by written comment on composition is generally ineffective “ (p. 61). However, Fathman and Walley [22] show that specific feedback on grammatical error has a greater effect on the improvement of grammatical accuracy than general feedback on content has on the improvement of content.

What constitutes appropriate written commentary within second language writing is an open question that deserves further research. However, we think that features providing immediate feedback, supporting drafting processes and, affording specific grammatical detection, diagnosis and alternative suggestions to language errors could become useful for the understanding/improvement of the target language.

3 A pilot study on second language writers using Granska

In this section a pilot study of three writers is presented. The pilot study had two aims; the first aim was to explore and evaluate a rather simple but naturalistic way to collect data from writers’ revisions during free text production. The second aim was to start to study how a spell and grammar checker for Swedish, Granska, should be adapted to second language learners.

In order to adapt Granska to second language writers, we wanted to know more about the detection capacity of the program when used by second language writers, and also more importantly, to study the

role of feedback from the program and if the users follow the advice from the program. We were also interested in how the users would react on false or misleading responses from Granska. The following questions were addressed and explored in the pilot study.

1. Does Granska support second language writers' revision process in their free text production?
2. What parts of Granska are most important to improve and develop further in order to adapt Granska to second language users?
3. Is the research method chosen suitable for studying second language learners' free text production?

3.1 Granska – a Swedish grammar checker

Research and development of language tools has been going on for many years at our department (cf. [23, 24, 25]). Recent work in the group has considered the integration of different functions such as grammar checking and proofreading, linguistic editing functions, language rules and help system into the processes of writing and document handling. This work has resulted in Granska, a prototype Swedish grammar checker and general language toolbox, which has attracted interest from both researchers and potential users [16, 26, 27, 28]. The current version of Granska, used in this study, is designed for native Swedish writers.

3.2 Granska and the revision process

There exist several cognitive models of the writing process, but one of the most well known is the model given by Flower and Hayes [29]. They divide the writing process into three subprocesses; planning, text production and revision. The writer goes from one process to another in a non-linear way following a complex pattern adapted to the writing task.

In the following we will focus on the revision process, which can be divided into three sub processes; detection, diagnosis and correction. A computer program for grammar checking can support the revision process by marking up the errors in the text (detection), give comments on the errors (diagnosis) and propose how the error should be corrected (correction).

