Standardising written feedback on L2 student writing

H. Louw Hons. B.A.

Dissertation submitted for the degree Master of Arts in English at the North-West University, Potchefstroom Campus.

Supervisor: Prof. A.J. van Rooy

January 2006
Potchefstroom

Acknowledgements

I would like to thank the following people and institutions for their assistance:

Professor Bertus van Rooy (promoter) with more ideas and plans than one human being should be blessed with.

The staff at the School of Languages in Potchefstroom for support and encouragement.

The NRF: Without the funding received from the NRF, this thesis would never even have been started.

Miss Audrey van den Bergh: You had to put up with a lot.

My parents: It is difficult to encourage a son 1300 kilometres away. You did it!

God Almighty: For too many things to mention.

Abstract

Keywords: feedback, second language student writing, computer-assisted language learning (CALL), language teaching, standardization, error, error classification

The primary aim of this study is to determine whether it is possible to standardize written feedback on L2 student writing for use in a computerised marking environment. It forms part of a bigger project aimed at enhancing the feedback process as a whole.

The study attempts to establish "best practice" with regards to feedback on writing, by establishing from the literature what works and what should be avoided. Also, an empirical study was launched to establish what lecturers focus on and what marking techniques they use. A set of randomly selected essays from the Tswana Learner English Corpus and the Afrikaans Learner English Corpus were sent to the English departments of different tertiary institutions across the country. The essays were marked by the English lecturers at the relevant institutions. The conclusion was that lecturers typically focus on surface structures, and use ineffective marking techniques.

The best practice (and data from the empirical study) was then used to create a set of standardised feedback comments (tag set) that can be used in a specially programmed software package in which students submit their texts electronically. Lecturers can then mark the student essays on the computer, hopefully speeding up the process, while at the same time giving much more detailed feedback. In later stages of the bigger project, students will get individualized exercises based on the feedback, and

there are experiments currently being run to try and automate certain parts of the marking process in order to take some strain off the lecturers when marking. The immense archiving abilities of the computer will also be utilized in order to create opportunities for longitudinal studies.

The effectiveness of the feedback tag set was tested in comparison to the marking techniques used by the lecturers in the empirical study and a self-correcting exercise. The conclusion was that the feedback tag set is more effective than the other two techniques, but students seem to perform weak overall when it gets to the revision of cohesive devices and supporting arguments. I argue that students are not used to revising these features, since lecturers seldom (if ever) comment on the structural elements of texts. However, the experiment proves that standardization of written feedback is possible to an extent.

The implications of the findings are discussed, and recommendations for further research are made.

Opsomming

Sleutelwoorde: terugvoer, tweedetaal-studente skryfwerk, rekenaargesteunde taalonderrig (RGTO), taalonderrig, standaardisering, fout, foutklassifikasie

Die hoofdoel van hierdie studie is om vas te stel of dit moontlik is om geskrewe terugvoer op tweedetaalskryfwerk van studente te standaardiseer met die oog op gebruik in 'n gerekenariseerde merkomgewing. Dit vorm deel van 'n groter studie wat ten doel het om die hele terugvoerproses te verbeter.

Hierdie studie poog om "beste praktyke" ten opsigte van terugvoer op skryfwerk te bepaal, deur vanuit die literatuur vas te stel wat werk en wat behoort vermy te word. Verder is 'n empiriese studie van stapel gestuur om vas te stel waarop dosente fokus en watter merktegnieke hulle gebruik. 'n Versameling ewekansig-geselekteerde opstelle van die Tswana Leerderengelskorpus en die Afrikaanse Leerderengelskorpus is na departemente Engels by verskillende tersiêre instellings landwyd gestuur. Hierdie opstelle is gemerk deur dosente in Engels by die betrokke instellings. Die gevolgtrekking was dat dosente tipies op oppervlakelemente fokus, en oneffektief gebruikmaak van merktegnieke.

Die beste praktyk (en die data uit die empiriese studie) is toe gebruik om 'n stel gestandaardiseerde terugvoeropmerkings (etikettestel) te skep, wat gebruik kan word in 'n spesiaal-geprogrammeerde sagtewarepakket waarheen studente hulle opstelle elektronies kan stuur. Dosente kan die opstelle dan op die rekenaar merk, en sodoende hopelik die merkproses bespoedig, terwyl hulle daarmee saam meer

gedetailleerde terugvoer gee. In die verdere verloop van die groter projek sal studente geïndividualiseerde oefeninge kry wat op die terugvoer gebaseer is, en daar word ook tans eksperimente gedoen om te probeer om sekere aspekte van die merkproses te outomatiseer, om sodoende van die druk van dosente af te haal. Die buitengewone stoorkapasiteit van die rekenaar sal ook gebruik word om geleenthede vir longitudinale studie te skep.

Die effektiwiteit van die terugvoerstel is getoets in vergelyking met die merktegnieke van die dosente in 'n empiriese studie, saam met 'n selfkorrigeringsoefening. Die slotsom is dat die terugvoerstel meer effektief is as ander twee tegnieke, maar dat studente oor die algemeen swak vaar wat die hersiening van kohesiemiddele en ondersteunende argumente betref. Ek argumenteer dat studente nie gewoond is daaraan om hierdie eienskappe te hersien nie, omdat dosente selde (indien ooit) terugvoer gee op hierdie strukturele elemente van tekste. Nietemin, die eksperiment bewys dat die standaardisering van geskrewe terugvoer tot 'n mate moontlik is.

Die implikasies van die bevindinge word bespreek en aanbevelings vir verdere navorsing word gemaak.

CONTENT

1. STANDARDIZING WRITTEN FEEDBACK ON L2 STUDENT WRITING

Feedback on written assignments is a given in most educational settings in South Africa. Teachers or lecturers "mark" students' writing and assign a grade to it. This is in spite of a global debate about the effectiveness of feedback. Some see feedback as ineffective and a waste of time, based on studies finding that students mostly ignore feedback or do not understand it. The long-term benefits of feedback are also questioned (see Hyland, 1998:255 for a discussion of the various conflicting ideas about the effectiveness of feedback). This dissertation argues that the ineffectiveness of feedback is due mainly to inconsistency and a lack of transparency. The dissertation reports on a project which aims to standardize written feedback to an extent, by utilizing "best practice" in feedback techniques and combining it with computer technology. The project is part of a bigger project with the aim of developing a computerised marking system which can provide individualized remediation for students and be used to automatically create a corpus of learner writing, thereby enabling longitudinal studies on a large scale.

One of the problems in the debate surrounding feedback is that "feedback" is not clearly defined. Many researchers on feedback mentioned in this study do not clearly say what they encapsulate under the heading of "feedback". For the purposes of this study, "feedback" refers to any indication of an error in, or any comment about, a written text. Apart from a clear definition, unfortunately, there are other problems with feedback as well.

1.1 Problems with feedback

The four main problems with providing feedback are the following:

1.1.1. Lack of consistency

There is inconsistency in feedback procedures and this inconsistency can cause problems. The method for providing feedback is not standardised. Different lecturers use different methods of providing feedback. In some instances all errors are not indicated every time. Certain types of errors are more likely to be treated than others and the more often a particular type of error is made, the less likely the teacher is to treat it. Teachers also sometimes indicate errors that have not been made or indicate more than one type of error in the same way (Ellis 1996:585, Nwaila, 1996:83 and Van der Walt and Van Rooy, 2002:115). As feedback is a form of input, this is a grave problem. Krashen (1985:43-52) indicates that inappropriate input may contribute to the problem of fossilization.

1.1.2. Feedback is labour-intensive and time-consuming

The feedback process is complex (Ellis, 1996:585) and the process of decision-making takes time. Spencer (1998) indicates that the method of providing written feedback that was found to be the most effective, also proved to be the most labour-intensive, increasing the workload on already overworked lecturers.

1.1.3. Students expect feedback but do not always know how to handle it

Students and parents expect feedback and error correction from academic institutions (Spencer, 1998:208). Nevertheless, it appears as if a lot of the effort going into the marking of student writing is wasted, as students do not know how to handle feedback, or simply discard it. This may be due to the marking technique. James (1998:236-238) indicates that by simply indicating an error, it is not placed in context of the *type* of error it actually is, and therefore the source of the problem is not found. Feedback should lead to correction that should lead to remediation. "Surface" correction is mere editing, but does not address the source of the problem. Many students cannot correctly identify the specific error due to untransparent marking methods, and if they can, they do not always know how to correct it (Monyaki, 2001:66). This vagueness may be attributed to teachers' lack of the skills to analyse and explain the problems students experience (Hyland, 2003:218).

Feedback may also be counter-productive for the learning process, if used incorrectly. Feedback is a type of consciousness-raising, whereby learners are reminded of where they do not have the target language features under full control. If handled incorrectly, learners will not see their errors (and the lecturers' feedback) as a learning opportunity, and may feel that they should strive for "perfect" language use. Instead of experimenting with the language, they then stick to what they know they are capable of. This leads to the undesirable effect known as avoidance (See Hyland, 1998:264).

1.1.4. Students do not recognise recurring patterns of errors in their own writing

Wible, Kuo, Chien, Liu and Tsao (2001:308-310) indicate that students may not notice a pattern of errors in their own writing as "each instance of the error appears to the student amidst a forest of other markings." This applies to the teacher/lecturer as well. Lecturers also get tired of indicating the same error repeatedly and stop doing it. The student may then not realise his/her error.

The main problems with feedback practice therefore have to do with the techniques of providing feedback (which are time-consuming and ineffective) as well as a lack of action in response to the feedback – both on the part of the students and the lecturers. There is also a psychological barrier to overcome – to see feedback as an opportunity for learning, and not as a slap on the hand. The questions can then be raised if and why feedback should be provided if it has such dubious value.

1.2. Why should feedback be given?

Despite the problems facing feedback, there are important reasons to continue the practice: Ellis (1996:50-54) indicates that errors occur due to **lack of knowledge**. Errors should therefore be corrected to increase the knowledge of the learner – i.e. provide input. Feedback is important input in the language learning process. Hyland (1990:285) claims that "Foreign-language students are often anxious about writing, and need to be encouraged to see it as a means of learning, rather than demonstrating learning". The writing process (and the accompanying feedback) should be seen as a

way to enhance SLA. The importance of receiving **comprehensible feedback** should then be obvious as feedback is one of the most common ways a teacher/lecturer can provide input.

The Kingman report (see Nwaila, 1996:51) states that leaving ESL errors uncorrected is a trap that keeps the pupils in their current social and ethnic sectors. It creates "a barrier to their educational progress, their career prospects, and their social and geographical mobility" (Nwaila, 1996:51, see also Titlestad, 1996, and Wright, 1996). Therefore it is very important to handle errors in such a way that learners stop making them as this can help prevent fossilization (Krashen, 1985).

The arguments in favour of feedback are closely linked to the medium it is provided in. The positive aspects of feedback may differ for writing, speaking and reading. In this study, there is a specific focus on writing. There are good reasons to focus on writing.

1.3 Why focus on writing?

Written language, apart from being a more conscious process than spoken language or non-verbal communication, is quite simply easier to archive and therefore easier to study. Feedback is possible in a variety of manners, on all language forms, but the effects it has on spoken and non-verbal communication is difficult to monitor or study. That is arguably why writing is such an important educational tool. Apart from being used for education, writing is also a tool to be used by students for self-education (monitoring). Writing is also used to illustrate and present academic knowledge, and may have advantages in a student's personal development. Last but

not the least; writing is in the public eye. The most obvious reason to focus on writing however is that there is a problem in writing. All of the above viewpoints will now be explained briefly.

1.3.1 Writing as education

Writing is the most obvious way of communicating apart from talking and body language. It is by means of writing that we communicate with society at large and even to those close by. In legal terms, writing is considered to be more permanent and binding than spoken language. In most teaching situations, writing is the main medium through which students indicate their knowledge. Especially in language teaching, writing is important for hard analyzable facts about the competence of students in the language.

Spencer (1998:5) found that writing is an "invaluable tool for education" in that it can be used to promote learning and to help students come to grips with new material. Spencer (*ibid.*) claims that writing creates the opportunity to reinforce language learning, as well as inviting them to be adventurous with language and develop strategic competence. Writing can foster audience awareness and teach students to evaluate material critically; it "promotes the integration of old and new information, is a means of power, a way to understand complex ideas and a route to self-discovery". (Spencer, 1998:5)

Spencer (1998:6) quotes White in saying that once the importance of writing and its many uses are accepted, responding to writing becomes very complex and calls for special thoughtfulness on the instructor's part.

1.3.2 Self-education

Written language is more permanent than other forms of communication. Therefore it provides more opportunity for monitoring and analysis – be it self-monitoring on the part of the student, or monitoring by an external person. Computer-mediated communication is written and can easily be archived electronically, so students have additional time to plan and reflect on their language use (Nel, Dreyer and Carstens, 2001:242).

1.3.3 Monitoring

Corpora provide researchers with the opportunity to analyse these texts in order to draw general conclusions about the usage patterns in the language. Granger (2002:6) indicates that despite a host of knowledge about learner variables such as motivation, learner styles, needs, attitudes etc., there has been a noticeable absence of knowledge of learner output. Granger (2002:6) finds it encouraging seeing the attention of Second Language Acquisition (SLA) and Foreign Language Teaching (FLT) research communities shifting towards descriptions of learner data, through which it may be possible to "rehabilitate" learner output. Coetzee-Van Rooy and Verhoef (2000) found that many students perceive their own language skills as much better than it really is. Due to the more permanent nature of written language, it could prove to be a very important tool in changing that erroneous perception in order to draw learners' attention to academic language proficiency (also see Parkinson, 2002).

Granger (2002:8) also indicates that a corpus of L2 student writing adheres to the requirement of authenticity, seeing as writing is such an integral part of academic activities. Once a corpus of learner writing has been error tagged, "the range of possible applications that can be derived from it is absolutely huge" (Granger, 2002:14).

1.3.4 Academic competence

Written language is often the means through which students have to illustrate academic competence. Coetzee-Van Rooy and Verhoef (2000:180) indicate that although students rate their English proficiency quite high, their actual academic language proficiency is much lower. Learners make no distinction between interpersonal language skills and academic language skills. For the purpose of academic institutions, academic proficiency is of great importance (see Parkinson, 2002). It is important to make learners aware of the difference between interpersonal and academic skills in a way that will spark action. Feedback (presented in the way proposed in this study) could assist in this.

1.3.5 Public expectations

James (1997:26-27) indicates that the public is aware of issues on "bad language". Evidence for this statement is provided by the large amount of "watch your language columns" in innumerable newspapers and magazines throughout the world. James argues that the publishers of magazines and newspapers do not give media coverage to minority concerns and the existence of these columns prove that the editors are responding to a public concern (also see Milroy and Milroy, 1992:30).

1.3.6 Personal Development

Writing can facilitate personal development. Katznelson, Perpignan and Rubin (2001) claim that in writing courses, on top of all the changes in the writing proficiency of learners, certain other changes also occur in students. These changes develop along with the writing outcomes of the course and are termed "by-products". Apart from the "normal" changes occurring when learning to write, such as better grammar and vocabulary development, writing also develops other characteristics and abilities of students. Katznelson et al. (2001:151) found that students experience training in writing as helping them to express themselves more coherently. It also helps them to follow a logical line of thinking in other languages and other disciplines. In addition Katznelson et al. (2001:152) found the following "by-products" students developed along with their writing skills:

- Other skills:
    - Oral presentation techniques
    - Analyzing scientific papers
    - Use of a specific on-line internet writing tool (computer skills)
- Affective outcomes:
    - Overcoming a fear of speaking
    - Starting to like the English language
    - Increased self-esteem
- Teamwork:
    - Generating ideas through group work
- Learning the Meaning of Learning
    - Feeling of belonging to the academic community
    - Feeling committed/dedicated to own work
    - Paying attention to the writing and not to the grade
- Listening to others:
    - Accepting positive criticism
- Increased Genre/Discipline/Media knowledge:
    - Encountering different genres of writing from different disciplines
    - Comparing the language of literature to the language of film
- Critical thinking:
    - Achieving a critical perspective
- Broadening of Knowledge Base:
    - Using opportunity to learn about work of other students

These "by-products" are all relevant to academic development (Webb, 2002:56) and writing should therefore not be taken lightly. Writing fulfils a host of academic functions, is important for the public and can even assist in interpersonal development. It is obvious that writing is important and it is almost unthinkable that such an important skill has shortcomings. But writing – and more specifically student writing – is far from perfect. There are numerous problems.

1.3.7 Problems in writing

In spite of the importance of writing as pedagogical and communicative tool, writing skills in South Africa are far from optimal. Van Wyk (2002) found the following problems in student writing:

- Poor sentence control
- Lack of an awareness of audience
- Clarity of expression and organisation
- Lack of relevance and logic

Van Wyk (2002:227) explains the problem as one of "structural irregularities that do interfere with communication". Van Wyk found that students' writing "can be difficult or impossible to understand" since the sentences are "out of control" and do not adhere to "any standard sentence structure". Webb (2002:54-55) gives an example of student writing, explaining that if one ignores the poor sentence construction, one will discover that the student had a reasonable grasp of the substance being examined, but the quality of the language was so bad that there "was no way in which these students could have succeeded in the examination." This is troublesome. If students

cannot express themselves well enough to present their knowledge in a comprehensible way, they may find that their progress at tertiary level is affected adversely (see Webb, 2002:55-56).

1.4 Conclusion

Writing is important for much more than only effective communication, but even effective communication is not always achieved in writing due to a number of problems. Therefore it is very important to ensure that the teaching of writing is optimised so as to reap all the benefits of effective writing.

1.5 Problem statement

This dissertation attempts to answer the following questions:

1. Is it possible to create standardized written feedback on second language writing?
2. Can standardization of feedback ensure (a) clarity and (b) consistency?

1.6 Objectives

1. To use best practice in feedback to create a set of tags for use in a computer marking interface.
2. To test and ensure that the feedback tag set is:

2.1. Clear

2.2. Consistent

1.7 Approach: how to provide more effective feedback

The problems with feedback are a lack of consistency and transparency, coupled with a lack of action based on the feedback. This dissertation argues that the problems facing feedback can be tackled with the use of modern technology. Feedback can be provided in a CALL (Computer-assisted language learning) environment in a consistent manner. Catering for the individual can then be done after a standardized process has been applied. Such a CALL environment is found in the IWill-system created and tested by Wible et al. (2001) for schools and universities in Taiwan.

In this system, students' essays are submitted, marked and archived for corpus use. The system solves two of the problems noted above – those of labour/time and noticing recurring errors. According to Wible et al. (2001:303), the system saves time in the marking process and is easy to use. The lecturer uses a special computer interface and a "comment bank" to provide feedback on the learners' writing assignments. This comment bank contains some frequently used comments that are supposed to be clear and appropriately detailed to be easily understood by the students. This is an example of a correction code – the method found by Spencer (1998) to be the most effective way of providing feedback, as well as the method preferred by students.

The computer is also able to archive marked errors and can create statistics, clearly showing the weaker areas of the different students. The analysis of these errors could be used to decide how to use teaching time more effectively (Ellis, 1996:48).

However, the IWill-system does not solve the problems of inconsistency and the way learners understand feedback. IWill has **one very important limitation.** Chapelle (2003:xi) indicates that a lacuna exists in the CALL field. This gap is that knowledge on SLA is not always used to design CALL systems. This is exactly what happened in the design of IWill. Every teacher or lecturer had the opportunity to create the entries in their own comment bank. The problems of inconsistency and student uncertainty due to vagueness were not addressed. By first doing research on what feedback exactly is needed in a marking system, and then designing the system, important insights from both SLA and CALL could be incorporated. The creation of a standardised correction code is the first step towards answering questions on consistency, accuracy and clarity and effectiveness.

It is very important to use a standardized set of comments for such a system as this would:

- make the system consistent.
- make it easier to explain the feedback method and tags to students and lecturers – i.e. equipping lecturers to mark more thoroughly and effectively and empowering students to understand the feedback better and act on it.
- make it easier to test the effectiveness of the supplied feedback.
- enable further research by the creation of a corpus.

This standardised correction code should be one compact, yet thorough list of feedback tags. This list should conform to the standards of SGML (Standard Generalised Markup Language) in order to enable computer processing in later stages (see Granger, 2002:10).

1.8 Overview of the study

This dissertation reports on a project aimed at achieving the above: A computerised marking interface is designed (similar to IWill). The computer interface makes use of a standardised set of feedback tags to facilitate consistency – the main aim of this dissertation. The tags are created by researching best practice in feedback and testing the effectiveness thereof. To ensure that action based upon the feedback, is taken, the system will provide individualized remedial exercises for students, based on the individual student's shortcomings. The remedial exercises fit into the next stage of the project, and fall outside the scope of this dissertation. However, it is mentioned here in order to sketch the context in which the feedback tag set was created.

Chapter two presents a literature review on the different techniques used to provide feedback, in order to find the do's and don'ts of feedback. The chapter looks at different feedback techniques, establishing what effective practice is. This percolates to a checklist of ideal feedback, which is later used to evaluate the feedback tag set.

Chapter three describes the methodology followed in order to create and test a feedback tag set. Two different methodologies were followed for different parts of this study: One for establishing what lecturers mark and creating the tag set (an empirical study), and one for testing the effectiveness of the tag set.

Chapter four describes what lecturers focus on when marking. The findings reported on in this chapter gives an indication as to why feedback is sometimes claimed to be ineffective – the guidelines for effective feedback are not adhered to.

Chapter five presents the provisional tag set used in the experiment to test the tag set. Some design decisions are explained in context of the system it will be used in.

Chapter six discusses the results of the experiment conducted in order to test the effectiveness of the created tag set. The results are positive, indicating that standardised feedback by means of a computer is possible to an extent.

Chapter seven presents the final version of the tag set. This version could still be refined, but it is suitable for current purposes.

The final chapter indicates where follow-up research and further development will be necessary. Since this is a new system and a new approach (at least in South African context) to feedback, some problems are anticipated. These are briefly discussed in chapter eight.