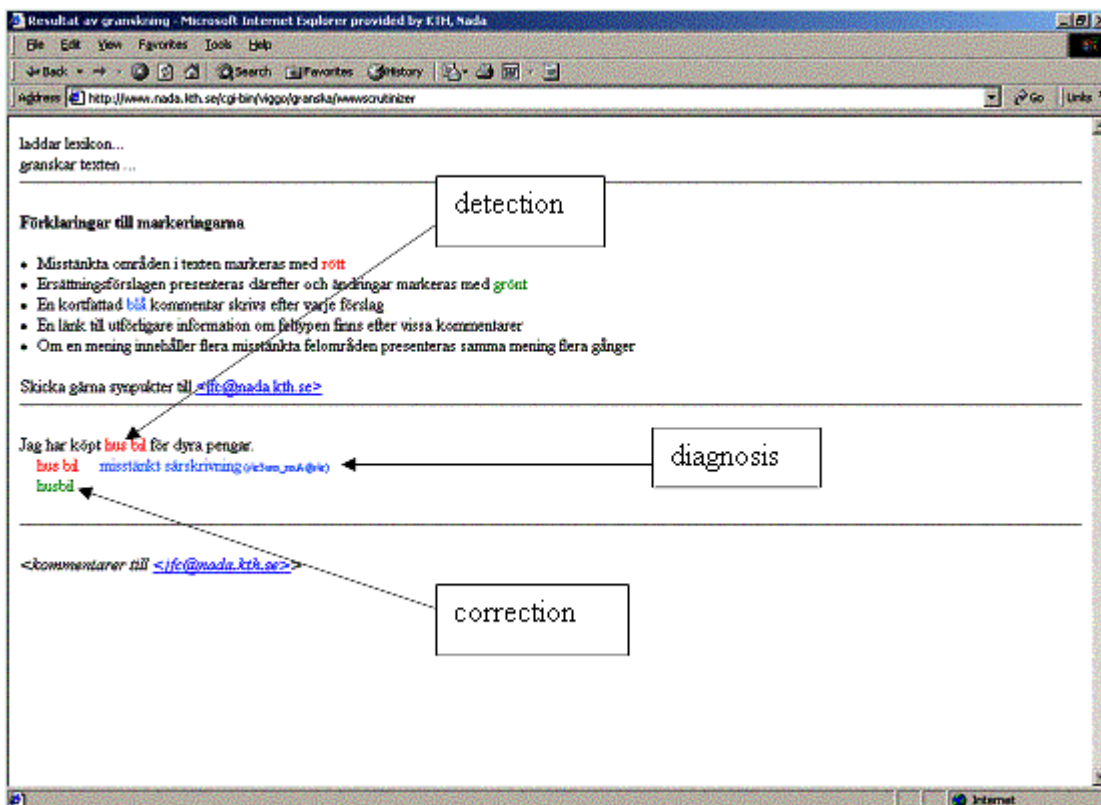
Our system Granska has been developed with several different interfaces with different levels of interaction with the user. In one version, Granska was only one tool among others for language support in a writing environment with word processing facilities. Using such system for studying free text production would mean a strong limitation for the writers in their choice of word processing environment. Moreover, we would have to build some kind of logging system in order to follow and study the users outside the laboratory. We have also developed a version of Granska as a plug-in to Microsoft Word, but using that fails on the same reasons.

In the study presented in this paper a web interface has been used in connection with Granska. All the limitations with system dependencies, described above, led us to a user interface on the web. The web interface only supports the revision process in isolation; the grammar checker is not integrated in a word processing environment. By doing so, we believed that the study of how Granska supports the revision process would be more straightforward and easier to follow both for the writers and for us.

The web interface is basically an interface where the users can submit their sentences or whole texts and in return get back a version of the text (an HTML-page) with errors marked up, supplemented with comments and corrections proposals (see figure 1). The web version allowed the users to write their texts in any text editor.

In addition, the web interface is accessible for all users, and independent of the system used. Through the web interface the users can submit sentences, links to text files on the web, and also text files from their own hard drives to Granska. Using the web interface at any time from any computer, the users can take part in the study whenever they want. In the web interface Granska's support on the three sub processes of revision (detection, diagnosis and correction) can be pointed out for the users, see figure 1.

Figure 1. Granska's web interface gives a clear view of how the revision process is supported by the program.



When studying the revision process in detail, we focused on the following questions:

- How do the users react on false alarms and misleading diagnosis and correction proposals?
- Are all steps in the revision process necessary to support, or can the users fill in what is missing in Granska's feedback?

These two questions are both related with the fact that current grammar checkers have a limited understanding of the text, and sometimes fail when analysing it. This leads to the programs giving erroneous feedback to the user, false alarms and misleading or inappropriate feedback. In earlier studies, false alarms have seemed to be problematic for the native Swedish writers [27].

3.3 Method and data collection

Studying writers in action is not an easy task. The writers want to succeed in their writing and can be hard to observe due to the cognitive nature of writing. One successful way to study writers has been with so called thinking aloud methodology [30]. With this method the subjects are talking aloud about what they are doing when performing a task. The writing task is often predefined and limited in time. Using think-aloud methodology often means an experimental setting in a research laboratory. We wanted to get out of the laboratory in order to study free text production.

The prerequisite for the users in the study was the following: use Granska whenever you want and when you think it will help you. The control of the data collection was thus left to the users. According to the instructions the user should save the original text scrutinized with Granska and also the final version, written after the revision aided by Granska. Collecting the two versions gave the opportunity to follow the users' actions. From the two versions we can answer the following questions.

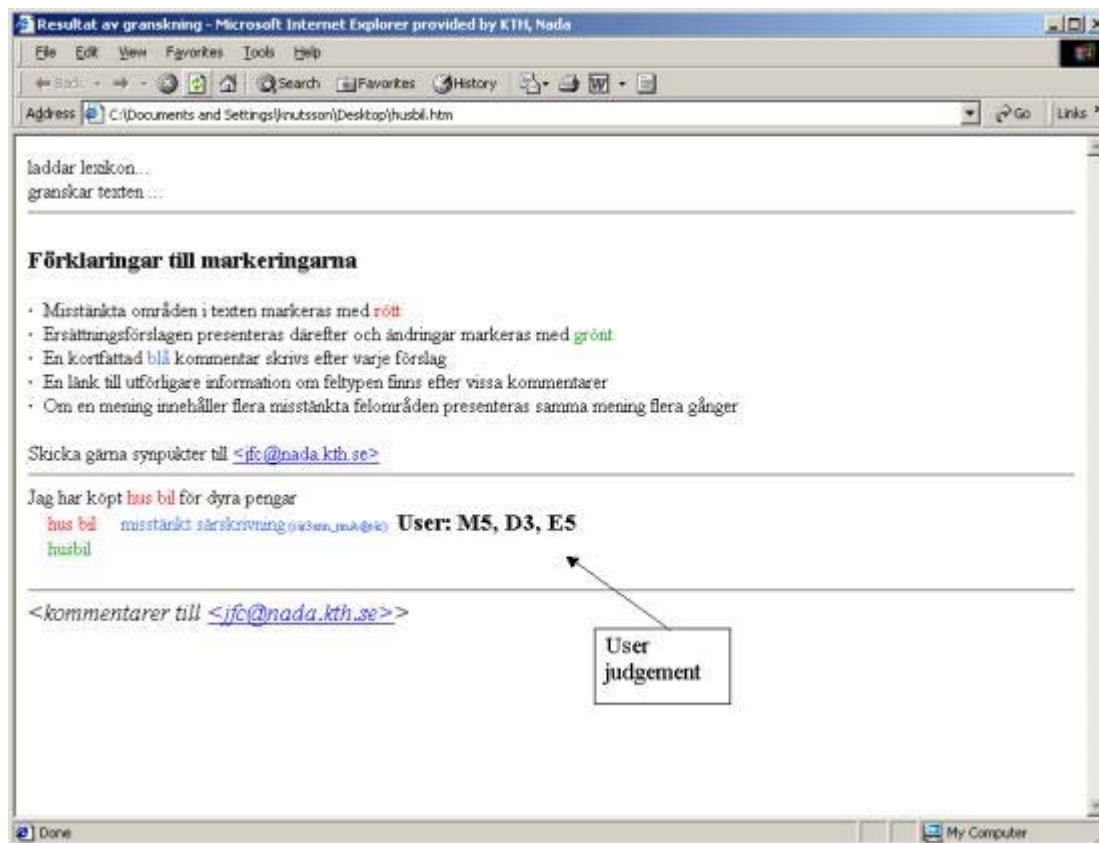
- Did the users follow the advice from Granska?
- How many errors did the users revise with limited or no feedback from Granska?
- How many errors were left unchanged in the text?

3.4 Users judging Granska's feedback

We wanted the users to judge the program's detections and feedback using grades. By using a grading procedure we wanted to see if grades could track some of the revision process and if the grades could give us some hints on what detection or feedback that was problematic for the user.