2. CONCEPTUALIZING AND INVESTIGATING FEEDBACK

If one starts to look for information on feedback on writing, a simple search will provide numerous articles and books. Some advocate the merits of feedback, some say it does not work, while still others say it may work if one focus on small subsections of writing and not on the total system. Accepting that none of these articles were written or researched without having elements of truth in them, how is it possible to have so many different views on the same subject? My interpretation of the situation is that all of these articles do have a contribution to make to feedback practice. Instead of looking at the "right" or "wrong" of feedback, I investigated the "why" and "how" of feedback. If an article says that feedback did not work, then there should be a reason for it. Likewise, if a researcher found feedback to work, then something had been done right.

By carefully analysing the different techniques and objectives of feedback, it could be possible to identify best practice when providing feedback. This chapter is therefore aimed not at fuelling the debate whether or not to provide feedback, but rather at taking feedback as a given and to establish how to provide feedback in the most effective way possible, in order to meet the objectives of "best practice", "clarity" and "consistency" as stated in chapter one. To answer those questions it is necessary to first establish what is meant by "feedback" and "error". The definitions are discussed in section 2.1.

In section 2.2, the reasons for identifying errors in the first place are also touched upon, after which the arguments behind providing feedback are explained. Problems with feedback are also touched upon, providing a reason why feedback should apparently not be provided.

Section 2.3 deals with the "how" of feedback, looking specifically at techniques currently in use and what experienced teachers do. This ultimately leads to a checklist of ideal feedback.

2.1 Feedback and Error

2.1.1 What is feedback?

When discussing feedback, it is important to take into account the context of the situation. In the L2 writing context, we are talking of feedback in the context of a learning situation. Feedback in the sense of everyday communication is not relevant here. Feedback in the context of second language acquisition and writing education, is.

At present, the literature on feedback seems to use two main (overlapping) frameworks for feedback:

- Any response or "evidence" by any reader.
- Any correction by a lecturer (error treatment)

These two views will now be discussed, whereafter a combined view of feedback in the learning process will be illustrated.

2.1.1.1 Feedback as any response or evidence

The approach to feedback as response or evidence is an all encompassing view of feedback, reminiscent of feedback in any communicative situation, with the exception that we are dealing specifically with writing here. In this approach, which is a popular one, it seems as if any response (both verbal or written) from a lecturer/teacher or reader on a student's text, is seen as feedback. It seems that most students and the broad public, as well as some teachers, feel that this category constitutes feedback. Evidence for this is found in the studies by Hyland (2003, 1990 and1998), Lyster and Ranta (1997) and Askew and Lodge (2000).

Hyland (1998:261), for example, found that her students considered all interventions in the text as feedback and did not differentiate when using feedback to revise their essays. She also found that teachers dealt with both meaning and grammar related issues at the same time when responding to the student texts.

One of the problems of linguistics is that one instance of something may be interpreted in many different ways. Even something as ordinary as feedback can be viewed differently, depending on who is using it. Lyster and Ranta (1997:38) indicate that the way competent speakers react to learners' language errors is seen as:

- Negative evidence (from the viewpoint of linguists)
- Repair (from the viewpoint of discourse analysts)
- Negative feedback (from the viewpoint of psychologists)
- Corrective feedback (from the viewpoint of SL teachers)
- Focus on form (from the viewpoint of SLA-studies)

Lyster and Ranta (*ibid.*) claim that the various views is the reason why neophyte L2 teachers find so little in the research literature to help them deal with the very practical issue of what to do when students make errors. It seems that in order to use a definition of feedback, one should first situate yourself in a specific field, identifying your purpose.

The purpose of feedback in a learning situation will obviously be to enhance the language use and writing skill of learners. The way in which to go about in doing so is unfortunately a more difficult question to answer and this is where the distinction between evidence and response comes into play.

Evidence is seen as information provided by "knowers" of the language, to learners of the language about the latter's attempts at reproducing it (James, 1997:239-240). This information (or evidence) is used subconsciously by the learners to revise their interlanguage (Ellis, 1994:702). Evidence can be direct or indirect, and positive or negative. An example of direct positive evidence will be a "knower" of the language

telling a learner that he/she has just produced a "good" sentence. Foreign language learners make use of all four types of evidence (James, 1997:240) in order to acquire a language.

Response can be seen as a sub-category of evidence, or even as a technique of evidence. Spencer (1998:10) argues that feedback is a sub-category of response, but many of the roles and effects she ascribes to response can also be interpreted as feedback. She describes feedback as not enough on its own, as response should result in a type of conversation between the learner and lecturer. Important to note here is the statement that "response is only as effective as the students' ability to grasp what has been conveyed, internalise the knowledge, and use it constructively in the learning process". This could then also be seen as **feedback-as-communication.** This response then, is obviously a tool for learning, and that is why I see the distinction as inconsequential.

The attributes of response (which may just as well be attributes of feedback) mentioned by Spencer are:

- Response should contain information that provides the performer with direct, **usable insights** into current performance, based on tangible differences between current performance and hoped-for performance.

Response aims at:

- helping the learner consciously to **identify and solve** writing difficulties;
- enabling the student to take any **corrective action** that is necessary to improve his or her writing;
- **stimulating practice** in the sense that students must apply the response to the revised work;
- allowing for transfer of internalised **problem-solving skills** to other writing;
- developing **metacognitive skills**;
- promoting the **independence** of the learner;
- increasing motivation, reducing anxiety, promoting self-esteem, encouraging risk-taking, fostering tolerance of ambiguity and showing empathy, thereby actively **reinforcing the affective factors** associated with language learning;

- enabling the learner to internalise the gist of the commentary thereby raising student judgement and promoting **self-evaluation** skills;
- modelling effective reading strategies that assist the learner to become a successful **critical reader** of his or her own writing;
- providing **models** of expert writing;
- **promoting learning**. Response should never be restricted to evaluation, but should strive to be part of the teaching program.

Spencer (1998:9-10)

It is clear that Spencer is aiming at response-as-communication – enabling the students to *learn* from the teacher commentary. The difference between feedback-as-communication and response-as-communication is inconsequential if both have the inherent aim of enhancing learning.

2.1.1.2 Feedback as any correction by a lecturer

If feedback is seen as evidence or response, it could be in the form of a correction – but not necessarily. This correction could come from anybody. However, feedback could also be seen as a correction (of any error) by a lecturer. In fact, Moletsane (2002:21) claims "it is the responsibility and duty of second language teachers to help students overcome their problems regarding the use of the target language when writing. They do this by correcting students' written work". This statement of Moletsane may be a bit blunt and will be addressed later, but it is important to note that Moletsane is voicing a popular view.

The problem is that **feedback-as-correction** is an area where there is no consistency at all in the literature. Feedback in the form of a correction may be evidence or response, but the problem is that a correction does not necessarily lead to learning. James (1997:236) explains that the term "correction" has been used in three senses:

1. Informing the learners that there is an error but leaving them to discover it and repair it themselves. James calls this **feedback** – the giving of knowledge/ results in the broadest sense, telling people whether their utterance or

understanding is right or wrong. This fits into the view of correction proposed by Lightbown and Spada (1999:172) as explicit corrective feedback.

2. Providing treatment or information that leads to the revision and correction of the specific instance of error (error token) without aiming to prevent the same error from happening again. The corrector can specify how or where the error is and give a hint on how to go about in correcting it. The correction is aimed at product enhancement and James refers to this as "**correction proper**". It is almost a type of proofreading.

3. Providing learners with information that allows them to revise or reject the wrong rule they were operating with when they produced the error token. The result would be to induce learners to revise their mental representation of the rule, so that this error type does not recur. The purpose is to improve the process for future productions rather than merely improve the present product cosmetically. This is **remediation** and this is when learning can take place.

James (1997:237) explains that feedback could be seen as an overture to correction, while correction is in turn an overture to remediation, "since having given feedback you can decide whether to stop there, or to enlarge your treatment by going on to correction and then remediation as well." It seems reasonable to argue that in a teaching situation, it is preferable to continue to remediation.

James then draws a further distinction between surface correction and deep correction. This is an important distinction. Surface correction is "putting right", but putting right does not address the source of *why* some bit of language was produced the way it is. Deep corrections address the source of the problem, prompting learners to reorganize their cognitive structures. Once deep correction has been achieved, learning takes place. It is in light of this distinction that the above statement by Moletsane could be judged as being a bit blunt.

James (1997:239) uses the above definitions of feedback, correction and remediation to explain how he sees the process and use of indicating errors: "Slips" are self-correctable without feedback, "mistakes" are self-correctable with feedback and

"errors" are not self-correctable by simply providing feedback. What is needed is correction proper, which would in effect be remediation. The remediation would stem from the fact that correction proper would imply explaining the target language structure to the learner. Of course, the value of explicit instruction is also a much debated topic, and techniques to effect remediation will have to be carefully selected.

The distinction made by James (1997) among the different levels of error correction (or feedback then) is in line with the discussion of "feedback" in Ellis (1996:584), with the exception that Ellis also makes room for affective feedback – motivational support that interlocutors provide each other with during an interaction.

In spite of all the different approaches and distinctions mentioned above, it still seems as if many distinctions are cosmetic, with one thing in common: all aim at providing the learner with information that would enable him or her to adapt to a certain standard of language use, with the teacher as "knower" of what that standard pertains. In short, there is communication between the text creator and some level of audience to optimize written communication. It seems the main difference between all the above approaches lies in the amount of information provided. An all-out correction may provide less information than a well-structured explanation, and it may provide less learning opportunity as well. The amount of information provided should therefore assist in the communication process between the text creator and the audience. In this case, the communication aimed at the text creator, also has the aim of facilitating his/her learning. This is a tall order for feedback.

Communication is never a one-way process. In the writing as process approach (see Krapels, 1990), much emphasis is placed on the meaning of the text and the fact that the meaning of a text is not simply created by the writer *but co-created by the reader.* Feedback given by the reader should be indicative of the way an audience would experience the text. The problem with this is that an audience of average readers would most probably not recognise all the errors in a text, and feedback is still a tool with which to provide learning as well.

Feedback is ideally a "ping-pong game" (Askew and Lodge, 2000) of comments going back and forth between the reader and writer until near perfect communication is attained. This is an almost unreachable ideal as writing in most academic instances

is not an extended cycle. Instead of looking at what is ideal then for all situations, one should rather look at what type of feedback is the most relevant for the specific purpose. In the context of a second language writing classroom, the feedback definition that should be used is not dependent on the distinction between errors, mistakes and slips. The fact of the matter is that writing is a slow process where plenty of thought and time is available in relation to speaking. If there is any "error" or "slip" in a text, this would at the least still be an indication of poor proof-reading skills or even the error of inconsistency. Even minor errors can hamper communication.

A summary of all of the above would lead to the conclusion that feedback is part of a communicative process, in which a knower of a language (teacher/lecturer) provides information on different levels of language use to a learner of the language. Feedback may be provided in different formats and may differ in intensity and purpose.
For the purposes of this study, the feedback would be given by a lecturer or teacher and may even be an all-out correction of the error. The process of feedback can be illustrated graphically as follows:

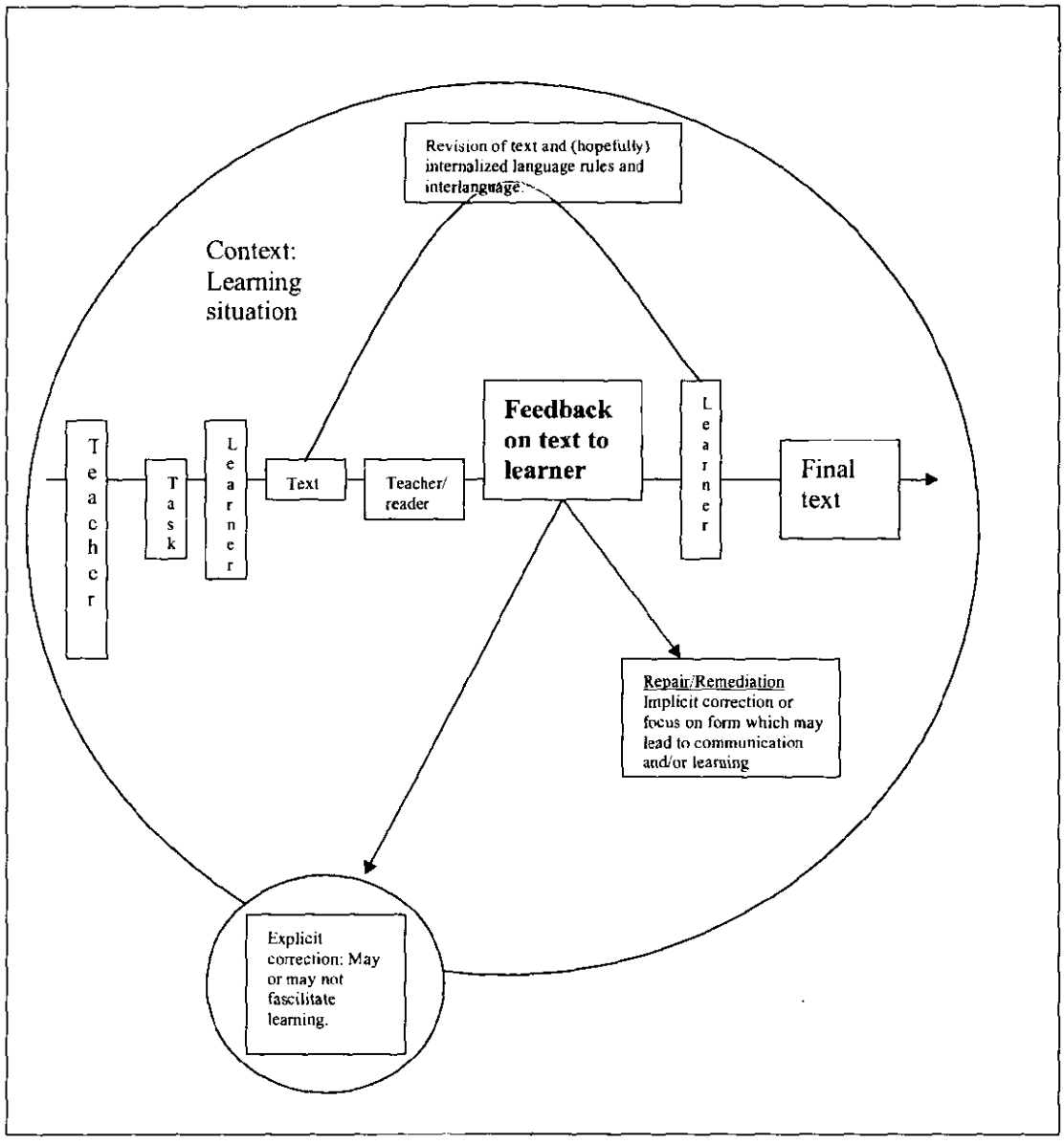


Figure 2.1. The communication timeline of feedback on writing in a learning context.

Within the larger context of a learning situation, a teacher or lecturer will give a writing task to a learner. The specific situation will determine the goal of this writing

task. The task may or may not have a specific language educational motive. It may only be to test a learner's knowledge of a specific topic, or it may be to evaluate the learner's writing competence, or even to test a specific style of writing. It may also be to see how well learners have understood a teaching objective such as mastering the passive voice for example.

The learner will get the writing task either orally, or in written form and may or may not know the specific goal of the exercise. The learner then creates a text, which may or may not contain certain errors, and gives it back to the teacher, or maybe a fellow student (reader) for their commentary.

The reader or teacher then provides feedback on the text. This feedback will differ in method and the specifics it focuses on. This feedback may be either (or a combination) of:

- repairing feedback leading to communication between the learner and reader, or
- it may be explicit correction, which has dubious value as a teaching tool.

In some teaching situations, this is where the process for a specific text stops. The first version of the text was the final version. In other situations (specifically the process writing approach) the learner will have to revise the text and give it back to the reader or teacher. During this revision process, the learner may learn something about writing or even about his/her internalized rules of language. Feedback can be considered effective if this step happens.

Does this mean that feedback can only be effective in the process writing approach? No. In situations where the first version of a text is the only version, the possibility still exists that students may use the feedback to revise their internalized rules of language. In effect, they will have to refer back to the feedback on a previous text, to enhance a different future text. This can get very difficult – especially if the student never looks at the feedback. In situations where a process approach to writing is not practiced, something should be done to ensure that students still utilize the feedback. The proposed marking system reported on in this dissertation, will eventually

incorporate a system to automatically create personalized exercises for students, based on the feedback in their respective essays. This is a step in the direction of ensuring that feedback is used effectively.

For the purposes of this study then, a working definition of feedback would be:

> **Feedback constitutes any mark by an external reader on the text. The mark may indicate something that is considered to be wrong, or something that is considered less than optimal. Feedback may also indicate instances where the reader is satisfied or impressed by something in the text.**

2.1.2 What is an error?

Feedback (as explained above) has the function of looking at errors. For obvious reasons, the term "error" then also has to be defined. The term "error", although seemingly simple, has no fixed definition in the linguistic literature, although there is much greater conformity than for the definition of "feedback". This section aims to arrive at a working definition of "error". I will first indicate what linguistic dictionaries see as errors, then what theorists see as the reasons why errors occur. Thereafter an example of error classification will be discussed and the use of errors in linguistics will be pointed out. This will lead to a working definition of error for this specific project.

2.1.2.1 Dictionaries definitions of "error"

In order to illustrate the lack of conformity of the definitions for error, a few dictionary definitions are considered here:

An Encyclopaedic Dictionary of Language and Languages.

Errors in error analysis are seen as unacceptable forms produced by someone learning a language, especially a foreign language. Errors are considered to be systematic, governed by rules, and appear because a learner's knowledge of the rules of the target language is incomplete. They are of particular interest in linguistic research because they provide evidence about the nature of the language learning process. A contrast is drawn with mistakes, which are unsystematic features of production that speakers would correct if their attention were drawn to them (e.g. those arising out of tiredness or a lapse of memory).
(Hartmann and Stork, 1997:116).

A Dictionary of Linguistics and Phonetics. 3rd ed.

In language teaching and learning, error analysis is a technique for identifying, classifying and systematically interpreting the unacceptable forms produced by someone learning a foreign language, using any of the principles and procedures provided by linguistics. Errors are assumed to reflect, in a systematic way, the level of competence achieved by a learner; they are contrasted with 'mistakes', which are performance limitations that a learner would be able to correct.
(Crystal, D, 2003:98)

Encyclopaedic Dictionary of Applied Linguistics

"Errors came to be viewed as a reflection of L2 learners' mental knowledge of the second language: their interlanguage grammars" (Johnson and Johnson, 1999:111).

"The description of the errors involved is not the same thing as an explanation for them" (Johnson and Johnson, 1999:113).

"...in the case of error taxonomies it has often been difficult to determine why an error should be classified in one way rather than another" (Johnson and Johnson, 1999:113).

"Studies on error evaluations indicate the necessity to consider exactly who is doing the correction; considerable differences exist between native speaker and non-native speaker teachers as regards the focus of corrections" (Johnson and Johnson, 1999:115).

> "An error is a breach of the language's code, resulting in an unacceptable utterance; with L2 learners this might occur because the learners have to yet internalize the formation rules of the code. Mistakes or lapses are the result of some failure of performance. They occur when the language user (who might be a native speaker) makes a slip such as a false start or a confusion of structure... The above use of these terms is the generally accepted one, though Corder draws the distinctions differently. He distinguishes between 'lapses' (performance failures) and 'mistakes' which are seen as the result of inappropriate usage. This usage has not become common" (Johnson and Johnson, 1999:117-118).

From the above few definitions from dictionaries, it should be obvious that there is no unambiguous definition for error, not even taking into account the distinction between slips, mistakes and errors, mentioned above. Ellis (1996:50-54) indicates that the difference between errors and mistakes is that errors occur due to lack of knowledge and mistakes are a failure to perform existing competence. The distinction would be very difficult to make when providing feedback, since tracing an individual learner's patterns through huge volumes of data is not possible.

In order to study the way lecturers and teachers react to errors meaningfully, it would be important to have an unambiguous definition for what is meant by error. For the purposes of marking student essays, errors would have to be anything that is not standard (or "wrong"), but this would open up the door for subjectivity (see Spencer 1998), especially since some problems in South African English may be attributed to the teachers (see Buthelezi, 1995 and Van der Walt and Van Rooy, 2002).

The distinction made between error, mistake and slip is useful when trying to understand how learners acquire a language. I will touch on this later in this section. The distinction is however not so useful for teachers who simply want to mark an essay. To distinguish whether something is indeed a slip and not an error would mean that intensive research would have to be done on the issue. This would not be useful in a marking environment. If there is anything that is not correct, then it is an error. Be it a slip or a mistake, the fact remains that it is not optimal language use in that it is a violation of the conventions of the target variety. However, one should be careful of simply focusing on incorrectness, as this would not be effective feedback either. The ability to see something that is not inherently wrong but could have been better, as an error as well, would provide much more opportunity for learning.

2.1.3 Why do errors occur and what do they tell us?

From the above definitions it is obvious that errors are "unacceptable" in that they break the rules of the language. However, errors are not created on purpose – unless for poetry or satire. Learners strive to use the correct forms of the language. The point of departure of error analysis is that by analyzing errors many valuable insights can be made into the "how" of language acquisition. Many researchers have therefore tried to look at the reasons errors occur. The purpose of this study is to improve feedback to such an extent that learners can improve their writing. The learners are therefore learning from their errors in order to acquire their language better.

One of the first ways of investigating L2 acquisition was through error analysis. This became popular in the 1960's and 1970's (see Corder, 1973). Early error analysis focused on determining the extent to which L2 acquisition was the result of L1 transfer or of creative constructions:

- Transfer: errors that mirrored L1 structures was seen as evidence of transfer
- Creative construction: the construction of unique rules similar to those which children form in the course of acquiring their mother tongue. The presence of errors similar to those observed in L1 acquisition is indicative of creative construction (Ellis, 1996:19).

It was found that:

- Many errors are caused by transferring L1 habits to L2.
- Learners also contributed creatively to the process of learning.
- Learners go through stages of acquisition – the nature of the errors they make varies according to their level of development.

But Error Analysis as practised in the 60's and 70's was not a precise tool as it could not show when learners resorted to avoidance, it ignored what learners could do correctly, and it was unclear what an exact definition of "error" should be (Ellis, 1996:19).