The users were given instructions on how to use a grading procedure on the alarms from Granska. They were supposed to use a scale from 1 to 5, where 1 is very bad and 5 is very good. The users were requested to use the grading in order to judge the detections, diagnoses and correction proposals from Granska. Definitions of detections, diagnoses and corrections were given in text also by using figure 1 above. Every scrutinized document was supposed to be annotated with user judgements on the alarms from the program. The scrutinized version of the text (see figure 1), an HTML-page, could serve as the medium for user annotations both on paper and electronically. The user should write their grades on every alarm from the program. An example on this is given in figure 2.

Figure 2. Example of a user judgement made electronically. M5 means that detection was graded with a 5, D3 that the diagnosis was graded with a 3 and E5 means that the correction proposal was judged with a 5.



3.5 User data

Three users participated in the pilot study. Two were PhD students and one was a research fellow. They used Granska when working at an academic department. Their knowledge and skills in Swedish were on different levels, see table 1.

Table 1. User data from the pilot study.

User	Age	Native language	Language level	Time of study	No. of texts	Length of writing	Graded texts
A	34	Spanish	Advanced	6 months	18	2240 words	4
B	37	Chinese/Spanish	Intermediate	4 months	16	244 words	4
C	34	German	Advanced	2 months	1	190 words	1

Text from emails was the major text type and the texts were related to the users' work. Most academic writing in the department is written in English, but when communicating by email Swedish is the main language at the department. The three users used three different mail clients when editing their texts.

The most active user was user A. She had used Granska for revising 18 texts, mostly emails. The final versions of the emails contained together 2240 words. The emails were often sent to several people and some were checked with Granska twice in order to detect new errors introduced by the new revisions. User A judged Granska's alarms in four texts. She made the grading annotations on paper. Most data in this study comes from this user.

User B had used Granska in a little different way than the other users by sending only smaller pieces of text, mostly parts of emails, to Granska for response of the sentences' grammaticality. He used Granska 16 times and wrote 244 words. Some pieces of texts include variants of formulations. User B has not submitted the final version of text but has instead made comments on nearly all text pieces, and all alarms from Granska were judged by the user. User B made the grading annotations of Granska's performance electronically.

User C had used Granska only at one occasion, mainly because the user did not write text in Swedish during the period of the study. User C had written a rather long and important email message, 190 words, and scrutinized it twice with Granska. One of the text versions with feedback from Granska has been

annotated with the user's judgments of Granska following the given grading scale. User C has made the grading annotations electronically.

3.6 Error typology

The focus in the pilot study is on the errors and what the users do when they encountered the errors by themselves or in interaction with Granska. All errors in the texts were manually analysed and annotated. In the following error typology the broad error types for all linguistic errors are presented, see table 1. Granska does not have the capacity to detect lexical, semantic and pragmatic errors, but we believe that it is important to present all errors produced by the users.

Table 2. The error typology with the number of Granska's detections, diagnoses and correction proposals.

Error type	Number of errors	Granska's	
		detections and diagnoses	corrections
Typographical	46	4	4
Orthographical	54	46	44
Morphosyntactical	22	8	8
Syntactical	81	24	20
Lexical	9	0	0
Semantic	2	0	0
Pragmatic	5	0	0
Style	4	0	0
All errors	223	82	76

There were totally 223 errors in the texts comprising 2700 words. Granska detected 36,8% of all errors and proposed corrections for about 34,1% of all errors. The users detected a few of the errors without Granska's help and repaired some of errors where Granska only gave a detection and diagnosis. About 15 false alarms occurred mostly caused by the program's limited knowledge of idioms, see below.

3.7 Error types and user's actions

Different error types are handled in different ways both by the users and the program; and also with different results. In the following section the broad error types are presented together with the findings on how the users treated them in interaction with Granska.

3.7.1 Morphosyntactical errors

Morphosyntactical errors concern errors within the phrase, in this study mostly agreement errors within the noun phrases, for example "en spöke" (eng. "a ghost"). Agreement errors were quite a few in the material, and when the program detected the error the user followed the given proposal. The program can nearly always provide the user with a relevant proposal for correction. About 68% of the errors were undetected; they were often more deviant and therefore hard to detect without causing a lot of false alarms when running the program on Swedish natives text. These errors should be detected with rules that are specially designed for second language learners.