More recent research has a different view of why errors occur, stating amongst other things that errors may be the result of:

- Insufficient input
- Materials and style of presentation by the teacher is not suitable to the particular student
- Problems with discrete language points
- Inefficient learning strategies
- Carelessness caused by lack of motivation
- Loss of interest
- Personal problems outside the classroom
- Poor attention in class
- Irregular attendance
- Particular macro-skill problems

(See Moletsane, 2002 for a more detailed discussion.)

The above is simply an example of an explanation for errors. There are other theories of why errors occur (for example Ellis, 1996:57-63), but for the purposes of this study, the **types** and **description** of errors are more important than their explanation. Note however that some of the findings above could be used as evidence in the case for feedback. For example, if errors are caused by the transfer of L1 habits to the L2, it would be important to point out to the learners where they are in fact "speaking Tswana in English", thereby obscuring meaning. This type of "response" to learner errors may be very similar to the "negative evidence" provided in real-life situations. (See discussion in 2.1.1.1.)

2.1.4 The classification of errors

2.1.4.1 Why classify errors?

As seen in the above section, there may be numerous reasons errors occur. Depending on the specific field of reference, there are various reasons why researchers would want to know why errors are created. There are also numerous classifications of errors that try to order errors on the grounds of *why* they occur. This is very difficult and often very subjective. In some instances it is very useful to know why errors occur, but for the purposes of providing feedback it is more important to classify errors in terms of the categories they fall into, so that something can be done to correct the problems in that specific category. For example, if there are many errors under the category "syntax" then the lecturer would know that he has to pay more attention to syntax. The reasons for problems may go back to a large variety of sources and it would be virtually impossible to locate all those reasons. If the lecturer does isolate those reasons, there is also a chance that he would not be able to do anything about the source, other than doing corrective work anyway.

Apart from the reason mentioned above, James (1997:96-103) says that it is possible to classify errors simultaneously according to a number of criteria but there are three purposes for describing errors:

1. To make explicit
2. To enable counting
3. Categorization

These may equally well have to do with education or research. Moletsane (2002:13) says that errors provide evidence of the system of the language that the learner is using or has learnt at a specific stage in the acquisition of the L2. (Also see Corder, 1973 and Ellis, 1994 and 1996). Moletsane explains that errors are significant in three different areas:

- For teachers: how far have the learner progressed and what remains to be learnt.

- For researchers: evidence of how languages are learnt or acquired and the strategies employed by learners to discover the systems of the language.
- For learners: errors are a device learners use to learn.

James (1997) looks at error description in terms of Error Analysis, but all three of James' reasons for describing errors fit the provision of feedback as educational activity as proposed by Moletsane (2002) as well. Feedback aims at making errors *explicit – to draw attention to the learners'* errors. This is known as consciousness raising, even though the way some educators provide feedback is not effective for this purpose. (For more on consciousness raising, see section 2.2.2.2.7.)

Errors could also be counted and categorized, so that a lecturer would know on what to work with his students. If students routinely make the same error, it makes sense to try to remediate the problem. Errors seem to be a readily available resource for this type of problem analysis.

2.1.4.2 Example of Error Classification

To fully comprehend the classification of errors, James' (1997) explanation of error classification will be discussed here. This example is an extremely condensed version of James' classification and for more detail, the reader is advised to consult the book. Keep in mind that the classification of errors is closely linked to the definition of error. Also keep in mind that this discussion is simply intended to highlight the amazing amount of information that could be gleaned from learner errors and that this error classification is not the one used for the creation of the tag set. Instead, the error types and categories in this study were classified in terms of a theory of language (Halliday and Matthiessen, 2004 and Givón, 1993) as explained in detail in Chapter five.

2.1.4.3 James' classification of errors

James (1997) classifies errors according to sections and levels. He then distinguishes between errors in text, lexis, grammar, discourse and pragmatics. Finally, the reasons why errors occur are also taken into account. These distinctions will be elaborated on shortly. It is important to note that different corpora have been annotated using different classification systems. Since some of the above categories may overlap, the purpose of analysis may prescribe which categories are used in the classification of errors for a specific project. It will be very difficult, if not impossible, to incorporate all of the above categories into one classification system.

2.1.4.3.1 Sections of errors

James (1997:100) identifies Fitikides's *Common Mistakes in English* (1936), although dated, as the prototype of error dictionaries. James focuses pertinently on error analysis so this classification system could be seen as a prototype for corpus linguistics. It divides errors into five sections:

1. misused forms:
    a. wrong prepositions
    b. misuse of tense
    c. miscellaneous "un-English expressions"
2. incorrect omission
3. unnecessary words
4. misplaced words
5. confused words

Taking these sections as a starting point, a more elaborate classification system can be created. James (1997:95) writes that a system used to describe errors must be:

1. well-developed and highly elaborated
2. as simple, self-explanatory and easily learnable as possible

The system by Fitikide is not at all elaborate. The system used in this study for standardizing feedback is not elaborate either due to a compromise that had to be

made to accommodate everyday use. It is not as elaborate as a detailed classification system such as one that could be used in a major research project. In this case the purpose of use defined the elaborateness of the error classification system.

2.1.4.3.2 Levels of error

The above "sections" of errors gives us a certain starting point, but it is not elaborate enough and does not provide for errors of discourse, for example. A finer analysis of errors can be (needs to be) made by means of levels. James (1997) dedicates a whole chapter to the different levels of error. He indicates that errors are classified in respect of their modality, medium and level. This is graphically illustrated in the flow-chart below.

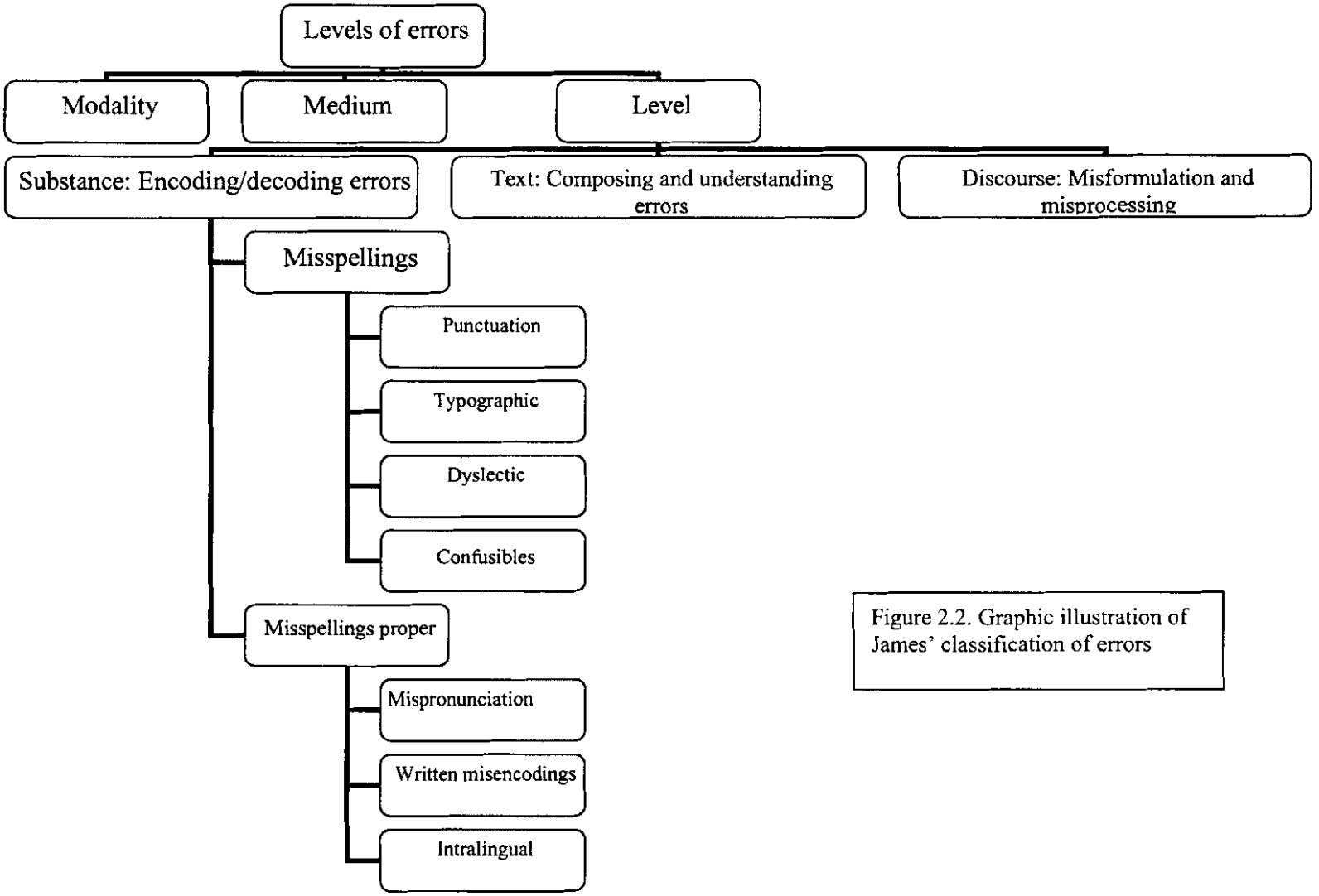


Figure 2.2. Graphic illustration of James' classification of errors

For a complete explanation of all the different levels, see James (1997). By combining modality and medium together we can specify which of the four skills the learner was using – speaking, writing, listening or reading. In further combining the above classification with the categories of omission, redundancy, misselection, misordering and blends, the range of errors that can be described gets to be much more versatile (James, 1997:130).

James (1997:140-154) in describing the difference between text errors and lexical errors, mentions spelling errors. The categories start overlapping here in that what James calls "distortions" are not that far different from "dyslectic misspelling". For the purposes of this study, it also does not make sense to distinguish too finely between spelling errors, simply for reasons of space in the final marking tool. (See chapter five for more detail). This will be discussed under error diagnosis.

James' category of discourse errors deals with coherence only, whereas the linguistic categories used in this study in chapter five use a much broader definition of discourse. Those which James (1997:164) labels "pragmatic errors" (social taboos and values of society), will also fall under the broader category of discourse in chapter five of this study. The reasons for the difference in classification no doubt have to do with the purpose of the classification. This brings me to the important distinction by James (1997) of errors categorised according to the reason they occur. This is known as "diagnosis-based categories". (See 2.1.3.)

2.1.4.3.3 Error Diagnosis

With regards to the reasons why errors occur, there are four major categories:

- Interlingual
- Intralingual
- Communication-strategy
- Induced

James (1997) dedicates a whole chapter to the reasons errors occur (diagnosis). Here one can once again distinguish between primary and secondary diagnosis. Primary diagnosis simply explains "why" errors occur, while secondary diagnosis explains the forms that errors assume (James, 1997:177). Though it may be useful to know how errors are formed in the remediating of learners with consistent problems, the relevance of such a diagnosis does not extend to this study. It will however be relevant for further research (see chapter eight).

Apart from noticing the errors, James (1997:204-205) indicates that noticed errors have to be evaluated to get priorities right. The smallest error need not be pounced upon, but the errors that really matter should not be ignored. In error analysis, there should be a set of criteria to determine the "priority" of the error. The criteria for error gravity according to James (1997:206-234) are:

- Linguistic criteria
    - Rule infringement
    - Rule generality
- Frequency
- Comprehensibility
    - Intelligibility
    - Communicativity
- Noticeability
    - Frequency
- The irritation factor

James (1997:209-211) indicates that to some extent the gravity of the frequency of errors is debatable as it is difficult to determine how the learners go about producing them. What is important though, is the error density, as even small errors in quick succession can jeopardize the intelligibility of a sentence. This is also an area for further research for the current project (see chapter eight), but at this stage it is important to note only that "frequency" has two different perspectives. There is in the first place, the frequency of the error and in the second place, the frequency of the construction in general. An error occurring in a very rare construction is not of such importance as an error occurring in a more frequently used construction. If the error

occurs frequently in a much used construction, it may indicate fossilization, which will affect the course of action taken from there to address the problem.

To further complicate the matter for the researcher, the data of error analysis comprises not two but three more variables to consider:

1. What the learners said/wrote;
2. What they were attempting to say;
3. What the native speakers would have said or written.

Due to the versatile nature of language, all three the above may be difficult to deduce and may influence the actions taken on account of them.

2.1.4.3.4 Conclusion on error categorization

The lengthy discussion (though still only an extremely brief overview) should prove the immense number of variables to take into account when working with language errors. It is therefore a completely unrealistic expectation that teachers or lecturers will be able to identify all errors, or know how to react to them. Before writing error classification off as a superhuman feat that will never realistically be achieved, it should be remembered that although it is a big task, most big things can be broken down into smaller parts. If these errors are categorised to an extent, it can be used for error analysis, which is still a difficult endeavour, but at least it is a start.

2.1.2 The importance of Error Analysis (EA)

Why go through all the trouble of categorizing errors? Errors have to have some use. I have indicated in section 2.2.1.3 that errors have a use in describing learning and annotating a corpus. An annotated corpus is a handy tool for a researcher trying to find out with what his/her students has trouble. By doing an analysis of the errors in the corpus, the researcher would be able to optimize the way language is taught. He/she could do this by looking at what is taught and analyzing how successfully that teaching is converted into learning. This is Error Analysis.

James (1997:1-3) indicates that a language error could be seen as an unsuccessful bit of language and Error Analysis is the process of determining the incidence, nature, causes and consequences of unsuccessful language. Error analysis is a branch of applied linguistics and not of linguistic theory. It is the paradigm that has replaced Contrastive Analysis. Error Analysis involves first independently describing the learner's interlanguage and the target language and then comparing the two. Initially it was thought that error analysis could be used without considering the effects of the mother tongue, but when this proved to be not the case, the scientific community gradually began to favour transfer analysis. Error analysis is compatible with transfer analysis as the errors caused by mother tongue influence could also be categorised. James (1997) sees transfer analysis as a sub-procedure applied in the diagnostic phase of doing error analysis. It is important to note that error analysis is a methodology for dealing with data, rather than a theory of acquisition. Therefore, EA can be used in studies on SLA, but is not necessarily part of a theory of SLA (James, 1997:7).

According to Cagneaux, Dennis and Granger (1998), error analysis has the following limitations:

1. EA is based on heterogeneous learner data;
2. EA categories are fuzzy;
3. EA cannot cater for phenomena such as avoidance;
4. EA is restricted to what the learner can NOT do;
5. EA gives a static picture of L2 learning.

The first two categories are methodological. The data used in EA studies were not collected under well-defined conditions so it was difficult to replicate the studies. The categories used in EA had the following problems:

- Often ill-defined;
- Rest on hybrid criteria;
- Involve a high degree of subjectivity;
- Error typologies often mix two levels of analysis.

(Cagneaux, Dennis and Granger, 1998:165).

With the advent of corpus linguistics, some of these problems are being rectified, mainly due to standardization. A more standardized method of analysing and classifying errors provides better defined categories and better defined criteria and as a result a smaller level of subjectivity.

2.1.3 Working definition of error

After discussing the definition and merits of errors at length, it should be clear to the reader that the definition of an error is a very sensitive issue. The definitions connected to "error" are also in a sense too limited in scope to apply to feedback as a total system (a system to enhance learning). For those reasons, I have to resort to a very broad working definition.

The definition of error used in this study will be: **An error is any instance in a text which is incorrect language use or language use which is not inherently wrong, but which could have been better.**

This definition links up sufficiently with the working definition of feedback, to make it useful for the current study.

2.2 Why identify errors and why/why not to provide feedback.

In this section, the reasons for identifying errors will be touched upon briefly, whereafter a discussion will follow of the arguments against the provision of feedback. The discussion will mainly focus on the problems with feedback. The counterargument (for the provision of feedback) will then be dealt with.

2.2.1 Three reasons for the identification of errors

Later on, there are sections on why and how feedback is provided and what could be done with feedback. For the purposes of this study, these are the main reasons to describe errors, but one could also identify three other reasons that are similar to reasons for classifying errors, but not so closely related to the classroom:

1. To improve the language
2. To describe learning
3. To annotate a corpus

One can argue that these categories overlap and they do: An annotated corpus of learner errors is useful in describing learning. A description of errors is useful in annotating a corpus or describing learning. The aim of this dissertation is to create a classification of errors that would combine the three – a list of errors with computer markup that could be used to establish the current status of the learners' knowledge.

2.2.1.1 To describe errors and improve language

James (1997:25-28) indicates that society is worried about errors and wish errors to be eradicated from the language. This is known as the "complaints tradition[1]".

The complaints tradition comes in two forms:

1. Type 1 complaints are concerned with upholding standards of correctness and discouraging misuse of specific systems. "Their aim is to root out linguistic usage that stigmatizes those who use it, like split infinitives, sentence-final preposition, double negatives, dangling participles, stranded prepositions and two score of the same order" (James, 1997: 29).
2. Type 2 complaints have to do with more "moralistic" complaints. These are seen as positive as they recommend clarity in writing and condone use of language that may confuse and mislead the public. The term "verbal hygiene" is also used to describe this (James, 1997:29-31).

[1] The complaints tradition (though summed up effectively by James, 1997) is discussed in much more detail by Milroy and Milroy (1991). For the purposes of this study, a detailed discussion is not warranted.

James (1997:31) indicates that although the above two complaints focus on different types of error, they have a lot in common as sometimes the two types overlap:

- Both are committed to protecting the public.
- Both employ the same type of logic.

One type of classification for this view of errors, comes from Moletsane (2002) who identified the following types of errors:

- Gross errors: affect the meaning of the entire sentence.
- Local errors: affect the interpretation or meaning of part of a sentence.
- Delicate errors: break an individual rule and do not lead to the generation of other errors.
- Breakdown errors: are likely to cause a breakdown in communication.
- None-breakdown errors: break the rule but do not interfere with communication.
- Critical errors: cause another error to be made in the same context.
- Persistent errors: errors which are continued to be made, despite remedial attention.
- Transient errors: do not persist for a long time. Similar to errors made by native speakers.

As in many fields of language, these types of errors may overlap. A local error and a breakdown error, for example, may seem very similar and for these categories to be useful, a detailed description would be necessary. These error types could be handy when doing error analysis with the purpose of remediation in mind, especially if effective communication is part of the goal of teaching.

2.2.1.2 To describe learning

Errors indicated in learner writing are aimed at indicating to the students where they are not yet fully capable of using the language correctly. In a sense then, this type of indication would describe how a learner "learns". There are numerous attempts at this

type of classification and a complete discussion will not be necessary for the purposes of this dissertation, since the description of learning is a major function of SLA.

What is important to note is that errors can be used to describe a learner's interlanguage, notify the teacher what needs to be taught and allow the learner to test hypotheses about the language (see James, 1997:2-12; Spencer, 1998:75; Ellis, 1996:50-57). Unfortunately, it is sometimes difficult to rewrite an error so that it makes sense – i.e. it is difficult to establish how proficient a learner is, based simply on his/her errors. Although errors may be useful, one should therefore be careful to read too much in them.

From surveying the literature, it is obvious that researchers differ on their views of how descriptions of errors would indicate learning in students. Errors are attributed to the learners themselves (poor attention), to the teachers and to common learning sequences. The end effect is that all agree that learner errors give clues as to how learning takes place, but there is little agreement as to what exactly these clues point to, or how to interpret them. It does not fall within the scope of this paper to attend to that problem, but it could make for interesting further research.

Errors are used not only by teachers and researchers to describe learning, but also to facilitate learning. Moletsane (2002:13) explains that in the 1950's and 1960's, errors were deemed to discredit teachers and learners, and learners were reprimanded for making too many errors. This attitude has changed with the realization that errors are inevitable and an essential part of the learning process in that they help the learner by providing him with "feedback on the process of concept formation in the target language. Errors are an indication that learning is taking place and they indicate the learner's proficiency in the target language."

Despite the caution not to read too much in errors, they do have a role to play in describing learning, although researchers differ to what extent. One method of describing this learning is by using an error annotated corpus.

2.2.1.3 To annotate a corpus

One cannot even start to address the question of error interpretation, unless one has a database of errors to work on. In this case, error classification is very useful. In order to use the vast amount of information inherent in errors, computerised corpus linguistics uses very specific classification systems. Cagneaux, Dennis and Granger (1998) describe how such a system would function: "The ideal error analysis system should enable researchers working independently on a range of language varieties to produce fully comparable analyses." A purely descriptive system would be the most effective. Errors are described in terms of linguistic categories. If errors were to be classified in terms of the source, the system would be too subjective (Cagneaux, Dennis and Granger, 1998:166).

The error tagging system used at Louvain for the ICLE-project is hierarchical: "...error tags consist of one major category code and a series of subcodes. There are seven major category codes: **F**ormal, **G**rammatical, Le**X**ico-grammatical, **L**exical, **R**egister, **W**ord redundant/word missing/word order and **S**tyle. These codes are followed by one or more subcodes, which provide further information on the type of the error" (Cagneaux, Dennis and Granger, 1998:166). According to the researchers, the system was found to be very effective and flexible, provided that the users knew exactly what each category referred to. Users also need to know exactly what to do in case an error could fit into more than one category. This would make it a very difficult classification to use in normal everyday marking. (This is discussed further in chapter five.)

As indicated above, there are numerous different ways of looking at errors and classifying them. The important thing is to clearly define what needs to be done with the specific classification system. The context in which the classification will be used is therefore the crucial variable. The present study attempts to create a dual-purpose classification system, answering the needs of standardised feedback as well as corpus annotation in order to enable error analysis. However, as mentioned, such a thorough breakdown of the error categories mentioned above would not be practical for everyday marking. Therefore it important to try and establish a tag set that covers a

wide range of these functions, without rendering it too big to be used for everyday marking.

Having looked at the purposes of identifying errors, it is time to distinguish between the identification of errors as done by a researcher and the type of identification as done by a teacher or lecturer when marking a student's text, as well as the identification of errors as done by the learners themselves. It is obvious that the "error identification" as done by a lecturer or marker and students, is not the same as error identification as done by a researcher. The researchers are (hopefully) much more detailed, accurate and have more time at their disposal. The researcher may or may not have more advanced electrical apparatus and software at their disposal and **have a different purpose with the identification of the errors.** A teacher identifies errors in the learners' texts in order to evaluate, teach and test, whereas a learner tries to identify errors to learn, so as not to lose marks for making avoidable errors. The researcher uses error analysis – the student and the lecturer use feedback (which may include a bit of elementary error analysis). Although the "why" of feedback and error analysis may be similar in some cases, the "how" differs more. The following pages will therefore first be devoted to the "why" of feedback. The discussion will not try to draw comparisons between error analysis and feedback, but will provide more information on why feedback should be provided. Thereafter the "how" of feedback will be described in order to identify successful techniques as practised at present.

2.2.2 Why provide feedback?

There are arguments for and against the provision of feedback. The arguments can be likened to the discussion of whether explicit teaching is effective or ineffective. Without being lured into that debate (since feedback stands astride that argument – see earlier discussion) this section will discuss problems affecting the effectiveness of feedback, thereafter contrasting it with reasons for the provision of feedback.

2.2.2.1 Reasons why feedback is argued to be ineffective

It has already been mentioned that feedback is expected by students and society alike, despite the problems with it – and there are problems. Spencer (1998:69-75) sums up research findings by other researchers on the issue of marking strategies. It seems that with regard to marking strategies in L2 context, the situation looks bleak:

- Marking strategies in L2 context have not yet been researched in great depth and no conclusive results have been reached.
- There is very little empirical evidence to suggest that students comprehend feedback or use it purposefully to modify their next efforts.
- There is little evidence to suggest that careful annotation of papers can help student writers improve.
- There are even doubts that the various types of correction techniques make any difference and that the use of any type at all may not be worth the effort.

Although not enough research has been done on the questions of feedback, the literature on the issue does indicate the following main problem areas:

1. A focus on form is ineffective
2. Feedback is often not individualized
3. Pattern recognition is ineffective
4. Importance of errors
5. Unclear purpose of feedback and teacher expectations
6. Confusion on the part of learners
7. Students cannot use feedback
8. Too little too late
9. Students are unwilling to use feedback
10. Feedback may lead to avoidance
11. Feedback does not lead to independent learners
12. Lack of revision
13. Time-consuming and not enjoyable for the teacher
14. Feedback may be insensitive and create negative competition
15. Grading is ineffective

These problems will now be explained briefly.

2.2.2.1.1 A focus on form is ineffective

Formal instruction does not lead to the kind of knowledge used to communicate. When students have serious problems to communicate, the programs used by teachers would primarily focus on form and correctness, which may lead to students being unable to experience the process of discovery and thorough revision (see Spencer, 1998:62;76 and Ellis, 1996:653 for a more complete discussion).

2.2.2.1.2 Feedback is often not individualized

James (1997:257) writes that when learners are expected to correct errors in exercise passages, the obvious problem is that "not every single learner in a class or group has committed all of the errors that are exposed. Those who have not made any or many of the errors might object to or be bored by the exercise". It is therefore important to know exactly which learners have trouble with which features of the language so they can be given personalised exercises or exposure. Before the age of computers, this would have been impossible – now it is merely difficult.

2.2.2.1.3 Pattern recognition is ineffective

When a teacher repeatedly marks a specific error in an individual learner's texts over the course of a year or more, the error may still not strike the learner as salient. The error still appears within a "forest of other marking" and this may prevent the learner from noticing a pattern of errors in an essay or in a portfolio of essays (Wible et al., 2001:308-310).

2.2.2.1.4 Importance of errors

Often in marking, no indication is given of the level of importance (status) of the error (Spencer, 1998). A student would have to guess how severe his/her problem is. A marker may also be fooled by plenty of typing errors, in an otherwise good text. Typing errors, for example, are not a high level error since they are more of a nuisance. Errors of syntax or word choice can however create much bigger problems to do with comprehension of the text.

2.2.2.1.5 Unclear purpose of feedback and teacher expectations

The impact of grading and feedback is short-lived. Feedback should be an intervention strategy and not simply aimed at the here and now (Monyaki, 2001:14-16). Hyland (1990:279) even found that feedback itself, rather than the learners acting on feedback, becomes the focus of the action. Moletsane (2002:27) also found that teachers tend to lose sight of the purpose of feedback, and teachers seem to see their role as to simply identify errors. This is clearly losing sight of the purpose of feedback.

A problem with lecturer comments is that the lecturer reads the text expecting something specific – it could be that he/she understood the assignment much differently from the student. Students also resent comments on the content of the paper, especially if the content could be considered "personal opinion" (Spencer, 1998:71-72)[2]. Spencer (1998:55) indicates that the way lecturers read student writing, is the wrong way round. Normally a reader reads a text, assuming that the text has coherence and is structured in such a way as to convey the intended meaning effectively. A lecturer, on the other hand, approaches student writing with a scepticism – going against the grain while reading.

[2] From conversations with my students I have heard that they dread "give your own opinion" questions. If the opinion of the student differs from that of the lecturer, they often get bad marks. Clearly the lecturer then did not want the students' opinions, but simply a rewrite of his/her own opinion.

2.2.2.1.6 Confusion on the part of learners

Feedback can be confusing. Learners get dejected or confused by a text which is full of marks – even if some of those feedback comments are positive, as there is simply too much to take in. Furthermore, many language aspects are marked in the same manner. This confuses learners, as even a trained linguist sometimes does not know what the feedback refers to. Students do not seem to know what the feedback marks mean either (Monyaki, 2001:66,74), even though they may sometimes be able to correct an indicated error (Moletsane, 2002:31 and Hyland, 1998). Feedback has also been found to be inconsistent and vague (Hyland, 1998:225 and Spencer, 1998:107).

2.2.2.1.7 Students cannot use feedback

Due to the confusion mentioned above, learners do not know how to react to feedback. Learners see an instruction to rewrite something as a punishment and not as an opportunity to improve their work (Monyaki, 2001:75 and Moletsane, 2002:30). The problem in understanding feedback may lie in students' inability to read (see Spencer, 1998:52 for a more detailed discussion), or they do not have the skills to understand and use the feedback (Hyland, 2003:218).

2.2.2.1.8 Too little too late

In the study done by Monyaki (2001:63;65) he reports that the majority of the teachers in his test group agree that feedback is more effective before the final product is delivered. This is rather obvious, but in most situations, feedback can only be given after the final product is delivered. In most cases, this feedback is then ignored as it would not change the grade obtained for the paper (also see Munchie, 2000:50-51).

2.2.2.1.9 Students are unwilling to use feedback

Students experience a vague "rubber-stamp" type of feedback (or error correction as only feedback) very negatively. This creates an inattention to feedback and may be an additional reason why students fail to see patterns of errors in their writing. (See Hyland, 1990:282.)

2.2.2.1.10 Feedback may lead to avoidance

Moletsane (2002:30) and Munchie (2000:49) speculate that students may be afraid to take a risk, due to previous feedback. Spencer (1998:56;62;109) also found that students are likely to omit an idea or construction if they are unsure about the correct action on a specific comment, even though the comment may be easy to understand.

2.2.2.1.11 Feedback does not lead to independent learners

Both Munchie (2000:49) and Monyaki (2001:76) found that learners often indiscriminately use a teachers' comment, which implies a lack of critical processing and evaluation of feedback. Munchie speculates that this lack of evaluatory and decision-making reasoning may reduce the impact and long-term effectiveness of feedback and revision. Spencer (1998:73) also found an uncritical attitude in students' reactions to feedback and she also warns that it may be counterproductive.

2.2.2.1.12 Lack of revision

Paulus (1999:266) indicates that an incorrect focus in feedback may contribute to ESL students' lack of writing and revision strategies. A feedback comment which prompts an incorrect revision, is also not effective (James, 1997:236). Revision based on feedback therefore also needs feedback.

2.2.2.1.13 Time-consuming and not enjoyable for the teacher

To mark student essays is time-consuming and not enjoyable for the teacher. This attitude towards marking, is worsened if it does not seem to bring about any changes in the students' performance (see Moletsane, 2002:21).

2.2.2.1.14 Feedback may be insensitive and create negative competition

Probably due to work overload and the unenjoyable side of marking, feedback is seldom "nice" to students. Comments are short, careless, exhausted or insensitive (Spencer, 1998). Askew and Lodge (2000:7) also indicate that feedback may encourage competition and comparison which could be negative as some learners would simply give up trying. But is this really the case with feedback as a whole or only with grading? This in an important distinction to make. Grading will therefore be discussed on its own later.

Feedback that is simply indicative of satisfaction with a learner's performance, may prove to be unhelpful if it is given in a general or indiscriminate way (Askew and Lodge, 2000:7). (This links with the "individualized feedback" mentioned above.) Askew and Lodge quote Brophy in saying that "Infrequent but contingent, specific, and credible praise seems more likely to be encouraging... than frequent, trivial or inappropriate praise."

Askew and Lodge (2000:7) have coined the phrase "killer feedback" for situations were the receptive-transmission form of feedback blocks learning. Such feedback discourages all further drafting, is too much and feels overpowering, does not give any help and encourages no dialogue between learner and lecturer.

Everything seems to come back to the issue of a relationship between the learner and the lecturer, but with big classes it is not always possible for the lecturer to know his students well enough to know who will be discouraged by a certain style of feedback.

2.2.2.1.15 Grading is ineffective

Grading is a natural part of the administrative process at universities and schools and the fact that it provides a value judgement about a text, qualifies it as feedback as well. Spencer (1998:101-103) found that grading is ineffective as feedback. There are several reasons for this: Grading is not an objective way of giving feedback because the grader is the product of a specific cultural value system. Furthermore, a grade overshadows everything else and creates false impressions. A high mark indicates a satisfied lecturer and a low mark a dissatisfied lecturer. Commentary on the other hand encourages rewriting. A lot of positive comments for the further development of a text, may be ignored if a learner is blinded by a lower than expected grade. Spencer also found that grading distorts power relations. Teacher response is often linked to the grade and is used as a means to justify the grade given. This is wrong. Feedback is supposed to facilitate learning.

Students interpret grades wrongly. Spencer (*ibid.*) writes that grades, instead of knowledge or learning, are seen as the mark of success. Students need to learn that a good grade is a by-product of the care and attention required by the writing process.

2.2.2.2 Why should feedback be provided?

There are a number of arguments for and against giving feedback on learner errors. This section will focus on reasons why to give feedback or correct errors. There are numerous reasons for feedback, in spite of the difficulties associated with it. Among these are the expectations from students and society, proof that it does help students, and numerous others. One of the main reasons is "Consciousness-raising" – making the learners aware that their English proficiency (especially academic English proficiency) is not what it should be (see chapter one). Since Consciousness-raising

(or CR for short), is so important, it will be discussed separately in section 2.2.2.2.7. First, the other arguments in favour of feedback will be discussed. These are:

- Society expects it
- Feedback enhances learning
- Feedback counters fossilization
- Learners expect it
- Feedback can improve writing
- Other reasons

2.2.2.2.1 Society expects it

Generations of language learners have attended school and university and received feedback on their language use. To suddenly do away with it completely, could seem to society as if lecturers and teachers were not doing their job. Teacher feedback is considered to be an integral part of the role of the writing teacher (see Storch and Tapper, 1997:245).

2.2.2.2.2 Feedback enhances learning

As early as 1972, Selinker and Lamendella claimed that extrinsic feedback may assist in language learning. This is however a controversial statement, as there are almost equal numbers of researchers who raise their voices for and against feedback. Lately it seems as if the tide is turning in favour of feedback. This may have to do with a change in feedback techniques based on the growing attention given to the subject since the 1980's. Techniques are improving and teachers are more aware of what NOT to do. The problem, it seems, had to do with connecting feedback and learning. Due to teacher awareness and techniques such as corpus analysis and error analysis, feedback can now enhance learning if used correctly.

Lately, more and more evidence points towards the effectiveness of feedback. Hyland (2003:219) mentions that "studies which measure student improvement longitudinally

after error correction in terms of accuracy ... suggest that students who receive error feedback over a period of time do improve their language accuracy". Askew and Lodge (2000:2) come to a similar conclusion when they report on seminars on effective learning. They found that feedback is important in supporting learning at individual, group and organisational levels. They observed that a focus on feedback was popular and that the notion of feedback seems unproblematic. Askew and Lodge (*ibid.*) also observed that people had different perceptions of feedback and its functions and processes based on their perceptions of learning. Feedback should be seen as a crucial feature of teaching and learning processes and should be considered an element in a repertoire of teaching techniques (Askew and Lodge, 2000:1) of connected strategies to support learning.

The way to support this learning is by simply being a guide to recognise what has been done wrong, in order to prevent new errors. However, feedback on its own is not sufficient. Askew and Lodge (2000:1) warns that "learning is supported by a whole range of processes, one of which is feedback".

2.2.2.2.3 Feedback counters fossilization

Feedback should be a guide to evading fossilization. If the target language is simply learnt through communication oriented instruction, communicative competence can be acquired before the grammatical structures of the language has been mastered. This increases the risk of fossilization (Moletsane, 2002:28). Feedback has to counter that risk. Krashen (1985:43-52) explains that fossilization may be the result of:

- Insufficient quantity of input
- Inappropriate quality of input
- The affective filter
- The output filter
- Acquisition of deviant forms: in order to ensure that deviant forms are not acquired, they should be eradicated.

Feedback is seen as a way of providing input. However, comprehensible input alone is not sufficient for successful L2 learning. Comprehensible output is also needed, and

that is mostly done through speaking or writing. If provided through speaking, the learner can negotiate meaning with the listener (which can be seen as a form of feedback). If done through writing, however, the learner has no way of judging comprehensibility, if not provided with feedback. Lyster and Ranta (1997:41) warns that subject-matter teaching does not provide adequate language teaching on its own, but that "language used to convey subject matter needs to be highlighted in ways that make certain features more salient for L2 learners".

2.2.2.2.4 Learners want to be corrected

Learners definitely want feedback and/or corrections on their work (Monyaki, 2001, Hyland, 1998:256, Hyland, 2003 and Spencer, 1998:88-90). One student in Hyland's (2003) study indicated that being corrected once or twice may not be enough. She would still forget occasionally, but "...the teacher should keep correcting me and some time I will remember". Spencer (1998:88-90) also found learners surprisingly eager to be corrected. She surveyed research on what learners expect from lecturers. Her findings were as follows:

- Learners want to be corrected more than lecturers believe is necessary.
- Learners feel it is necessary to correct learners' errors in English so that they can use the language more accurately and fluently. Learners are very negative towards an approach of "no comments at all".
- Teachers should not overlook errors and simply focus on communicative effectiveness of the text.
- Errors should not only be corrected when the majority of the students struggle with the grammatical feature.
- Learners do NOT feel that constant error correction can cause frustration and discourage the learner. (This contrasts sharply with Krashen's findings.)
- Learners do not want selective error correction.
- Learners found that the following areas requires the most attention:
    - Organisation
    - Grammar
    - Pronunciation
    - Vocabulary

- Spelling and punctuation were not seen as so important by the learners, as they felt they could correct those themselves.
- The most favoured form of correction was "error indicated and cue for self-correction" followed by "error and answer" and using errors as examples in the classroom.

Storch and Tapper (1997:259) found that second language writers often see their texts in terms of the number of errors they make. Whether this is a good thing or not is debatable, but it makes sense then to show learners where they do not meet their own standards for correctness. Luckily, students do not seem to want their work simply corrected. Spencer (1998:208) indicates that what students want from feedback or commentary on their writing, is the following:

- Input from their lecturers in the form of detailed commentary.
- A mark.
- Specific, encouraging, honest criticism.
- To take responsibility for their own writing. They are open to the concepts of computer-assisted instruction and peer review.

Overall, there is enough evidence to establish that learners themselves feel they will not be discouraged by feedback. Learners definitely want feedback, but then it should be honest, thorough criticism that can enable them to take responsibility for their work.

2.2.2.2.5 Feedback to improve writing

If learners' errors are not corrected, learners may assume that their writing has communicated their meaning in an acceptable way and they may see no need to revise their work. If errors are left uncorrected, learners may become unable to distinguish between erroneous forms and correct ones (Moletsane, 2002:22-27). Hyland quotes research by Ferris and Roberts who found that students whose errors had been marked with a correction code or where errors were simply indicated, were able to correct 60-64% of their errors, but students who were left on their own, only managed 18%.

Hyland (2003:218) states straightforwardly that "It has been found that feedback on error can improve students' writing in the short-term". Askew and Lodge (2000) have also found that writing can facilitate long-term improvements (see section 2.2.2.2.2). Students usually see teachers' corrections as being helpful if they *are able* to use them to revise and correct their errors. The students must be able to understand the corrections (Moletsane, 2002:30).

2.2.2.2.6 Other reasons for feedback

Feedback is useful to draw attention to errors. This will be discussed under the issue of Consciousness Raising (section 2.2.2.2.7). If an error is left uncorrected, learners may interpret it as correct (see Moletsane, 2002).

Feedback can help to correct errors. Learners who find feedback of lesser importance, are often weaker students (see Spencer, 1998:184). In addition, students sometimes already know the correct form to use, but they need to be pushed to use it (Lyster and Ranta, 1997:41).

Feedback can facilitate communication: In some cases, the negotiation of meaning is not the main reason for correcting a structure produced by a learner (Lyster and Ranta, 1997:42). A poorly structured essay may still provide the reader with all the relevant information, only with more effort. Lyster and Ranta (1997:41) see the importance of feedback in this process, as very important, saying that "... producing comprehensible output entails the provision of useful and consistent feedback from teachers and peers...".

Hyland (1998:267) found that in some cases, evaluation (giving a mark) can be a motivating factor for students and should not be regarded as having only negative effects.

The question of whether all errors should be corrected, should also be mentioned here. The problem is that all errors are not of the same seriousness. Some errors may create much bigger problems for the learner than others. This is referred to as "error

gravity". Error gravity refers to the error's intelligibility, acceptability and irritation. The urgency of the treatment of an error depends on its gravity and seriousness (Ellis, 1994:701). However, in a learning situation, the gravity of an error is not the issue at hand. The aim is to indicate *all* errors since a lot of "small" errors may in the end create just as much problems as one "big" error in that clarity and comprehension is influenced.

2.2.2.2.7 Consciousness Raising

Second language acquisition theories state that learners use a variety of processes to learn a L2. The theories to explain this are cognitive in nature explaining the mental processes that enable learner to work on input and the knowledge systems which they construct and manifest in output (Ellis, 1996:30).

The best known theory is the interlanguage theory. The term refers to the interim grammars which learners build on their way to full target language competence. The theory is constantly changing, but a standing theme is that of hypothesis testing: learners "form hypotheses about what the rules of the target language are and then set about testing them, confirming them if they find supportive evidence in the input and rejecting them if they receive negative evidence" (Ellis, 1996:30). This process is mostly subconscious.

Ellis (1996:30) explains that during hypothesis testing, learners have certain mechanisms they use to

- notice features in the input,
- compare these features with those that are currently part of their mental grammars or interlanguages and
- integrate the new features into their interlanguages.

In order to enhance this process, teachers try to raise the awareness of the learners with regards to the features they wish the learners to acquire. This is known as consciousness-raising.

Consciousness-raising (or CR for short) is a sensitive issue, since it touches on the question of whether or not formal instruction is useful. The term "formal instruction" refers to grammar teaching. It shows the importance and centrality attached to grammar in SLA (Ellis, 1996:611). Although there are researchers who claim that formal instruction is not useful, research findings prove otherwise (see Ellis, 1996:611, Ellis, 1992:232-241, and James, 1997:244). Grammar has a practical and a theoretical importance. According to Ellis (1996:611) it has helped teachers understand the factors that determine whether instruction is successful. It has also helped researchers explore issues of importance for theory building. It has been especially useful to investigate the relationship between the linguistic environment and the learner's internal processing mechanisms.

Formal instruction has been found to be advantageous for children and adults alike, and for intermediate and advanced students. Formal instruction has helped in acquisition-rich and acquisition-poor environments. This has been proved by evaluating formal instruction by means of different tests (see Ellis (1996:613-614).

Despite numerous arguments in favour of formal instruction (see Ellis, 1996 and James, 1997:246 and further for a detailed discussion) there is one limitation: Formal instruction cannot alter the ROUTE of acquisition, but only the rate (Ellis, 1996:631). In the end, learners who have had instruction, demonstrate higher ultimate achievement (James, 1997:244).

Some researchers claim that CR is not useful as it creates implicit knowledge, while learners use their explicit knowledge more. CR does have a distinct advantage, as it has been found that once learners' consciousness of a particular feature has been raised by formal instruction, they continue to be aware of that feature in subsequent communicative input (see James, 1997:257 and further for a discussion). Learning is also not developmentally constrained, "which means that there is no fixed order in which it (language) must be learnt: it can be learnt and taught in any convenient order," (James, 1997:257).

Another advantage of CR, is that it can assist both students and teachers. Moletsane (2002:21-22) writes that if teachers are aware of the aims of marking and are

conscientious in the application thereof, marking can be of invaluable assistance to them. This means that by indicating errors in students' work, the teachers themselves become aware of the areas their learners have trouble with. If they are aware of the weaknesses of their students, they could adapt the input they produce for the learners. One of the aims of the current project is to utilize the amazing record-keeping abilities of the computer for exactly this purpose, in order to create individualized exercises or input based on the specific needs of individual learners.

Consciousness-raising is clearly linked to effective feedback. The trick is in providing feedback in such a way that it creates the opportunity for CR to occur.

Now that the reasons for the provision of feedback have been established, it is time to discuss the status quo of feedback practice. How is feedback provided?

2.3 How feedback is provided

In this section the "how" of feedback will be discussed, looking specifically at techniques currently in use and what experienced teachers do. Since I argue that most of the arguments against feedback can be traced to ineffective feedback practice, this section will ultimately lead to a checklist of ideal feedback. I believe that if the guidelines of this checklist are adhered to, feedback will be much more effective. In order to get to that point, an investigation of current practice is needed.

2.3.1 Varieties of feedback

Feedback on written work is given in a variety of manners and can come from a variety of sources. These methods can be broadly categorized based on the medium it is presented in (spoken or written), the source of the feedback (teacher, group, peer or computer), the level of interactivity between the learner and the one providing the feedback, the scope of the feedback and the level of standardization of the feedback.

As yet, there are no hard and fast rules what "good" or "bad" feedback practice is, as it is situation dependent. Using standardized feedback comments may work in one situation, but not at all in another. (This will be discussed shortly.)