3.7.2 Syntactical errors

Syntactical errors are divided into many error types. Common for the users with Spanish as their mother tongue was word order errors. There is very little support for detecting these errors in the system. In some cases the program was successful and in these cases the users followed the advice from the program mostly with good results. On word order errors the program gave good proposals for correction, which seems to be an important condition if the user should follow the program's advice. Other errors caused by syntactic violations are real word error (spelling mistakes resulting in a real but wrong word), missing words and also problems in choosing the correct preposition or to choose between indefinite or definite form of words or

phrases. These errors are mostly missed by the system, and much research effort must be put on these problems in order to help second language learners.

3.7.3 Orthographical errors

The users follow the program on nearly all alarms on orthographic errors (spelling mistakes). One explanation is that the program is not only good at detecting orthographic errors, in many cases the program also gives perfectly ranked correction proposals. In some cases the correction is wrong according to context, for example the program gives a word in indefinite form when it should have been definite. There are also other errors important to handle, which violate the spelling norm of Swedish in such a substantial way, that make them hard to correct in the current version of the program.

3.7.4 Lexical and semantic errors

Lexical and semantic errors were not too many in the material. The users seemed to have a good vocabulary, but some words are used in a way, far from the Swedish norm. An integration of an on-line dictionary might provide some help. The grammar checker can hardly detect these errors.

3.7.5 Pragmatic errors

There were also some problems in language use, mostly caused by interference of English, but these problems were limited. The program is totally unaware of these problems, but in the future some direct translations of English idioms might be detected and corrected based on corpus observation.

3.7.6 Typographical errors

Typographical errors are mainly errors in punctuation, capitalization, time and date format and abbreviations. Errors in this group are rather frequent in the material, which is a bit unexpected. In the current version of the Granska prototype only a few of these kinds of errors are detected. For different reasons we have focused on spelling and grammatical errors when developing the last versions of Granska. The findings in the study lead us to the fact that we have to consider typographical errors again. These errors seem to be related to different conventions between languages when it comes to typographic issues in writing. One of the users in the study repeated an error by always putting a space before question and exclamation marks. Afterwards, the user commented that it is the convention in French. A computer could easily support the correction of these kinds of errors.

3.7.7 Style errors

Style errors were not frequent in the study, but might be important to point out for the user who wants to adapt to the norm for written Swedish.

3.8 Ignored alarms and alarms causing revision

3.8.1 Some alarms are always ignored

The type of error that the program detects but that was always ignored by the users was “finite verb missing”. The program signaled this alarm when a finite verb was missing in a clause. On this error type the program gives an error diagnosis full of linguistic terminology and with no proposal for correction. The explanation and diagnosis of the error can be improved, but not in a straightforward way and therefore one has to construct them in a way that leads to understanding and a starting point for self-correction. To give correction proposals on these types of errors is very hard without limiting the language domain and by doing so leave the context of free text production.

3.8.2 False alarms

We also have to admit that some of the alarms may seem stupid to the users by remarking on idioms and short and correct expressions, for instance that there is no finite verb in the expression “Gott nytt år” (eng. Happy new year). Many of these alarms can be avoided by using heuristics in the program. None of these false alarms have caused the users to make changes in their texts. When looking at the comments from users they seem to get annoyed when the program continues to mark and comment on sentences without finite verbs. The interface of Granska must provide the user with clear possibilities to turn on and off rules in the program.

3.8.3 What makes users revise their text

Especially the proposal for correction seems to be critical if the user should change the text according to directives from the program. One exception to this is errors on the word level, where the user in some cases can use a dictionary to look-up the relevant correction. Indications on this are given in the material several times. These indications are based of the following observations.