From the literature, it seems there are at present roughly 14 techniques for providing feedback. (These will be explained shortly.) Note that even within a set technique, the feedback technique may have various different versions. (This is illustrated in table 2.1.) For example: a class discussion can occur in spoken form or written form, or even in written form on a local area computer network. Every student may get one chance to make a statement, or it may result in a lively discussion. Another example is self-assessment: a student may do self-assessment during a class discussion on insight gained during the discussion, and the whole nature of a class discussion is a group activity. In that sense, a class discussion can have more than one option within all the above categories. The fourteen feedback techniques are listed below.

1. Comprehensive marking
2. Selective marking
3. Correction code
4. Minimal marking
5. Grading only
6. Computerised marking
7. Class discussion (workshop type)
8. Conferencing
9. Taped feedback
10. Peer review
    a. Oral
    b. Written
    c. Networked computers (computerised workshop on a Local Area Network)
11. Self-assessment
12. Self-correction
13. Student annotation
14. Computerised feedback
    a. Grammar checkers and Spell checkers
    b. Nihongo-CALI (Mentioned here as an example)
    c. IWill (Mentioned here as an example)
    d. Networked workshop (See 10c above)

Table 2.1 illustrates how all these techniques fit under the categories of Medium, Source, Interactivity, Standardization and Scope.

Technique	*Medium*	*Source*	*Interactivity*	*Standardized*	*Scope*
Comprehensive marking	*Written*	*Teacher*	*Situation dependent*	*No*	*Comprehensive*
Selective marking	*Written*	*Teacher/ peers*	*Situation dependent*	*No*	*Selective*
Correction code	*Written*	*Teacher/ peers*	*Situation dependent*	*Yes*	*Selective/ comprehensive*
Minimal marking	*Written*	*Teacher*	*Situation dependent*	*Situation dependent*	*Selective*
Grading only	*Written*	*Teacher*	*Situation dependent – probably minimal*	*Situation dependent*	*Minimal*
Computerised marking	*Written*	*Teacher/ peers/ computer*	*Minimal to comprehensive*	*System dependent*	*Minimal/ comprehensive*
Class discussion (workshop)	*Oral/ written (computer)*	*Peers and teacher*	*Comprehensive*	*No*	*Minimal/ comprehensive*
Conferencing	*Oral/ written*	*Teacher*	*Comprehensive*	*No*	*Selective/ Comprehensive*
Taped feedback	*Oral (and) written*	*Teacher*	*Minimal*	*No*	*Selective/ comprehensive*
Peer review	*Oral/ written/ computer*	*Peers*	*Comprehensive*	*No*	*Selective/ comprehensive*
Self-assessment	*Written*	*Student*	*Comprehensive*	*No*	*Selective/ comprehensive*
Self-correction	*Written*	*Student*	*Selective*	*No*	*Selective*
Student annotation	*Written*	*Student*	*Selective/ comprehensive*	*Situation dependent*	*Selective/ comprehensive*
Grammar Checkers/Spell Checkers	*Written*	*Computer*	*minimal*	*Yes*	*Selective*
Nihongo-CALI	*Written*	*Computer*	*Intermediate*	*Yes*	*Selective*
IWill	*Written*	*Teacher*	*Selective/ comprehensive*	*Limited*	*Selective/ comprehensive*

Table 2.1: Classification of different feedback techniques

These techniques can all be discussed at length, but that is not the purpose of this section. The purpose is to find what works from these techniques. The different techniques will therefore now be described briefly in order to identify their strengths and weaknesses. This will lead to the checklist of effective feedback.

2.3.2 Written Feedback

Written feedback is by far the most widely-used medium in the educational context in that it is used (sometimes additionally) in all fourteen techniques. Written feedback is used mostly in comprehensive, minimal and selective marking, as well as with a correction code and in computerised marking. If only a grade is awarded (grading only), it can also be considered written feedback.

Comprehensive marking is simple. Every single error is indicated, based on the assumption that if an error is not indicated, learners could be under the impression that they did not make any errors. Learners seldom look at their writing and they would much less look at it if they thought that it was error-free.

Comprehensive marking has the advantages that it gives learners an indication of the number of errors in their work and that it can motivate learners to reduce errors. It may also avoid confusion in that something is not marked an error in one place and ignored in another. It also counters fossilization in that learners are not left with the impression that incorrect language use is in fact correct (Moletsane, 2002:32-33).

On the other hand, comprehensive marking faces the problems of a large variety of errors that has to be marked and it can happen that a text is so full of feedback marks that it is difficult to know what to correct. A text so covered in feedback may also discourage learners. Due to the large amount of feedback (and the excessive time spent on marking) the feedback itself may not provide sufficient information and learners may not understand all the corrections. It may also happen that the feedback is an all-out correction, which discourages independent thought and action from the student. Learners may also be afraid to experiment with the language.

Teachers find comprehensive marking tiring and complain that it takes too long to return the work to the students. This results in students getting less work as teachers are still busy marking the previous. Due to the variety of errors, teachers may be indecisive on what exactly they want to achieve with their feedback and learners may not analyse the errors they actually look at. Lastly, comprehensive marking can result in writing being viewed as a product and not a process, resulting in lecturers focusing on what is bad on not on the positive aspects of a text.

Selective marking on the other hand has been found to be more effective than comprehensive marking (Moletsane, 2001:1). Selective marking concentrates on specific structures, marking only errors of a certain kind selected beforehand. Only the important errors are marked. This allows the marker and student to approach errors systematically, instead of having to face all the errors at once. Moletsane (2002:1-2) found that selective marking has the advantage of allowing the learner to build up confidence. Moletsane also claims that it eliminates corrections which might not be understood by the learner, though this is probably more due to a degree of standardization. (Due to the fact that only a few error types are focussed on, the learners will not be so confused with feedback marks. In other situations, the marks and words used to give feedback may refer to more than one error type. See chapter 4.)

Selective marking enables learners to aim at a specific goal and allows teachers to emphasize a specific aspect of language, concentrating learners' attention to constructions and rules the learners do not yet understand. It avoids a disheartening amount of red ink, has fewer distractions and reduces the marking time. Moletsane (2001:1) also claims that it can boost morale.

On the downside, it is difficult to categorize all errors while marking, and even more difficult to decide which errors are important enough to mark. Selective marking still provides little information on the errors and may create the impression that the text has few errors. Some learners may think that the teachers are not doing their work and parents may even think that teachers are incompetent. Ignoring some errors may induce fossilization.

Minimal marking is another version of selective marking. In this technique, an error is simply indicated by means of a (it seems) non-specific symbol – a line under the error, a cross in the margin, circling the word or phrase, a question mark or an "insert" sign. The idea is that this method encourages active correction since the students have to identify and correct the errors themselves. It prevents excessive correction (keeping the essay clean) and shifts the responsibility to the learner. Using minimal marking is quicker, so it can present the lecturer with more time to spend on substantial problems. The implication is that errors were caused by oversight and that students are capable of self-correction (Spencer, 1998:119).

Problems with minimal marking is that there is no research that demonstrates convincingly that the formalistic minimal marking approach results in improved proficiency over a period of time (Spencer, 1998:120;163). The possibility also exists that students will not be able to identify the errors themselves. (In chapters 3, 4 and 5 of the present study, I came face-to-face with minimal marking and in some cases I found it impossible to deduce what was in fact the error indicated. This is a real problem.)

Selective marking, comprehensive marking and minimal marking may be done by using a correction code. A correction code is a marking system in which errors are indicated using a code in the margin. For example, a spelling error may be underlined and provided with the code, "SP" in the margin. This is a type of "response shorthand" and is good practice in the sense that it shows exactly where the error is and gives a hint how to correct it. It is the marking technique used by more than 80% of lecturers (Spencer, 1998:114).

Spencer (*ibid*) indicates that there are a lot of negative feelings in the academic society regarding a correction code method. It has been labelled as laborious (Spingies, 1990:26), of dubious value (Spingies, 1990:26), mechanised and formalistic. It is error-oriented (Spingies, 1990:26) in that lecturers can believe that once they have picked and indicated all the errors and given a mark, they have finished their job. This does not mean that they have actually **read** the essay (Spencer, 1998:114-115).

In spite of the criticism by Spingies, both Spencer (1998:145) and Hyland (1990:280) found a correction code to be effective, helping to improve content as well as grammar. The criticism cited against the correction code (see Spencer, 1998:116-118 and Spingies, 1990) is common criticism levelled against feedback practice as a whole.

An advantage of the correction code is that it can briefly and efficiently indicate the correct route – like a road sign. It is feedback with minimal human intervention that informs student when **and how** they are diverting from the path. There is of course the limitation that students have to comprehend the grammatical issue identified. The errors should also only be identified and not corrected in full (Spencer, 1998:115-116) and the code should be kept simple and easy to understand (Spingies, 1990).

An even more minimalist approach to feedback than minimal marking, is grading only (see Monyaki, 2001:46). In this case, students receive no feedback other than a mark, though some overall comments may also be included. This presents students with an indication of how assessors value the work they produce. This can be seen as a crude substitute for many of the mechanisms that provide feedback in naturally occurring language learning situations. A problem here is that learners often do not know why they get a good or a bad mark, errors are still ignored and students may write for a mark and not to communicate.

The overall problems with written feedback techniques are therefore:

- Too much feedback discourages and confuses students.
- Too little feedback may induce fossilization and create the false impression that a text is flawless.
- Written feedback overall may not provide enough information for students.
- Written feedback may take up a lot time.
- To strike a balance between understandable "corrective hints" and all-out corrections, is difficult.
- Error gravity remains a problem.
- Marking may degenerate to "looking for errors".

- There is no guarantee that students will respond to it.
- There is not a guarantee that it will result in communication between the learner and the lecturer.

2.3.3 Discussion techniques

During class discussions (see Monyaki, 2001 and Spencer, 1998), a teacher or lecturer simply discusses certain problems or maybe a specific essay in class. This method encourages the feeling of audience and group discussion, but has the disadvantage that some students are shy and dislike being put on the spot. It is also very time-consuming, with the result that some students do not get feedback at all. Students also do not have a written commentary to go back to for revision, unless they make their own notes.

The same problems apply to conferencing (a one-on-one interview between the student and the teacher) but it has been found to be very effective (Monyaki, 2001:53) because a large number of problems and strengths can be discussed. Individual attention can also be provided.

Another type of discussion technique, is Audio-taped feedback (ATF) (see Spencer, 1998:122 and Monyaki, 2001:53-55) in which a lecturer reads an essay and then tapes his/her comments. The learner then listens to the tape. ATF is quick and can give detailed commentary. It encourages reader-based responses (rather than editorial or proof-reader stance). Learners find it interesting, comprehensible and emphatic and lecturers claim their ability to respond to complex texts improved.

Audiotapes are well-known to most people and are therefore not threatening. They are relatively cheap, are reusable and are an easy way to create the "personal interview feeling". Lecturers found it to be a fuller form of feedback, but very time-consuming and an additional administrative burden. Students have a record to go back to later on, but they find it annoying not being able to interrupt the lecturer and indicate that accents may be a problem (see Spencer, 1998 and Monyaki, 2001.)

The discussion techniques therefore have the distinct advantage that they can present a proportionally larger amount of individualized feedback than written techniques, but according to Monyaki (2001), it is too time-consuming to be practiced continually and learners may be uncomfortable with all the personal attention.

2.3.4 Self-feedback

Self-feedback refers to any type of feedback that learners do independently. I also classify peer review (peer assessment) under self-feedback. The level of interactivity between learner, lecturer and peers may vary, as well as the medium used to convey the feedback.

Self-assessment and self-correction are very closely linked. In self-assessment, the learners have to determine the relevance of the concepts discussed in class and apply them to their own writing. It is an evaluation from the learners' perspective, an internal activity. Self-correction, on the other hand, is more a type of proof-reading exercise in which learners are actively looking for their own errors. The distinction seems a bit superficial, since the main idea remains the same: to promote learner autonomy. In both situations, the concept of the learner as active participant in the learning process and the shared responsibility of assessment and quality control, are brought to the fore. It is possible to boost students' motivation by making them more responsible, raising their awareness, clarifying appraisal criteria and raising metacognition (Spencer, 1998:126-127, Monyaki, 2001:43-45 and Hyland, 1990).

The disadvantages of self-correction and self-assessment are that lecturers are afraid that it would seem as if they are not doing their work. It also creates a lot of extra work for students and administering control over the process can be difficult. Students will also not be able to identify all their errors, necessitating the need for the lecturer to mark it again in any case. Some training will also be needed to make sure that learners understand what they are doing, and this will take time and could be difficult, creating student apprehension (see Spencer, 1998 and Monyaki, 2001).

Peer review has the same kind of shortcomings and advantages, in that it is still a process executed by students. The difference is that they comment on each others' work, thereby obtaining experience in identifying problems in writing. They help each other plan and draft and create the impression of an authentic audience. Students could be honest and less evaluative than a teacher (Monyaki, 2001:45-46). The process is very versatile and can be done at home, in class, on networked computers, over the internet or verbally. It has been found to be effective in a process approach to writing, but is time-consuming and students still expect teachers to look at their work. The technique is not effective enough on its own and should be seen as an additional help and not an independent technique.

The advantages linked to peer feedback (on top of the advantages mentioned for other self-feedback techniques) are:

- Social support from peers;
- It is informative because peers are at the same level;
- Students learn by helping each other;
- It creates the opportunity for independent prioritization;
- Peers provide a much broader audience;
- Interaction can build a broader knowledge-base;
- Students find the generation of ideas and the spotting of trouble easier;
- Students can learn from the mistakes of others and understand feedback better;
- Feedback is less threatening, can be more encouraging and is simply quantifiably more;
- Peer feedback may result in more changes in content, organization and vocabulary than teacher feedback (Paulus, 1999:268).

The disadvantages are that students need explicit instructions, otherwise they could simply be wasting time (Jacobs, Curtis, Braine and Huang, 1998:310). Learners may also have different expectations from the feedback, their peers may not seem qualified enough and students prefer the "expert" feedback provided by the teacher (Paulus, 1999:269).

The final technique in self-feedback is student annotations. In this method, students annotate their own writing, indicating where they had difficulties. According to Storch and Tapper (1997:260), this technique has the advantages of individualized response and the promotion of a dialogue between the learner and the lecturer. It may also reflect areas that did or did not receive enough attention in class. Since students are the initiators of the feedback, it is hoped they will pay more attention to it, and it is illuminating for students to isolate their own strengths and weaknesses (Storch and Tapper, 1997:260).

The downside is that teachers could be lured into only focusing on the areas highlighted by the learners. Learners may also not always know what their own weaknesses are, or not know how to put them in words. Students might also not be willing to do the "additional work".

The advantages and disadvantages of self-feedback highlight the following qualities of good feedback: Learner autonomy and student motivation should be promoted. Learners should be trained to be able to identify and evaluate errors in their own work. A sense for the reader should be promoted, and a dialogue between the learner and the lecturer (or the learner and a reader) should be established. Unfortunately, self-feedback still has the disadvantage of the time it consumes and the problem that students still do not necessarily use the feedback.

2.3.5 Computer-based feedback

Computer-based feedback falls under the fields of Computer-assisted language learning (CALL), Computer-assisted language instruction (CALI), Intelligent Computer-assisted language learning (ICALL) and a few other computer-related fields. Most computerized systems are not yet advanced enough to function effectively independently. Computer-assisted exercises have the shortcomings that they cannot give hints (only all-out corrections), they do not provide for alternative correct answers, and they are difficult to program, since error anticipation is difficult if not impossible. The solution to this lies in natural language processing (a field developing quickly), but it has not yet developed sufficiently to enable independent

computer marking. (See Nagata, 1993:330 for a description of some of these techniques.)

Above, the two programs Nihongo-CALL and MINIPROF (see Nagata, 1993) were mentioned as examples of computer programs that can assist in marking. These programs can only be used on small amounts of text. These systems are effective in targeting specific deficiencies in students' performance, but only if those deficiencies are pointed out by a human. After the 1993 article on Nihongo-CALL and MINIPROF, I found no further information in the literature, so further comments would be useless, taking into account the immense speed at which technology advances.

Grammar and spell checkers are common tools integrated into most modern word-processing software. Although these are helpful in finding little errors, they are of little use for foreign language learners. The reason is that foreign language learners' errors differ widely from native speaker errors (Cagneaux, Dennis and Granger, 1998:165). These checkers are also relatively ineffective in identifying errors of syntax and cannot be trusted to identify organisational problems.

The final computer-based feedback technique mentioned here as an example, is the IWill system developed by Wible et al. (2002) and used in Taiwan. The system (Intelligent Web-based interactive Language Learning) is a computer interface and database in which students' essays are submitted, marked and archived.

Wible et al. (2002:299) designed their system based on certain assumptions:

- The system should be interactive.
- The system should exploit the computer's ability to record and track previous interactions between student and lecturer.
- Pinpoint searches of these interactions should be possible.
- The system should be seen as a component of a larger language learning program and should not function independently.

Wible et al. (2002:300) state that the advantages of their system are that it can archive:

- Learner writing.
- Feedback provided by lecturers on the learner writing.
- Corrections made by the learners, based on the feedback given by the lecturers.

The last-mentioned advantage is interesting as it can be used as a direct trace of the effectiveness of feedback and the way that learners interpret and use it. The project reported on in the present study, will also enable this type of monitoring so that the effectiveness of feedback can be traced for purposes of refinement.

Wible et al. (2001:303-304) describe the marking environment on the computer. They mention that teachers have the opportunity to insert their own comment each time they want to make a correction or give feedback, or the teachers can simply select the relevant comment from a "comment bank". In this comment bank, some frequently used comments are stored. "Unlike the case with traditional 'red-pen' corrections, with the Comment Bank it is no more time consuming for the teacher to give a clear, appropriately detailed comment than a highly abbreviated, cryptic one... In this respect the system encourages the teacher to provide feedback that is sufficiently clear and detailed to be useful to the student", (Wible et al.., 2001:303). The claim that it is no more time-consuming to mark essays using this system, is of course a big advantage, and it is hoped that there can be built on that to make the system reported on here, quick and easy to use (see chapter five).

Wible et al. (2001:303-304) say that their comment bank is essentially under the control of the teacher. The teachers can add, edit and delete comments in the comment bank to suit their own needs. The result is that (combined with the data of what the students do with the comments), the effect and effectiveness of certain types of comments can be investigated. Wible et al. provided a "starter-pack" of comments to some teachers. They say the advantage of such a "starter-pack" is that it can be used to train teachers in that it shows them:

- What type of difficulties to expect in learner writing;
- What types of feedback to give to students.

The problem with these comments is that they are not standardised. It seems that Wible et al.. (2001) used their own experience to create the comment bank. They make no mention in their article of testing for the effectiveness of the comments in the comment bank.

Wible et al. (2001:305) do say that the convenience provided by the comment bank in their system was an incidental advantage. The main advantage (and the main difference between their system and "track changes" in word processing software) is the ability to archive and track corrections for later reference and research. Their aim was therefore more oriented towards research. In this study, the aim is towards creating better feedback. Research will be the additional advantage, but for research purposes, a standardised comment set is even more important.

IWill has a function which creates statistics stating how many and what type of errors a specific student makes. This can enable the student to see patterns of errors in his/her own work (Wible et al., 2001:308-310). This is a very handy feature, programmed into this study's system as well, with the added feature that later on exercises can be automatically created based on that personal error statistics.

Wible et al. (2001:310-311) mention the further great advantages of their system as the (easy and automatic) creation of a corpus of learner writing. They quote other researchers (especially Granger) in saying that a corpus of learner writing is a very promising field of research. (For more on the creation and purpose of a corpus, see Granger, 2002.) In creating such a corpus, the data can be scanned, typed or collected electronically. In each of these instances, there is the possibility of making transfer errors from the original to the corpus. In the IWill system, the data stored in the corpus is exactly that of the original as everything is done automatically and digitally. In effect, a corpus of learner writing is created by simply archiving everyday teacher/student interactions in a systematic way. Furthermore, the additional data required by a corpus (such as age, sex, first language, education etc.) are filled in by the students themselves when they register for using the system.

Wible et al.. (2001:311) explain that one of the most fundamental aspects of learner corpus annotation is error annotation. Unfortunately, it is one of the "most tedious, subjective, time-consuming and labour-intensive aspects of corpus annotation". By creating an error-tagset that could double as a comment bank, the marking of an essay would double as initial error-annotation of a corpus, saving time and furthering research.

The biggest drawback in IWill is that the comment bank were not standardised from the start. Therefore, students still get non-standardised feedback.

In short, computer-based feedback techniques have the advantages of possible standardization, quick marking and archiving possibilities. Unfortunately, it takes a lot of research and programming to create an effective system, and teacher and student attitudes towards the computer may differ. Overall though, the computer creates opportunities which have never been possible. Utilizing the computer to the full could result in much more effective feedback.

2.3.6 What do experienced teachers do?

The effectiveness of teacher feedback may depend on a number of factors, such as students' attitude toward feedback and the nature and timing of the feedback (Storch and Tapper, 1997:245). Nonetheless, a number of studies have been done to establish the difference between the marking strategies of experienced and inexperienced teachers. Spencer (1998:82) gives a more detailed discussion of this, but the main differences can be summed up as follows:

- Teachers marked the same number of errors, but differed in *which* errors they marked.
- New teachers failed to mark important errors as they were busy with unnecessary or mistaken corrections.
- Experienced teachers ignored local infelicities or asked questions that might lead the student to come up with his or her own revision. They never marked a sentence as "too long," except when grammatical control had been lost.

- Experienced teachers are able to distinguish between serious errors and those errors where correction involves internalising new and difficult rules.
- Experienced teachers marked almost three times as many content errors. Inexperienced teachers focused more on form.
- Experienced teachers made more personal comments – making the teacher seem "not as impersonal correctors but as genuine readers who wrote substantive, constructive, text-specific suggestions, comments, and reactions".

Although the research may have some weak areas (see Spencer, 1998:82), the findings fit the criteria of effective feedback as identified in the various feedback techniques. What needs to be added though, is the notion of positive and negative feedback. Askew and Lodge (2000:7) indicate that positive feedback is seen as judgements implying satisfaction with the learner's performance. Negative feedback implies criticism and the need for changes. In their view, however, any feedback that helps learning is positive. Positive feedback should motivate by:

- increasing confidence:
- making new meaning;
- increasing understanding;
- helping to make links and connections.

In this view, negative feedback is feedback that demotivates by:

- discouraging;
- being overly judgemental;
- critical;
- giving unclear or contradictory messages;
- encouraging dependence on others for assessing progress;

The problem with this view is: "it is the experience of the recipient of the feedback which determines whether the gift [of feedback] is positive or negative", (Askew and Lodge, 2000:7).

The crucial point is that the teacher's role in student writing is not the last event in the process of writing. Marking should always provide a platform from which students

can reassess and redraft their work. Feedback must be interactive to be genuinely effective, and this requires us to find ways of correcting papers which both encourage students to think about what they have done and lead them to improve on it" (Van der Walt, Van Der Walt and Dreyer, 1994:15).

2.3.6 Conclusion on feedback techniques

Having discussed the different techniques available for feedback, and looking at what experienced teachers do, I have come to the conclusion that the notion of feedback as such is not the problem. Problems come in with the execution of the process. Sometimes, this execution is hampered by the specific technique, while other times it is hampered by the way it is received by the learner. It could be possible then to establish a checklist for the execution of feedback that could eliminate a lot of the problems related to feedback. From the above discussion, I have deducted the following checklist for ideal feedback:

2.3.7 Feedback checklist

2.3.7.1 Feedback should be:

- Consistent
- Understandable
- In a written form so that students can go back to it later
- Provided by a competent lecturer who is not simply a proof-reader
- Individualized
- Linked to teaching
- Useful for research
- Systematic
- Used by learners
- Used by lecturers
- Used by researchers

2.3.7.2 Feedback should:

- Set the learner thinking
- Place responsibility on the learner
- Encourage rewriting
- Encourage exploration of language
- Distinguish between grave and less important errors
- Should see the text as a process and not a final product
- Encourage communication between the learner and the teacher
- Should involve students in their own learning
- Raise student awareness of language features
- Sometimes focus only on specific errors (minimal marking), without ignoring the others.

2.3.7.3 Feedback should not:

- Be a complete correction
- Be "nasty" to students
- Miss any errors
- Only look at errors, but also at aspects of the text which could be better (inaptness)
- Look only at surface errors
- Be scribbled all over the text, making the feedback and the text illegible
- Waste time on the part of the lecturer

The checklist is a summary of all the commendable traits of feedback in the literature. It will be used later to evaluate the suitability of the tag set feedback (see section 5.3).

2.4 Conclusion

The term "feedback" invokes many different definitions in the literature. Part of the confusion as to the effectiveness or ineffectiveness of feedback may be attributed to this inability of researchers to agree on a specific definition. In spite of this problem (and using a working definition for the purposes of this study) it has now been established that despite problems with feedback *practice*, the *concept* of feedback is sound. Feedback has been found to be expected by learners and society, effective (to a degree) in countering fossillization, able to improve writing skills, and an effective method of consciousness-raising.

Feedback problems can broadly be blamed on learners' inability or reluctance to use it, the time it takes to provide effective feedback, the difficulty in providing sensitive, individualized, and timely feedback in such a way that the pro rata importance of specific errors are clear, the inability to recognise a clear purpose in feedback or recognise recurring patterns of error. These problems with feedback can in most cases be traced to ineffective practice. If feedback is practised in a more effective way, the concept of feedback will work better. The "best practice" checklist in section 2.3 will also answer the objectives (see chapter one) of clarity and consistency. One method of

ensuring that best practice is indeed practised is by using a standardised and computerised system such as proposed in this dissertation.

The next chapter will describe the methodology used to create and test the feedback tag set. The tag set has to answer to the above criteria of effective feedback in order to be effective *practice* of the concept of feedback (also see section 5.3).

3. METHODOLOGY

The aim of this thesis is to create a standardised correction code or set of feedback tags that could be used in a computerized marking process to provide

1. consistent,
2. effective, and
3. clear feedback.

The aim is not to start a whole new debate on feedback practice, but instead to combine already existing knowledge of both CALL and SLA in order to create possibilities for enhanced feedback within the constraints of a computerised system.
In this chapter I will explain how I went about creating and testing such a tag set. The creation and testing occurred in two phases. Phase one was the creation of a tag set. It describes where information was obtained to create the tag set, focusing mostly on the thoroughness of the set. (Chapter 5 will give more detail about the wording and structuring of the tag set.) The conclusion is reached that the set is thorough enough to meet the needs of lecturers, but admittedly, more research will be needed on an ongoing basis in order to ensure the effectiveness and streamlining of the tag set.

Phase two comprised the testing of the tag set. Explained here is the experiment used to test the effectiveness of the tag set. I will indicate exactly how I went about selecting and structuring the test, also indicating perceived shortcomings of the experiment. The conclusion is that the experiment was successful and provides an objective evaluation of the effectiveness of the tag set.

3.1 Phase One: Creating a Tag Set

In order to create a comment bank to meet the goals of effectiveness, consistency and clarity, the following was needed:

- A representative picture of errors – all the different types of errors second language learners make in writing in both form and content related issues.
- An idea of the different needs of lecturers when providing feedback – what are the elements and categories of the text or errors they need to comment on.
- A list of the most common errors commented on during feedback.

In order to realize all three criteria, information on what to include in the tag set (and how) came from three sources:

1. An empirical study.
2. A literature study on error types and frequencies.
3. The tag set used at Université Catholique de Louvain to tag the International Corpus of Learner English (ICLE) as described in the ICLE revised tagging guidelines (2003).

This section will first describe the empirical study and the data collection procedures used, whereafter short comments will follow on the ICLE tag set. Note that the influence of the ICLE tag set on this tag set is described in more detail in chapter 5 and this chapter only provides short notes. The literature study referred to will not be discussed, since the findings from the literature study were included in chapter two and has more to do with the wording of the feedback tags and the aim and focus of feedback overall.

3.1.1 Data collection procedures

3.1.1.2 Empirical study

An empirical study was conducted to identify the different types of errors that second language learners make in writing and the way lecturers react to them. Five essays from the Tswana Learner English Corpus (TLE) were randomly selected, as well as five randomly selected essays from the newly created Afrikaans Learner English Corpus. These ten essays had a combined word count of 5400 words, averaging out to 540 words per essay. The reason for using essays from these corpora, is that Afrikaans and Tswana are the two largest language groups at the Potchefstroom Campus of the North-West University.

Copies of these essays were sent to 15 different lecturers at five different universities, asking them to mark them as they would mark essays of their own students. The only requirement I imposed was that students should be able to revise the essays using the feedback given. The markers were all in the English departments of the various tertiary institutions around the country. In keeping with a promise of anonymity, these institutions will not be named.

The essays were then analysed to determine **how** the markers marked. I categorised feedback marks in four categories – "complete correction", "hint", "explanation", or "example how to correct". (From the literature study, it appeared as if these four categories should cover most feedback.) The results were analysed using a rubric in which every error was listed and a tick indicated when a lecturer used a specific category of feedback. I hoped to find an even distribution between the categories, with

more ticks going to "hint" and "example", since those two seem to be more desirable than a complete correction and less trouble than an example. Unfortunately, it turned out that by far the most feedback were either complete corrections, or simply lines, circles and strike-through lines. These lines, circles and strike-through lines I labelled "hieroglyphics", since making sense out of it was quite tricky. For an example of such a hieroglyphically marked essay, see Addendum A. After analysing the feedback in 60 essays, nothing new was yielded. No new feedback categories or techniques emerged.

The next step was to investigate what type of errors were marked. This could also give a clue as to why some claim that feedback is not effective. For this step, I analysed every single feedback mark I found in the essays. I wrote down the exact error, with a possible solution as indicated in table 3.1 below.

Page	Exact error	Possible solution	Error superordinate	Error subordinate	Marker 1	Marker 2	Marker 3	Marker 4	Marker 5	Marker 6
1	the working world **outside there**	the working world outside the university	Lexis	Vagueness: unclear or bad style	0	1	0	0	0	0
1	isn't	is not	Style	Unnecessary contraction	1	1	0	1	1	0
	most of the time	Mostly	Lexis	Verbosity	0	0	1	0	0	0
	most of the time it lies on the practical side	it is mostly practical	Lexis	Wrong word	0	0	1	0	0	0
1	most of the time it is	most of the time the real world is	Discourse	Vagueness: unclear or bad style	0	1	0	0	0	0
	in school they taught...	in school one was taught	Style	Be impersonal	0	0	1	1	1	0

TABLE 3.1. Example of analysis sheet for an essay

Errors were categorized according to the superordinate (grammar, discourse and presentation), the domain (morphology, syntax, lexis, punctuation etc.) and types (omission, superfluous, incorrect and inaptness). For the purposes of this study, a more detailed error description would not be necessary. (See chapter 5 for more detail on the error categorization.) Thereafter, I indicated which of the markers had indicated that error. (The results are discussed under "Discussion of marker behaviour".) Using the analysis an expanded tag was created for use by lecturers, together with a feedback label, which is the label the student would see when the system is used on computer. However, since it was expected that not all conceivable error types would occur in these essays, the data were supplemented from two additional sources:

- The ICLE tag set created at Louvain for error tagging the International Corpus of Learner English.
- A literature study on errors and the marking of errors.

3.1.1.3 The ICLE tag set

The ICLE tag set and the effect it has had on this tag set, are described in more detail in chapter five. This section is just a brief indication of why the ICLE tag set is important for this project.

The ICLE tag set was created at the Centre for English Corpus Linguistics at the Université Catholique de Louvain. This tag set has been used to annotate the International Corpus of Learner English and adheres to international standards. However, comments on the content of writing are not found in this tagset. The ICLE

tag set is important for this study, in that it provides the opportunity that the envisioned corpus archive of marked student essays, will be easier to convert to a conventional corpus. The ICLE tag set is simply a set of tags, focusing specifically on learner errors. It is not exhaustive and does not make provision for errors of inaptness, presentation or textual cohesion. For that reason I also expanded the sources of information to create the new tag set by consulting the literature available on the marking of errors. The findings are reported in chapter two.

These findings were then combined into one compact, but hopefully complete, list of errors with a specific tag assigned to each of these errors. These tags conform to standards for corpus annotation as explained by the Expert Advisory Group for Language Engineering Standards (EAGLES), taken from their website at http://www.ilc.cnr.it/EAGLES96/home.html. (See chapter 5 for more details.)

3.2 Phase Two: Testing of the Tag Set

The effectiveness of such a list should be tested. The main idea with the list is to incorporate it in a computer marking system so it was difficult to test the tag set without first creating the computer interface. The idea is that the computer would automatically insert the relevant tag when and where the lecturer instructs it to do so. I propose a dual type of insertion. The computer inserts an XML tag in the text being archived, while giving a more "full-text" version of feedback to the student. For example, <GTSVA> could indicate "Grammar, Tenses, Subject/Verb Agreement" in the corpus, but would be difficult to comprehend for the student who would find more use from a more full-text comment. It is of no use to try and "clarify" feedback by

making it even more cryptic than before. A student version of the same tag will therefore rather be: "The form of your verb should agree with the subject it refers to." It is much more straightforward, but still is not a complete correction, thereby forcing the student to figure out the correction for him/herself.

3.2.1 Structure of the test

To test the effectiveness of the tag set, three of the ten essays sent to markers, were chosen. These essays[3] were chosen to represent as many as possible of the error tags. As it turned out, the essays were not all of the same quality – in the language use and the content. To distinguish between the quality of the essays, they were labelled X, Y and Z in declining quality.

These three essays were then marked in three different ways, to create three different tests on the same data. The test marked with the first marking technique (further referred to as "Tag set"), was the three essays marked with the newly created tag set. The test marked with the second marking technique (further referred to as "Hieroglyphics") was the same three essays, marked by copying techniques similar to the ones used by the markers. This includes encircling and underlining errors as well as writing short comments or questions. The test marked with the third marking technique (further referred to as "Blank") was not marked in any conventional way. Learners simply received an indication of the number of errors on each page.

[3] The complete tests used for this experiment are added as addendum B

All in all, there were 20 tests of each type, which adds up to 60 tests. Each test contained 49 errors. The Hieroglyphic and Tag set tests contained some errors which were not marked. This is simply due to the fact that it would have been duplication to mark those types of error again. Some errors were also ignored due to space constraints. In marking the Blank test, only the errors marked in the other two tests were taken into account. Revision of other (unmarked) errors was ignored in all tests.

The tests were given to a first year class of Academic English, as well as to a first year class of Practical English. Both classes had about 45 minutes to complete the test. Both classes had about 50 students present at the time of testing. They wrote their answers on the test itself. (I will therefore not speak of "answer sheets" but of tests.) The tests were analysed using a four part scale, indicated by the letters A, B, C and D. The following criteria applied to each of these values:

A: The student identified the revision to be done correctly and did the revision correctly.

B: The student identified the revision to be done correctly but did the revision **in**correctly. For example: Trying to correct a spelling error, but spelling the word wrong again in a different way from the original spelling mistake.

C: The student failed to identify the revision to be done, but attempted a revision nonetheless.

D: The student did not attempt any revision whatsoever.

3.2.2 Examples of the application of the marking scale

The following table illustrates the application of the scale used for analysis. It may be argued that there may be a degree of subjectivity involved, but of course there is more than one way to revise any error effectively. Moreover, there exist even more possibilities when the text was not "wrong" to begin with, but simply not optimal. Subjectivity will therefore always be a risk, but the categories are sufficiently clear that few occasions of difficult classification occurred.

Error	*A-Type answer*	*B-Type answer*	*C-Type answer*
...students become faster grown-up...	*...students can mature faster...* *...students can grow up faster...*	*...students are more experience...* *...students become intelligent...*	*...students become more insecure...*
...the student will know why he is... (Gender bias)	*...the students will know why they are...* *...the student will know why he/she is...*	*...the students will know why they is...* *...the student will know why she is...*	*...the student will know why lucture (sic.) is...*
...and the other half for class discussions, an act, writing of something, visiting a hospital or court, etc.	*...and the other half for class discussions, an act, writing something, visiting a hospital or court, etc.* *...and the other half for class discussions, an act, the writing of something, visiting a hospital or court, etc.*	*...and the other half for class discussions, an act, visiting a hospital or court, etc.* *Note: Leaving something out is not an effective revision strategy.*	*...writing off something...* *...writing by visiting a hospital...*

Table 3.2: Example of the application of the analysis scale

A and B-type answers were considered a sign that the tag set or hieroglyphics worked. A C-type answer or D-Type answer was considered a fail for the tag set or hieroglyphics marking system, but it should be taken into account that students may

possibly not know how to identify and correctly revise some errors, or simply chose not to revise. This will be discussed in detail in later sections.

3.2.3 Shortcomings of the Test

Before any discussion of the data can take place, some possible shortcomings of the test should be brought to the attention of the reader.

The test was conducted during a normal two-period class. Students were told that I was conducting an experiment to find the most effective method of marking. They were told that they had to correct the errors indicated on the tests. This word choice ("errors") could have conditioned them to simply look for instances of language use that was wrong and not consider the text as a whole to be revised.

The three tests (Blank, Hieroglyphic and Tag set) were randomly handed out to students, with the explicit instruction that they were to do their own work. The students did not always obey this instruction and were frequently caught looking at each other's work, especially during the first few minutes. If a student was caught looking at the person next to him/her more than once or twice, they were asked to move to a different seat. The possibility still remains that students did help each other, since they knew that this was not a formal testing situation and they would not face any prosecution for helping each other.

Another problem was with the positioning of the tag set marks. The "Insert Comment" function of MS Word was used to insert the tags. This is not the way that

the computerised marking system is going to display these tags. I want the tags to display as pop-ups (mouse-point pop-ups) at the exact position of the error. (See chapter 5 for more detail.) MS Word inserts the comments in a way that makes it difficult to see which comment is linked to which word in the sentence, especially if there is more than one comment in a line. Steps were taken to try and eliminate this problem, but it may still have caused some inconvenience for the students. If this problem is rectified, it may lead to an improved performance of the tag set in the actual application.

Students taking the test were asked orally and instructed in writing to indicate what they considered the errors to be. Most of them also ignored this request, so the few answers obtained could not be used in this discussion. In addition, students could not be forced to answer all the questions. This resulted in a lot of D-type answers where no attempt at a revision was made. One cannot be 100% sure whether students did not know how to do the revisions, or if they were simply in a hurry to finish the test and leave.

Students were not working on their own texts. This could have influenced their attitude towards the test, but it is almost impossible to determine the effect of attitude on the performance of a student in a testing situation.

Finally, the complete tag set could not be tested due to constraints on time. The 49 errors included in the test, were however chosen to give a representative picture of the tag set as a whole.

3.2.4 Conclusion

The experiment was conducted on a representative group of students, with random sampling (to ensure an equal number of tests) and clearly formulated categories for classification. The problems with the test should not influence the results to such an extent as to make the findings irrelevant, ensuring that an objective evaluation would be possible to ensure that the tag set answers to the objectives of clear, consistent feedback and best practice.

4. DISCUSSION OF WHAT LECTURERS TYPICALLY FOCUS ON IN MARKING

4.1 Introduction

The creation of the tag set was the primary goal of the exercise. However, it provided a good opportunity to analyse the behaviour of the markers and to see on what exactly they focused[4]. In order to establish best practice in feedback techniques, it is also important to know what NOT to do. An evaluation of what markers typically focus on would provide the opportunity to see whether the markers satisfy the criteria of best practice as identified in section 2.3. The initial experiment (to see what lecturers mark) was conducted on three essays, marked by six different lecturers. This was a small test group, but the markers were volunteers from three different universities, and all of them work in the English department. They also had ten essays each to mark, so a stable pattern would emerge by that time.

4.2 Error Domains Marked

Table 4.1 illustrates on what the lecturers who had marked the essays had focused. Please note that the domains used in this chapter are more descriptive than the

[4] The results in this chapter are discussed in terms of the categories that are only presented in subsequent chapters (see chapters five and six) but the information in this chapter was needed to draw up those categories. This chapter (and subsequent ones) were developed together, iteratively, to develop the tag set.

domains used in chapter 5. The reason for this is simply to enable a finer analysis of what the lecturers had focused on.

Error domain	Total marks	%
Lexis	97	37.6%
Syntax	84	32.6%
Capitalization	16	6.2%
Punctuation	16	6.2%
Style	10	3.9%
Factual correctness	8	3.1%
Punctuation error	6	2.3%
Cohesion	4	1.6%
Morphology	3	1.2%
Coherence	3	1.2%
Presentation	2	0.8%
Layout error	1	0.4%
Layout	1	0.4%
Layout inhibits reading	1	0.4%
Wrong word	1	0.4%
Syntax and lexis (combined)	3	1.2%
Syntax and punctuation (combined)	1	0.4%
Wrong word form	1	0.4%
	258	100%

Table 4.1: Error Domains Marked

From the above table (4.1), it is clear that the most ink was shed on lexis and syntax with both scoring above 30% and with no other domains scoring nearly as much. This focus of the lecturers is not wrong. Some sentences are truly so ungrammatical and full of errors, that it is difficult to understand what the learners were trying to say. Overall, attention mostly went to the surface structure of the text and not to the content and ordering of information. Coherence and cohesion did not get that much feedback either. Some lecturers made a little note at the end of the essay advising students to look at their structure. I found this too vague to be useful. If one therefore judges texts by simply looking at feedback statistics, one might make the erroneous assumption that the texts are clear, well-structured and all have good arguments

backed up by hard facts or innovative thinking. This is exactly the impression we want to counter. If a text is full of surface errors, it is very difficult to ignore those errors and simply look at the ordering and reasoning. If the computer takes care of most of the surface errors, it could be quite a weight off the mind of the marker (see section 8.5).

4.3 Consistency in noticing similar errors

Marker correlation on subjective topics is difficult. However, since my interest was not on correlating the scores of the essays (which did differ quite substantially) I was more interested in seeing a correlation between the numbers and types of errors indicated by the different markers.

Error Domain	Marker 1	Marker 2	Marker 3	Marker 4	Marker 5	Marker 6	Average	Total	Percentage marked
Syntax and punctuation	0	1	1	1	1	1	0.8	1	83%
Capitalization	8	14	14	14	11	14	12.5	16	78%
Wrong word	0	1	1	1	1	0	0.7	1	67%
Morphology	0	2	1	2	1	2	1.3	3	44%
Syntax	14	45	34	34	40	50	36.2	84	43%
Syntax and lexis	0	1	0	1	1	2	0.8	2	42%
Cohesion	0	3	1	2	3	1	1.7	4	42%
Lexis	17	41	50	39	36	43	37.7	97	39%
Style	4	3	2	5	6	2	3.7	10	37%
Punctuation error	1	2	2	3	1	2	1.8	6	31%
Punctuation inaptness	0	6	6	4	3	8	4.5	16	28%
Presentation	0	1	0	0	1	1	0.5	2	25%
Factual correctness	0	4	2	1	1	3	1.8	8	23%
Coherence	0	1	0	1	1	1	0.7	3	22%
Layout error	0	1	0	0	1	3	0.9	5	84%
Wrong word form	0	0	0	0	0	0	0.0	1	0%
All groups	44	127	114	108	109	134	106.0	258	41%

Table 4.2: Consistency of markers in noticing the same errors

From table 4.2, one can deduct that **all** the lecturers focused much more on the surface elements of the texts. It is therefore obviously not an isolated phenomenon. Subsequently I looked at the attitude the lecturers marked with. For this, I classified the errors according to four general types: incorrect, inaptness (could have been better), omission and superfluous (see section 6.1 for more details). I considered inserting a type "praise", but decided against it. Praise was normally kept to a small note at the end of the essay. Table 4.3 indicates the results.

4.4 Error types marked by the markers

Table 4.3 illustrates the error types the markers focused on. Error types are explained in more detail in chapter 5. Briefly anticipating that discussion, error types refer to the focus of the errors. Was the indicated error an error due to incorrectness, inaptness, omission or overinclusion? It should be obvious that feedback on something which is not inherently wrong, is in fact pointing to an error of inaptness – it could have been better.

Error type	Count	Percent
Incorrect	125	48.1
Inaptness	57	21.9
Omission	48	18.5
Superfluous	30	11.5

Table 4.3: Error types

From this we can infer that lecturers focused on incorrectness. Less often, they tried to improve something and very seldom they indicated anything as unnecessary. I am aware that these categories do not have hard and fast borders. It may therefore be

possible that an omission may render a sentence "incorrect". However, the overall attitude of focusing on obvious errors (incorrectness), remains clear.