- The user writes a misspelled word that leads to detection and proposals for correction.
- The program generates a word form that the user seems to understand is wrong.
- The user looks up the correct word form in a dictionary and writes it in the revised text.

3.9 Results from the user judgments

The users have to a varying degree judged the detections, diagnoses and corrections. According to the instructions given all alarms from Granska should be judged. In table 3 below, 97 judgements from the users are presented. Most judgements were made on the detections and diagnoses and fewer were made on the corrections. The explanation to this is that Granska gave fewer correction proposals than detections and diagnoses.

Table 3. The user judgments on the detections, diagnoses and corrections from Granska.

Error type	Detections		Diagnoses		Corrections	
	Number of judgements	Mean value	Number of judgements	Mean value	Number of judgements	Mean value
Typographical	0	-	0	-		-
Orthographical	18	4,6	18	3,9	16	4,2
Morphosyntactical	4	4,0	4	3,8	3	3,7
Syntactical	13	3,5	13	3,2	8	4,5
Lexical	0	-	-	-		-
Semantic	0	-	-	-		-
Pragmatic	0	-	-	-		-
Style	0	-	-	-		-
All errors	35	4,1	35	3,6	27	4,2

On orthographical errors the users seemed to be most satisfied with the detections and the corrections. The diagnoses are the same on all orthographical errors, and are very short “unknown word”. This diagnosis seemed to be too short for the users, especially when no correction proposal was presented.

The users seemed to approve the corrections proposal on syntactical errors when Granska could present them. The diagnoses got a rather low grading, and this is a signal that they must be improved. When scaling up the study we have to consider automatically collecting and grouping the judgments. An extended study with many judgements will give more information on the individual rules in Granska.

4 Discussion

Using the two versions of the writer’s text seems to be a good and simple tool to collect data about the users’ decisions when they are revising their texts. Complemented with the users’ judgement some of the revision process can be traced. The judgements from the users gave fine-grained information of the different steps using Granska (detection, diagnosis and correction). Just looking at the final result in the text gives little help to understand what has led to the decision by the user. The judgments also seem to give rise to other comments from the user, which often are richer than just the grades.

There seems to be a trade-off between the number of judgements and the number of sessions with Granska. It sometimes seems difficult to incorporate the judgement process in the writing process.

False alarms were not a big problem for the users in the study. The program did not seem to fool the users when one compares the output from Granska with the users’ final text versions. User A and C could in some cases construct their own corrections, when they seemed to realize that the program was misleading them. User B seemed to trust the program too much in some cases, according to the judgements on Granska’s alarms made by the user.

The diagnoses must be rewritten and evaluated with users. It might also be good to involve teachers in this process. Some of the diagnoses seem difficult to understand and annoyed the users. If there was an understandable diagnosis and a correction proposal the users gave both high grades. If the correction proposal was missing, the diagnosis was often assigned a lower grade. Therefore it is important to improve Granska's capacity on generating correction proposals.

Did Granska support the second language writers? On one hand it did when the two text versions from the users were compared. The program detected many errors and the user improved the text by following the advice from the program. On the other hand there were still a lot of errors that were undetected by both Granska and the writers. Many of those were syntactical errors and much research effort must be put on these problems in order to support second language learners free text production. An effort in this direction is a system based on statistical error detection recently developed at our department [31].

When the errors that Granska searches are compared with the errors the teachers said were important, there is an overlap. The error categories that the teachers point to are word order errors, use of prepositions, verb inflection and agreement errors. In the current version of Granska, agreement errors and to some extent word order errors are detected. Future efforts in the development of Granska must improve the detection and feedback on these errors, and also start work on the other errors that both are pointed out by the teachers and are frequent in the study presented here and in other second language corpora.

In this paper we have presented preliminary results from an ongoing study. The study will be extended in the near future. There seems to be a need for alternative and easier ways for grading and for the submission of the sessions with Granska. One user suggested that the judgement should be integrated in the user interface. This is a good idea, which we have to consider before taking the next step in the study. One other user suggested that it would be much easier to send the emails directly to Granska without using the web interface, and get the scrutinized text as a reply. This would both simplify the process for the users and facilitate the collection of data.

5 Acknowledgements

We want to thank teachers and users who have participated in the pilot study. The work has been funded by Swedish Agency of Innovation Systems (Vinnova) and the Swedish Research Council (VR). Ola Knutsson's work has partly been funded by a scholarship from the Swedish Academy and the Swedish Language Council.

References

- [1] Hammarberg, B. & Viberg, Å. (1976). Reported speech in Swedish and ten immigrant languages, SSM. Rapport 5. Stockholm: Stockholms universitet, Inst. för lingvistik.
- [2] Hyltenstam, K. (1983). Data types and second language variability. In Ringbom, H. (ed.). Psycholinguistics and foreign language learning. Papers from a conference held in Stockholm and Åbo. Oct, 25-26 1982. Åbo Akademi. 57-74.
- [3] Viberg, Å. (1981). Svenska som främmande språk för vuxna, In Hyltenstam, K. (ed.). Språkmöte. Lund. Liberläromedel, 21-65.
- [4] Cerratto, T. and Borin, L. (2002). Swedish as a Second Language and Computer Aided Learning Language. Overview of the research area. Technical report TRITA-NA-P0206, NADA, 2002.
- [5] Kotsinas, U-B. (1982). Svenska svårt. Några invandrades svenska talspråk. Institutionen för nordiska språk. Stockholm.
- [6] Kotsinas, U-B. (1983). On the acquisition of vocabulary in immigrant Swedish. In Ringbom, H. (ed.). Psycholinguistics and foreign language learning. Papers from a conference held in Stockholm and Åbo. Oct, 25-26 1982. Åbo Akademi. 75-100.
- [7] Kotsinas, U-B. (1985). Invandrare talar svenska. Liber, Malmö.
- [8] Jansson, G. (2000). Tvärkulturella skrivstrategier. Kohesion, koherens och argumentationsmönster i iranska skribenters texter på svenska. Institutionen för nordiska språk Uppsala universitet. Skrifter utgivna av Inst. för Nordiska språk 49.
- [9] Vygotsky, L. (1962). Thought and language. Cambridge, MA: MIT Press.
- [10] Vygotsky, L. (1978). Mind and Society : The development of higher psychological processes. Cole, M.; Steiner, J., Scribner, S. and Souberman, E. (Eds). Cambridge, MA : Harvard University Press.
- [11] Downing, J. (1987). Comparative perspectives on world literacy. In Wagner, D. (Ed.). The future of literacy in a changing world, Pergamon Press, Oxford, pp. 25-47.
- [12] Olson, D. (1995). Writing and the mind. In Wertsch, J., Del Rio, P. and Alvarez, A. (Eds.). Sociocultural Studies of Mind. Cambridge University Press, Cambridge. pp.95-123.