TYPE	Marker 1	Marker 2	Marker 3	Marker 4	Marker 5	Marker 6	Average	Total	Percentage
Incorrect	29	65	71	65	58	67	59	125	47%
Inaptness	11	30	15	17	25	34	22	57	39%
Omission	2	23	22	13	19	20	17	51	33%
Superfluous	3	11	8	14	9	13	10	30	32%
All Groups	45	129	116	109	110	134	107	260	41%

Table 4.4: Error types per marker

From table 4.4 above we can deduce that the markers do not vary much with regards to the number of errors they had picked up. Marker one ignored/missed the most errors, while marker six found the most. It is possible that marker one decided to focus on some errors only, but she gave no indication of that being the case. Researchers (such as Moletsane, 2002) have indicated that this type of selective marking may induce fossilization and should be discouraged. It seems as if the other markers tried to find all the errors. This no doubt took a lot of time and thinking, but the system proposed in this study should speed up that process (see section 5.1.2 and section 8.5 as well).

4.5 Conclusion

It seems clear that markers have a tendency to look for incorrectness in the surface structure of the text. Markers do not attempt to provide feedback on textual organisation even though it seems as if they try to mark all the errors in the text. Markers seem so flustered by the large amount of surface errors that they do not comment on the things which could be seen as correct, but could most definitely have been better (compare the feedback checklist in section 2.3.8). However, markers are not the only ones focusing on the surface structure of language. Studies to test the language proficiency of students also utilize testing techniques focused on the surface elements of a text (see Hatting 2005 for an example). This could imply that language proficiency is commonly seen as only a surface element, enforcing the need for more attention on the surface structures of a text. Lecturers aim at bettering their students' language proficiency and then simply focus on the most obvious level.

In short then, markers are not able to adhere to best feedback practice as identified here. The inability may be due to personal ineffectiveness, an incorrect focus or being overloaded with work.

These are important facts to know. In the first instance, it could point to teacher training which is not as practical and effective as one would have hoped. On the other hand, the possibility is there that these lecturers were trained in marking techniques, but were overworked, or possibly not aware of writing theory. Either way, additional training for lecturers could well be necessary. It is also important to try and get the marking system as automated as possible in order to take some of the load from the

lecturers. By this is meant that the computer system should be able to be of assistance in identifying possible errors (see section 8.5.). This could possibly drastically reduce the amount of time that lecturers have to spend on marking.

5. WHY DID THE INITIAL TAG SET LOOK THE WAY IT DID?

The method followed to create the tag set has been described in chapter three. This chapter presents the reasoning behind the original version of the tag set used in the experiment. I will try to justify the inclusion of the tags, and explain why I structured the tags and error categories the way I did. This will be done by placing the error categories within a theory of language, seeing as feedback has been established as an important tool for language learning (see section 2.2.2.2).

I will also give an indication of how the tag set fits the criteria of "best practice" as identified in chapter two (section 2.3.7). The reasoning behind the tag structure is explained in terms of user friendliness for the marker in order to save time (see section 2.2.2.1.13) and clarity and comprehensiveness for the learner (see section 2.3.7). Since the idea was to first refine the tag set before doing the extra work of creating the XML tags (section 5.1.1), this version of the tag set does not contain XML tags yet. It does contain examples from the original empirical research, as well as the tips and examples which will be present in the final version. The provisional tag set (henceforth TS1) is included in full as Addendum C.

5.1 The tag structure

Each tag has three parts, which will be visible to different users (students, lecturers, and researchers) of the system. The tag has a label, the actual feedback tag, and an

XML tag. Each tag also has an example or explanation attached to it, to assist in marking. Each tag is classified according to the error type and the error superordinate it falls under. (Error type and error superordinate will be discussed in the section on error classification in section 5.2. Also see section 2.1.4.2.)

A typical tag will therefore have a breakdown as illustrated in table 5.1:

Superordinate	Domain	Label	Type	Feedback tag	Example	XML
Discourse	Style/register	Bias	Inapt-ness	Be careful of gender bias.	Use this when a learner makes use of language which could be considered sexist.	DSRB

Table 5.1. Example of a tag in the feedback tag set

5.1.1 XML-tag

As mentioned above, the XML tags are not yet included in this version of the tag set, since the tag set first had to be tested. The final version of the tag set presented in chapter 7, do have XML tags. The specifications for XML will be described here. Note, however, that XML is not used *only* for language-related computer applications. A full discussion of XML falls outside the scope of this dissertation and only the reasons why XML is at all relevant to this project will be discussed here. The full use of XML for corpora annotation is a project on its own. This discussion of XML is simply for purposes of context and background. For more detailed information on XML, visit www.xml.com.

The simple reason to use XML is that XML tags are the ones that will ensure that the marked essays will be partly annotated and searchable in the archive or newly created corpus. These tags will be utilized by the researcher, who has to analyse student writing from the corpus, but it can also ensure that the corpus will adhere to international standards of corpora.

5.1.1.1 What is XML?

XML is a computer markup language that is used extensively on the Web to create interchangeable, structured documents. It is a type of computer language that is used to define parts of a text. XML is derived from SGML, but is cheaper, since it is easier to use. It is still compatible with SGML.

The abbreviation, XML, stands for Extensible Markup Language, meaning that the markup language can be manipulated for specific purposes. This manipulation of the markup language does not affect the basic structure of the markup language.

XML is reusable (more on that in section 5.1.1.2) in the sense that any computer can read the information, as long as you provide the computer with enough information about it. This information is described in a document called a "document type description". The XML specifications in the tag set are the document type description for this project, in that it explains what a specific tag refers to.

5.1.1.2 What XML is useful for? Reusability

In spite of the fact that XML can be read across computer platforms, it also has other uses, the most important of which is that it promotes reusability. EAGLES (the Expert Advisory Group on Language Engineering Standards) is a group of European language experts. (For more information, consult the EAGLES website.) Their goal is to promote reusability in the language sciences. Reusability has gained significant ground in the past few years. It implies that the sources available for research and development should be used, organised, and equipped in such a way as to be multipurpose. This is where XML comes into play.

Reusability depends on:

- Common guidelines
- Adaptability of products to each other
- Standards

EAGLES see standards as the most important of the above three, in that by adhering to standards, the other two requirements can be met. EAGLES tries to create standards which will ensure that the other two factors are met. The more common and accepted these standards are, the better the possibility that programs could be written that are adaptable world-wide. Comparative studies will then be possible and multi-purpose programs could be created.

Reusability is important for the current project in the sense that the feedback should have multiple uses. The feedback should in the first place be easy to use for the lecturer doing the marking. It should be used by the students to learn from their

mistakes, and it should be useable for a researcher to do research. This last use is the only one that really applies to the XML specifications, and it is the main reason why XML is taken into consideration in the creation of the feedback tag set. The XML specifications will ensure reusability in a much bigger context – that of global reusability. A corpus annotated with XML tags will be searchable in most corpus research tools.

The claim that XML creates an opportunity for global research usability is not an exaggeration. The guidelines created by EAGLES for standardization are a joint attempt by a large group of linguists to create good guidelines for, amongst others, corpora. The resources for which EAGLES attempts to create standards for, forms the basis of a lot of modern research and the importance of the work done by EAGLES should not be underestimated. (Please note: Some of the documentation created by EAGLES refer to SGML which is also a markup language. SGML is, however, more expensive to use since the design is more difficult. For that reason, XML will be used in this project.)

5.1.1.3 How does XML look?

XML is basic computer markup language. An XML tag will therefore always start and end with a triangular bracket. XML tags will be inserted on both sides of the text object being tagged, with the ending tag being indicated as such by the use of a front slash. A tag in a text (if seen) will therefore look like this: <tag>object to be tagged</tag>. It is important to note that these tags will not be seen by the students in their feedback interface, or by the marker in the marking interface. These tags are

invisible (not displayed) to the human reader, but will tell the computer how to handle certain parts of the text. This type of insertion of the error tags will be done automatically by the computer, which is why the marker will have to highlight the specific part of the text that has the error in it, to indicate to the computer where to insert the relevant tags.

These tags are always "nested". This means that tags are to be closed in the order in which they were opened. If you have a tag structure such as <tag1><tag2><tag3> for example, it should be closed in the order </tag3></tag2></tag1> giving you the total structure of **<tag1><tag2><tag3>***thing to be tagged***</tag3></tag2></tag1>**. It is this kind of structure which makes XML reusable.

It is important to distinguish between the tag structure and the tag itself. The tag structure has been illustrated above in bold, while the tag content is everything within the triangular brackets. It is this tag content which is designed along with the error tag set. The tag structure is standardised to the same extent as XML is standardised (which is quite extensive, seeing the wide use of XML on the web). The tag content is not standardised (which is why there is a need for a document type description). This is the reason why I try to link it as closely as possible with the tag set used in the ICLE project to provide a degree of standardization – at least with regards to language engineering.

The tag content (much like the tag structure) has a hierarchical structure. This means that a tag will have an indication of the specific category it falls under, where after more information will follow. A tag will therefore first indicate the domain it is

relevant to, where after the rest of the tag will follow. In the ICLE tag set (to be discussed in the next section) the letter "G" is used to indicate "grammar" The ICLE tag for a capitalization error of a noun, is GNC, which is derived from Grammar, Nouns, Case. Note how the tag is structured giving the more contextual information first. First you know that the error is a grammar error, then you know that the grammar error has to do with nouns specifically, and finally you get to what the error is. (Note that in the example, it is simply the tag letters and not an indication of the full XML error tag.) This hierarchical structure is a reason for structuring the tag set into domains as well. (See 5.2, on the categorization of errors.)

5.1.1.4 XML, corpora and error tagging

It is important to note that by marking essays with the marking interface, the essays will not be completely annotated to be in corpus form. It is not yet completely annotated, since that would entail annotating all special characters (such as "-") and inserting paragraph markers and numerous other annotations. The idea is simply that by marking with this system, the lecturers are creating a partly error-tagged corpus. Inserting the other markup tags will be all that is necessary to have a fully error-tagged corpus.

The marking system will therefore result in a lot of tedious error identification and classification (normally reserved for the creators of a corpus) to be done in the process of providing feedback, which is expected of lecturers anyway. In order to adhere to the idea of reusability, the tag set tags are linked as closely as possible to the tag set used by Louvain. The marking system also have the additional benefit that it is not only the creators of corpora who would save time – according to Wible et al. (2001),

such a marking system also saves time for the marker in that a lecturer can highlight and click rather than write an extensive comment.

5.1.1.5 The tag set used by Louvain in the ICLE project

Granger and Meunier (2003) distributed the refined tag set used to error tag the corpora in the International Corpus of Learner English (ICLE) project. The ICLE project (like the name says) has the purpose of creating an international corpus of learner English. Given that the different error-tagged corpora will all come from different countries, it was important to adhere to a strict set of standards. These standards were created for a different purpose and are therefore not directly applicable to this project.

The ICLE tag set contains 40 tags and makes provision for formal errors, grammar errors, punctuation errors, lexico-grammar errors, lexis errors, register errors, style errors and word errors. These differ from the domains identified in this study. The main reason for the difference is due to the structuring of the tags to ensure the user friendliness for the marker. More on this in the next section. Although the error categories used in the ICLE tag set do not correspond on a one to one basis with the error tags in the current project, they are similar in some cases. Where they are similar, the same XML tags were used to ensure comparability. The ICLE tags do not lend themselves to reusability in the way that the tags in this project do. As a result, the similarities turned out to be less than expected, probably due to differences in the purpose of the two tag sets.

5.1.2 The tag label – what the marker will see

The tag label is basically what the lecturer will see while marking the essay. Each label will be a freestanding button, or a button on a drop-down menu. Taking into account that lecturers will not want to waste time thinking about the nature of the error, I tried to make these "buttons" as transparent as possible. A button labelled using specialized grammar terms will not be user friendly. This is especially important if one takes into account that the system might later be used by teachers, who, according to Webb (2002) and Coetzee-Van Rooy and Verhoef (2000) might not be so very proficient themselves.

The problem of transparency for the marker was solved by looking at the level at which the error occurred. With this I do not mean the analytical context (such as the error superordinate) but rather where it occurred physically. If the error was in a word, the label will start with "Word-" after which the rest of the label will follow. For example, if a learner has used the wrong form of the adjective, the label will be "Word – word form wrong – adjective". Although the form of the adjective is dependent on the sentence context, the error itself is not in the whole sentence, but simply in the word. The same goes for the omission or overinclusion of words. If a word has been omitted from a sentence, the label will start with "Word-" and not with "sentence" since it is a word that is missing. If sentence construction is wrong, the label will fall under "Sentence".

I realise that for some errors, these "levels" may overlap and it would be necessary to include a tag from more than one category. This will of course be possible, but markers will be advised to add their own comments in some cases, in order to explain

what they see as problematic. This approach to the structure of the tags corresponds to the approach used in the ICLE project, with their hierarchical structure. The difference here is that my hierarchy is not based on grammatical categories but on the surface level of the text. This was done simply to ease the process of marking, so that a marker does not need to keep thinking of grammatical categories the whole time.

An additional reason behind this, is that this labelling technique can keep the interface uncluttered. The marker simply has to decide if the error is on the level of "word", "sentence", "layout", or "coherence" and he/she does not have to look through all the available buttons. Once he/she has decided where the error lies, he/she can click on the relevant button, opening up a list of other buttons all with relevance on that specific level. The decision-making is therefore limited, in that all the buttons do not need to be read looking for the relevant one, but only the buttons in the specific level. The idea is graphically illustrated in figures 5.1 and 5.2 below. Note, however, that these two figures are not exactly the way the final version will look, but simply serve as illustration.

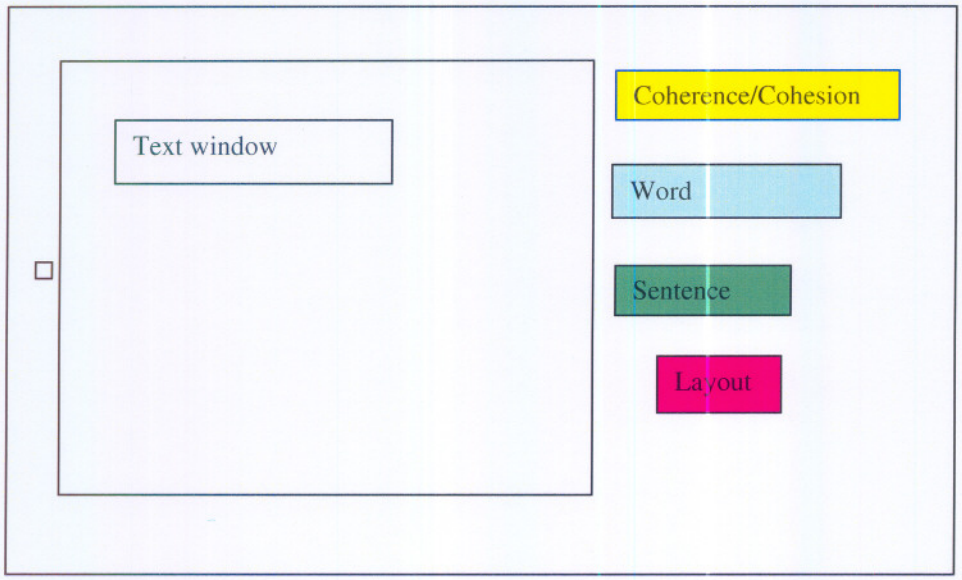


Figure 5.1: Uncluttered view for marker, showing only the "level buttons"

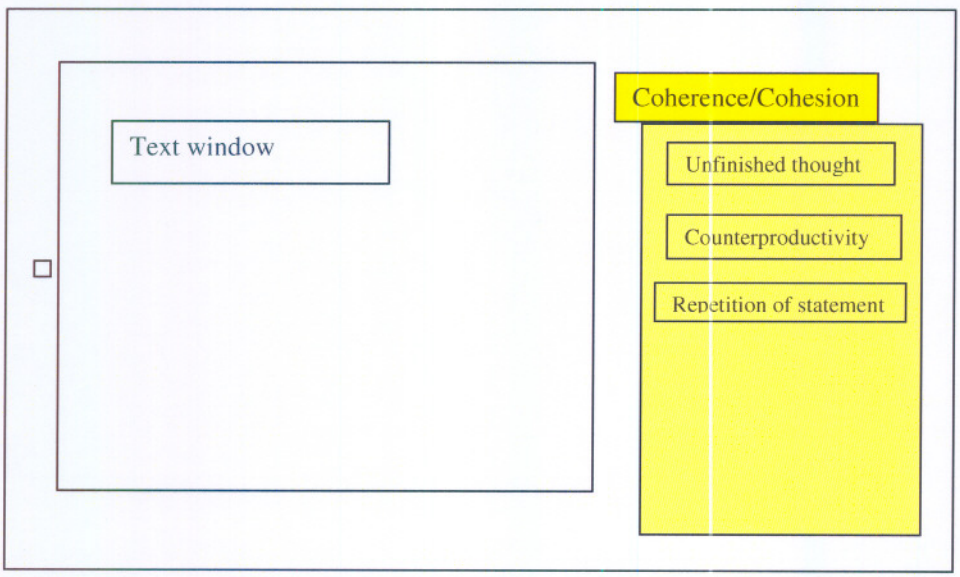


Figure 5.2: A drop-down menu for the level of Coherence/Cohesion. Note that the other "level buttons" are now obscured, making room for the drop-down menu and limiting the choices for the marker to go through.

5.1.3 The feedback tag – what the student will see

The final part of the tag structure is the tag seen by the student. The idea is that students will receive their feedback in an html document which can be read in any web browser, such as *Microsoft Internet Explorer* or *Mozilla Firefox*. Marked errors will not be shown completely, but their positions will be indicated by using colour. If, for example, a word has been left out, the two words on either side will be coloured. The student then has to move his/her mouse pointer over the error to see a "pen-point popup" in which the complete feedback tag will be written.

The tag itself has been written to be as simple, clear and to the point as possible and the experiment (chapter 6) proved that it was successful in the sense that students seemed to understand the tags. In a later stage of development of the marking system as a whole, the idea is that a student will be able to ask for additional information on the error, by asking for an example. Due to constraints on time, funding and availability of research partners, this feature could not be programmed into the first version of the marker system. These examples will need to be researched intensively before they can be added onto the system, to ensure their usability and pedagogical effectiveness.

The idea with the coloured indication of error positions is that the lecturer can decide which errors the students should focus on in a specific writing assignment – mimicking selective marking. If the lecturer simply wanted the students to pay attention to sentence construction, the rest of the errors do not need to be displayed in colour. However, I would still advise lecturers to mark all relevant errors for purposes

of archiving and in case the students are curious and want to see the rest of their errors.

I do not assume that the tags are perfect as they are. The system will be tested and refined some more. This continued testing applies to the tag content, the way the tags are written, as well as the way the feedback is presented to the students in the end through the web browser.

As mentioned, the tag set provides three different tags which will be visible either to the student, the marker or the researcher. But the tags must also be classified into categories, which brings me to the categorization of the errors.

5.2 Error classification: Where do errors fit into language?

Thousands of books have been written on the nature of language, ironically using the very medium they are describing. To describe something as complex as language is very difficult, but may be very useful, especially in a teaching or research environment. For those reasons, the errors in the tag set were positioned in a theory of language – the functional language theories of M.A.K. Halliday and C.M.I.M. Matthiessen (1990, 2004) and T Givón (1989). They are of supplementary use for this tag set in that the theories can facilitate research.

5.2.1 Towards a theory of language aimed at errors

Halliday and Matthiessen (2004:24) explain that language has strata. They explain the interaction between these strata in graphic form as illustrated in the following figure:

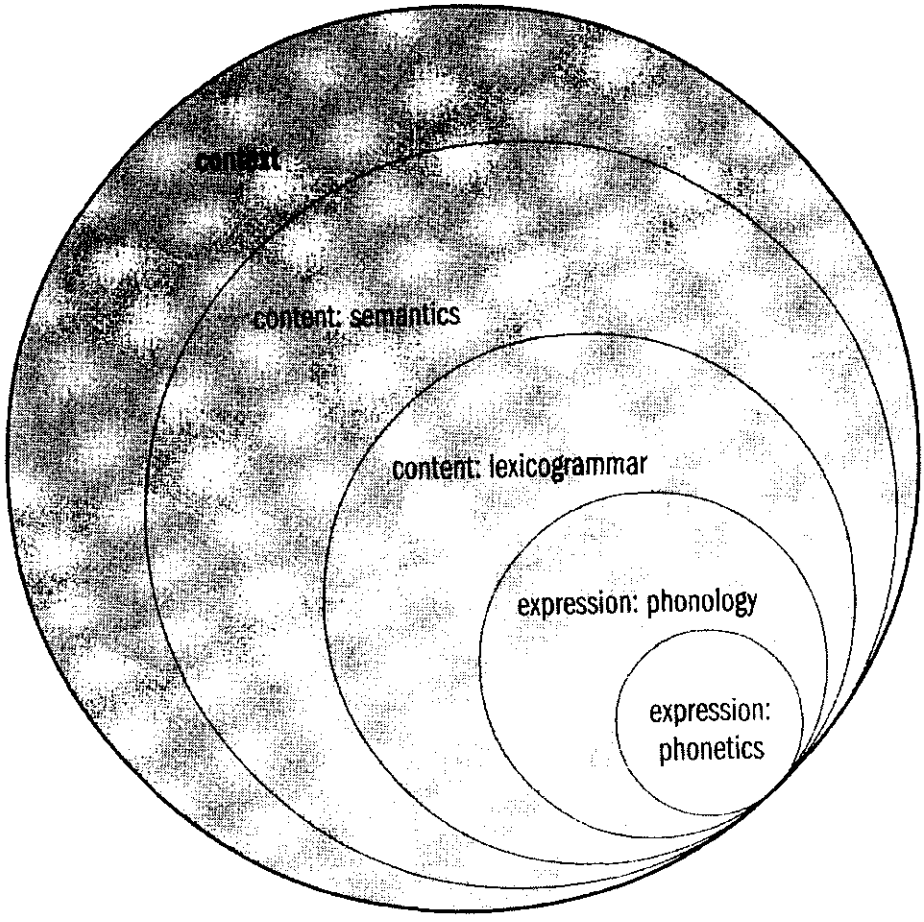


Figure 5.3. Halliday and Matthiessen's strata of language.