- [13] Arppe, A. (2000). Developing a grammar checker for Swedish. In Proc. 12th Nordic Conference in Computational Linguistics, Nodalida-99. Trondheim, pp. 13-27.
- [14] Birn, J. (2000). Detecting grammar errors with Lingsoft's Swedish grammar checker. In Proc. 12th Nordic Conference in Computational Linguistics, Nodalida-99. Trondheim, pp. 28-40.
- [15] Sägvall Hein, A. (1998). A chart-based framework for grammar checking. Initial Studies. In Proc. 11th Nordic Conference in Computational Linguistics, Nodalida-98, Copenhagen, pp.68-80.
- [16] Domeij, R., Knutsson, O., Carlberger, J. & Kann, V. (2000). Granska - an efficient hybrid system for Swedish grammar checking. In Proc. of the 12th Nordic Conference in Computational Linguistics, Nodalida-99, Trondheim, pp. 49-56.
- [17] Laurillard, D. And G. Marullo (1993). Computer-based approaches to second language learning. In P. Scrimshaw (Ed.) *Language, classrooms and computers*, Routledge, London, pp. 145- 165.
- [18] Warschauer, M. & Meskill, C. (2000). Technology and second language learning. In J. Rosenthal (Ed.), *Handbook of undergraduate second language education* Mahwah, New Jersey: Lawrence Erlbaum, pp. 303-318.
- [19] Sjögren, A. (1996). Språket, nykomligens nyckel till samhället men också en svensk försvarsmekanism. In Sjögren, A Runfors, A. and Ramberg, I. (ed.). *En 'bra' svenska? Om språk, kultur och makt*. Mångkulturellt centrum, pp. 19-38.
- [20] Scott, M. (2001). *Written English, Word processors and Meaning Making. A semiotic perspective on the development of adult students' academic writing*. In L. Tolchinsky (ed.). *Developmental Aspects in learning to write*. Kluwer, Dordrecht.
- [21] Leki, I. (1990). Coaching from the margins: issues in written response. In B. Kroll (Ed.). *Second language Writing. Research insights for the classroom*. Cambridge Applied Linguistics, Cambridge pp. 57-68.
- [22] Fathman, A. and Whalley, E. (1990). Teacher response to student writing : focus on form versus content. In Kroll, B. (Ed.). *Second Language Writing. Research insights for the classroom*. Cambridge University Press, Cambridge, pp. 178-210.
- [23] Severinson Eklundh, K. (1991). *Språkliga datorstöd för redigering under skrivprocessen*. IPLab Arbetspapper-31, Nada, Kungliga Tekniska Högskolan, Stockholm.
- [24] Cedergren, M. & Severinson Eklundh, K. (1992). *Språkliga datorstöd för skrivande: förutsättningar och behov*. Report IPLab-58, Nada, KTH.
- [25] Domeij, R., Hollman, J. & Kann, V. (1994). Detection of spelling errors in Swedish not using a word list en clair, *Journal of Quantitative Linguistics* 1:195-201.
- [26] Domeij, R., Knutsson, O., Larsson, S., Rex, Å., Severinson Eklundh, K. (1998). *Granskaprojektet 1996-1997*. Report IPLab-146, Nada, KTH.
- [27] Domeij, R., Knutsson, O. och Severinson Eklundh, K. (2002). Different Ways of Evaluating a Swedish Grammar Checker. In the proceedings of The Third International Conference on Language Resources and Evaluation (LREC 2002), Las Palmas, Spain.
- [28] Knutsson, O. (2001). *Automatisk språkgranskning av svensk text*. Licentiate thesis, TRITA-NA-0105 (IPLab-180), Nada, KTH.
- [29] Flower, L. S., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, pp. 365-387.
- [30] Hayes, J. R. & Flower, L. (1983). *Uncovering Cognitive Processes in Writing: An Introduction to Protocol Analysis*. In Mosenthal, Tamer & Walmsley (Eds.), *Research on Writing: Principles and Methods*. New York: Longman.
- [31] Bigert, J & Knutsson, O. (2002). Robust Error Detection: A Hybrid Approach Combining Unsupervised Error Detection and Linguistic Knowledge. In the Proceedings of Romand02, 2nd Workshop on RObust Methods in Analysis of Natural language Data , Frascati, Italy.

Authors

Ola Knutsson, Ph.Lic.
 IPLab, Department of Numerical Analysis and Computer Science
 Royal Institute of Technology
 100 44 Stockholm
 knutsson@nada.kth.se

Teresa Cerratto Pargman, Ph.D.
 IPLab, Department of Numerical Analysis and Computer Science

Royal Institute of Technology
100 44 Stockholm
tessy@nada.kth.se

Kerstin Severinson Eklundh, Professor
IPLab, Department of Numerical Analysis and Computer Science
Royal Institute of Technology
100 44 Stockholm
kse@nada.kth.se