A textual breakdown of this would present us with:

Context

...which contains **semantic content**

...which contains **lexicogrammar content**

...which contains **expression** by **phonology**

...which contains **expression** by **phonetics**

...gives us **language**.

Figure 5.4. Breakdown of Halliday and Matthiessen's strata of language

Halliday and Matthiessen (2004:24) explain that adult language has more strata than child language. Child language has a content and expression stratum, whereas adult language has two content strata (lexicogrammar and semantics) and two modes of expression – writing and speaking. In this study, expression can only refer to writing on the area of layout, punctuation and capitalization, which has more to do with the way a text looks. Although layout, punctuation and capitalization may have grammatical value (punctuation could be argued to belong to syntax, for example) they are elements of written texts only. One cannot hear punctuation or capitalization.

A similar breakdown of the (also concentric circle) theory of language presented by Givón gives us a less elaborate view of language:

Discourse

...which consists of **clauses**

...which consists of **words**

...gives us **human communicative code.**

Figure 5.5 Breakdown of Givón's theory of language

It is clear that both Halliday and Matthiessen and Givón make provision for something outside of language. Halliday and Hasan (1990:5-6) make this explicit in their explanation of "context of situation", a term coined in 1923 by Malinowski. Halliday and Hasan (1990) describe this context of situation as "the kind of commentary that placed the text in its living environment". From this very basic beginning, Hasan (2002) has refined a model (figure 5.6) of context in relation to culture and language.

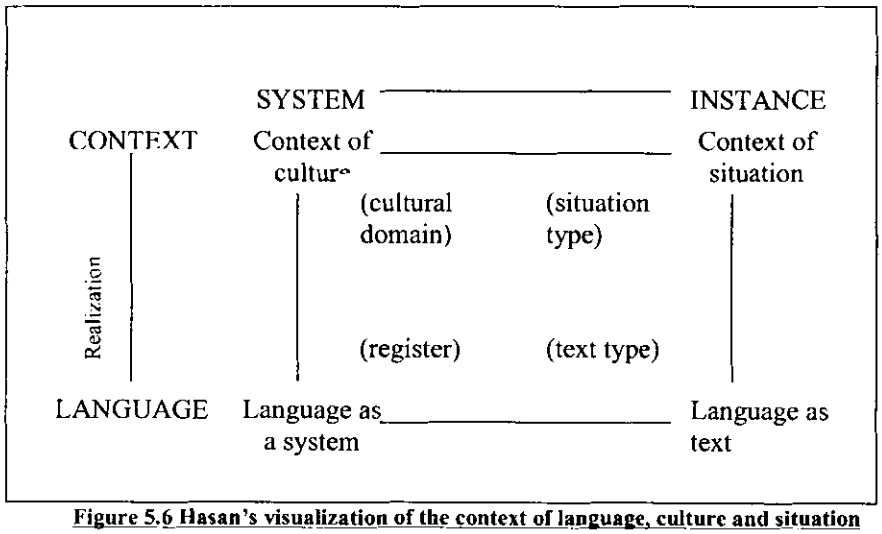


Figure 5.6 Hasan's visualization of the context of language, culture and situation

In this model, language is seen as a relative of a realization of context. Similarly, language as a system is seen as a relative of a realization of the context of the culture. Language as a text, which is dependent on language as a system, is also seen as a relative of a realization of the context of the situation.

This very interrelated and abstract view of the interrelation between language and context makes it easier to understand that errors made by learners fall within a certain context. Errors (although not ideal language use) is still an attempt at language use. Although the finer details of the context will differ for each individual, the context of the teaching situation is the same for all. It is important to remember that feedback can be a teaching tool, and that is why the feedback tag set was structured to enable teaching facilitation.

One can apply the strata proposed by Halliday and Matthiessen and the "components of human communicative code" of Givón to the domains of my error tag set. The different domains of my error tag set (in alphabetical order) are:

- Capitalization
- Coherence
- Cohesion
- Factual correctness
- Layout
- Lexis
- Morphology
- Positive comments
- Punctuation

- Style/register
- Syntax

To fit them all into the stratums presented by Halliday and Matthiessen, would fit them here:

Stratum proposed by Halliday and Matthiessen	Corresponding domains in tag set	Givón's communicative code	Tag set superordinate
Context	Factual correctness	Discourse	Discourse
	Positive comments	Discourse	Discourse
Content: semantics	Style/register	Discourse	Discourse
	Cohesion	Discourse	Discourse
	Coherence	Discourse	Discourse
Content lexicogrammar	Syntax	Clause	Grammar
	Morphology	Words	Grammar
	Lexis	***Words***	***Grammar***
Expression	Layout	Not handled by Givón	Presentation
		Not handled by Givón	Presentation
	Punctuation	Not handled by Givón	Presentation
	Capitalization		Presentation

Table 5.2: Error Domains and Error Superordinates applied to Halliday and Matthiessen and Givón's theories of language

In my tag set, I make provision for superordinates of grammar, discourse, and presentation. The positions occupied by these superordinates will be discussed shortly. First it is important to explain the above table. Since these strata fit into each other (as shown by Halliday and Givón), some of these domains may seem to fit in another stratum. Lexis, for example, may be closely linked to style and register, but on a smaller scale overall. The question to be answered here is which is the biggest stratum a specific error domain will fit into.

Style/register, cohesion and coherence fall under the superordinate of discourse in my tag set. This corresponds to the major functions of language, as described by Givón (1993:22), which identifies these as part of discourse. This specific categorization is especially important if one remembers that *academic* competence is important here. Knowledge of the academic register needed at university is important. Proficient communicative (social) skills will not pass as sufficient academic communicative skills at University despite students perceiving their communicative skills as adequate (Coetzee-Van Rooy, 2000:161-166).

Feedback on factual correctness has to do with the context in which the text falls. Positive comments (although directed at a small part of the text) fall in the broad context of teaching (as feedback should) and is therefore classified in the stratum of context.

Style/register may also be considered to fall under the stratum of context, but a finer distinction is at work here. Academic discourse is a context, but the specific style and register of an academic text has much to do with how a text will shape meaning and understanding. An informal style of writing will not make the writer seem credible.

Cohesion and coherence have to do with the way in which an argument is presented as clearly as possible, to make it easier for the reader to follow and understand – creating meaning.

Syntax, articles and morphology are all grammatical concepts, whereas lexis is on the stratum of lexicogrammar. It is part of the content.

Phonetics has to do with the sounds of words. In writing, we do not work with sounds, but over ages, man has devised a system of upper and lowercase letters and punctuation to assist a reader in knowing how to read a text. The layout of a text is also a way to ensure that a reader reads it the order it should be – an easy example of this that we read from left to right in English. The way a text is presented on the page may influence the way a reader experiences it, but in the end the layout has very little to do with the meaning of the text.

5.2.2 Error Superordinates

My tag set is divided into three overarching categories, or superordinates. These are "grammar", "discourse" and "presentation". They correspond closely to the categories proposed by Givón (1993), with the exception that I group his "words" and "clauses" under the category of grammar. When working with errors, the distinction between an error at word level or clause level is sometimes very vague and can therefore much more effectively be grouped under one heading.

My superordinate of "presentation" is not handled by Givón, but may fit Halliday and Matthiessen's "Expression". Halliday and Matthiessen's phonetics has to do with how the language sounds. In written language, we are not working directly with sounds, but unlike in spoken language, we have a very definitive way the language

"looks" in a text. Therefore, the superordinate can be argued to fit Halliday and Matthiessen's model.

The superordinate of presentation in my tag set also includes typing conventions. One may argue that this is not relevant to the development of writing skills, but from the empirical research done, it seems that some lecturers want to be able to comment on these issues. These lecturers probably want a text to be professionally finished and since the system is aimed at assisting lecturers as well, their needs should also be taken into account.

My superordinate of discourse is the same as Givón's discourse, but as is apparent from the above table, I make provision that this discourse is affected by context and the semantic content of a text. As seen from the positioning of these factors in Halliday and Matthiessen's model, this is a relevant area of concern. (See the explanation of the position of "style/discourse" above.)

It should be apparent that the tag set and error categories created in this project are in line with functional theories of language. This type of theoretical foundation for error categorization presents us with the opportunity to see errors within the bigger picture of language as a whole. The finer breakdown of the tag set further allows us to investigate problems in student writing from a bird's eye view, right down to the bedrock of individual errors.

Unfortunately, since feedback should be used to facilitate teaching, even these fine categories of error are not enough. They do not make provision for the "why" of the

error, but only for the "what". To account for the "why", I divided the errors into error types. During the empirical study, it was noticed that the identified errors fall into four broad types: Errors where something has been left out, errors where something unnecessary has been included, errors where something is incorrect, and errors where something is not formally incorrect, but could have been better. These four error types are therefore named "Omission", "Superfluous", "Incorrect" and "Inaptness".

This distinction is very important to make. An analysis of the error types may indicate to the lecturer if he/she is focusing on incorrectness or if he/she is paying enough attention to raising the students to the next level – i.e. focusing on the areas where a student can improve his/her writing (labelled inaptness here). An analysis of these error types may also provide clues as to where students are avoiding using new or difficult constructs. The latter will unfortunately only be possible if one is sure that the lecturers marked effectively and thoroughly. With time and further research, the computer may be able to assist both the marker and the researcher in finding these areas of avoidance.

These error categories are evident in the ICLE tag set, but are not spelled out as concretely as they are here. Also, the ICLE tag set does not seem to make provision for inaptness. In the corpus scenario, something seems to be either incorrect, or correct. This does not facilitate education, but then again, the corpus was designed to describe and analyse; not to teach. Here is evidence of where feedback fulfils a bigger function than simply error correction.

Having established that the theoretical foundations of the error tag set is sound, I turn my attention to the proposed use of the tag set in feedback. Is the envisioned use of the tag set based on sound feedback theory? Does it conform to best practice?

5.3 Feedback using the tag set and marking system: Is it "Best Practice"?

The qualities of "best practice" in feedback were established in section 2.3.7. To establish whether the feedback procedure as outlined in this dissertation fits those criteria, those findings will now be used as a check list to determine the quality of the proposed marking technique.

CRITERIA Feedback should be:	Marking-system performance	Comment
1. Consistent	Complies	
2. Understandable	Complies	See chapter 6
3. In a written form so that students can go back to it later	Complies	The system will force students to look at their feedback in order to do revision and exercises.
4. Provided by a competent lecturer who is not simply a proof-reader	Not system dependent	Although the system will be used by professional language educators, there is no way to ensure that this will always be the case. The system will be developed further to assist the lecturers also.
5. Individualized	Complies	Although a standard tag set is used, the opportunity to provide additional comments is there. The revision exercises will also be focused on the weaknesses of the individual.
6. Linked to teaching	Not system dependent	The feedback on the writing will enable the lecturers to focus their teaching, but the onus is still on the lecturers to make use of the analysis from the computer.
7. Useful for research	Complies	The system will automatically archive marked student writing, creating a big corpus of written learner English which will be useful for research.
8. Systematic	Not system dependent	Although the system will assist in marking, the way the different lecturers go about their marking is not the responsibility of the system.
9. Used by learners	Complies	Learners will get revision exercises based on their personal shortcomings. The idea is that students' marks will not be accepted before their revision is done. The degree to which these revisions will assist them in their future essays, will have to be seen.
10. Used by lecturers	Not system dependent.	The system provides the opportunity for lecturers to make use of the information provided by the feedback, but the lecturers themselves will have to decide how or if to use it.
11. Used by researchers	Not system dependent	The system provides the opportunity via the corpus, that researchers can use the information, but once again it depends on the researcher and not on the system.

CRITERIA Feedback should:	**Marking system performance**	**Comment**
1. Set the learner thinking	Complies	The tag set was designed so that all-out correction is not possible. In order to revise the problems, students will have to figure out for themselves how best to solve the error.
2. Place responsi-bility on the learner	Complies	The possibility exists that once students realise they will be given exercises based on their errors, they will be more likely to proof-read their work effectively. One will have to guard against avoidance though.
3. Encourage rewriting	Not system dependent	Many students will still only write and submit one essay – the final one. To create a "writing process approach" will depend on the lecturer or the institution.
4. Encourage exploration of language	Does not comply	The danger exists that students will revert to avoidance if they see that they get revision exercises on their errors. Although I am aware of this shortcoming, I have not been able to come up with a solution to build into the system.
5. Distinguish between grave and less important errors	To be added as system assisted, but not system dependent	The system will be adapted (once sufficient research has been done) to encourage lecturers to not only focus on errors in the surface structure.
6. See the text as a process and not a final product	Not system dependent	See the comment at 14.
7. Encourage communica-tion between the learner and the teacher	Does not comply	At present, the system does not make provision for the learner to comment on the lecturers' comments. At the North-West University (Potchefstroom Campus), there are other facilities available for two-way communication, so it was not considered necessary. In a later stage, the system may be adapted to facilitate writing-related communication.
8. Should involve students in their own learning	System-assisted, but not system dependent	The system can provide revision and consciousness-raising, so it does assist the student in seeing where they have problems. However, learning processes are complex and can therefore not be required of the system yet.
9. Raise student awareness of language features	Complies	It is hoped that the summary of errors given to each student after his/her essay has been marked, will assist in consciousness raising. This will be dependent on the effective marking of essays by the lecturers.

10. Sometimes focus only on specific errors (minimal marking), without ignoring the others.	Complies	The system will offer the opportunity to display only errors in a certain domain, thereby mimicking minimal marking. Lecturers should still mark all the errors they pick up, in order to assist in longitudinal monitoring of the students.

CRITERIA **Feedback should NOT:**	**Marking system performance**	**Comment**
11. Be a complete correction	Complies	See comment number 1 above.
12. Be "nasty" to students	Complies	The feedback tags seen by the students tell the student what is wrong in a neutral way, or gives advice and hints. Sarcastic comments were not included, but unfortunately nothing stops the lecturers from adding or creating their own derogative comments.
13. Miss any errors	Not system dependent.	The system is not an automatic marking system, but depends on accurate marking done by lecturers. The system will try to assist by prompting lecturers to not focus on surface errors only. The effectiveness of this will have to be tested.
14. Only look at errors, but also at aspects of the text which could be better	Complies, but not system dependent	The tag set provides tags for these situations, but the human marker will be the one who has to insert the tag.
15. Look only at surface errors	Complies, but not system dependent	The tag set provides tags for these situations, but the human marker will be the one who has to insert the tag. The system will prompt for further feedback, but cannot force the lecturer to do so.
16. Be all over the text, making the feedback and the text illegible	Complies	The colour-based feedback mentioned explained above, combined with the cursor-point pop-up feedback will result in a much neater marked essay than one can expect for pen and paper marked essays.
17. Waste time on the part of the lecturer	To be tested	The system at present should not require more time to mark an essay than on paper. Further development (inserting a custom-built grammar and spell checker) may even speed up the process. Wible et al. (2000) found that their system sped up the marking process, so it is reasonable to expect the same to happen here.

5.4 Conclusion

Based on research of what currently constitutes good feedback practice, the proposed system fares very well on paper. The system has been designed to assist both lecturers and students, but I expect that more refinements will be necessary after the complete system has been tested. The tag set al.so fits into a general theory of language, proving its validity and can therefore be regarded as justifiable from a theoretical perspective.

The provisional tag set has been tested, and performs well in practice as well. The complete system will be tested over a much longer period. The next chapter, which describes this testing of the tag set, will serve as proof of the effectiveness of the tag set and serve as impetus for the further refinement of the tag set.

6. EXPERIMENT: TESTING THE EFFECTIVENESS OF MARKING STUDENT ESSAYS WITH THE TAG SET.

In order to compare the proposed method of marking with two other techniques, an experiment was conducted. Theoretically, the tag set and marking technique answer to almost all the criteria for effective feedback as summarized in section 2.3, but concrete evidence is necessary to indicate if the theory works as well in practice as it does on paper. An experiment such as this would also indicate possible shortcomings. This chapter reports on the experiment. The structure of the test has already been explained in chapter three. This chapter will focus on the results of the test.

The initial findings prove that the tag set is more effective than conventional marking, but also that there are plenty to be done to get students up to standard with regards to the organization of a text. Students in this experiment did not seem to be able to identify errors accurately and consistently. It is therefore necessary to identify errors for them. Once errors were identified, students were able to correct them, but the tag set proved to be more effective in this regard than conventional marking techniques.

The results of the experiment will be discussed in terms of the error types, the superordinate categories, the error domains and the error labels. This organisation will present the data from the more overarching level, right down to the level of the individual error.

6.1 Error type distribution

All the "errors" marked in the test and all the categories of feedback do not specifically refer to things that are wrong. Feedback is intended to point out those wrong uses of language or incorrect facts[5], but if that was all, then the common grammar checker could identify half of the errors in student writing. Feedback is much more than simply the identification of incorrectness. Therefore, when creating the tag set, a distinction was made between errors of the following types:

- Inaptness (Could be better)
- Superfluous
- Incorrect
- Omission

(These categories were explained in section 4.3.)

One reason for this distinction is to test the suspicion that students have been conditioned to believe that "revision" refers to "correcting". Students seem to think that once a text is on the paper, (or computer screen) it is done and all that should be revised is the parts that are formally wrong, such as spelling errors. The reorganisation and rewriting of sentences to make the text overall better (but not necessarily more "correct") do not seem to be part of students' understanding of revising an essay. This suspicion was confirmed by the observation that students often simply exchanged a word with a synonym when told to provide *support* for an argument. The data also confirmed it. The error type distributions for the different error types are presented in the tables below (6.1 – 6.4).

[5] A lecturer pointing out incorrect facts in a literary essay is also providing feedback.

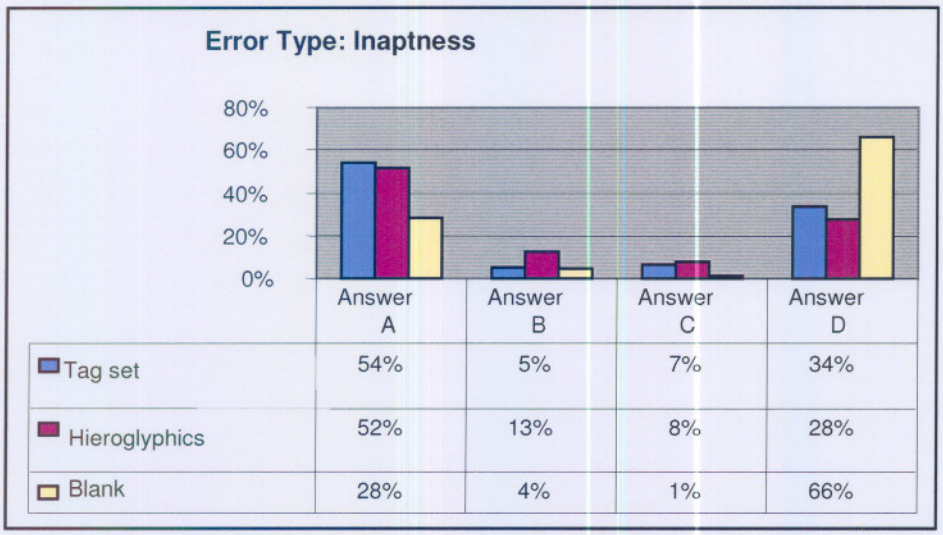


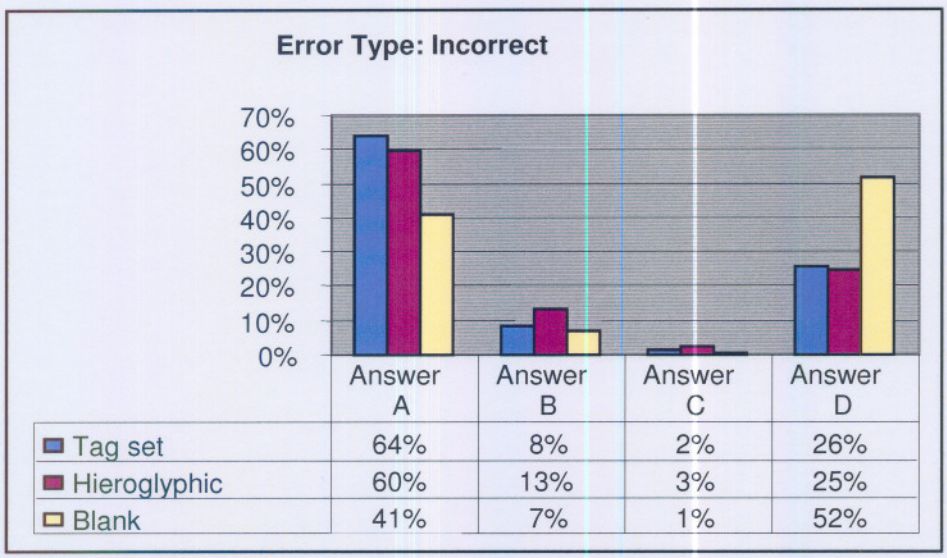
Table 6.1: Error Type Distribution for Error Type: Inaptness

Errors of the type "inaptness" refer to errors where the aim of the feedback was not to point out any explicit errors. Rather, these "error" markings were intended to lead the learners towards a higher level of language use. A classic example is where the learner gets a prompt to find a better word for the context, instead of overusing one specific word. The aim is therefore not correcting, but enhancing the text.

From table 6.1, one can see that tag set feedback in this case was slightly more effective in that students did more revisions "correctly". Their failed attempts are also fewer than for the hieroglyphic test. However, they did leave out more revisions than the hieroglyphic test subjects did. A possible reason for this may be that learners did not understand what was expected of them, or they decided against the advice from the feedback. Some of these enhancements also had to do with the organising of the text, and as will be indicated later on, these were not done to an acceptable level in any case.

It is clear that the students with the Blank tests either did not pick up where they could improve the text, or they did not see this as part of the exercise. Either way, both of these possibilities are worrisome. If students see only "error fixing" as the aim of revisions, there are a lot of psychological barriers to overcome before they can truly become better writers. On the other hand, if they did not spot opportunities to better a text, it could point to poor reading abilities (compare Pretorius, 2002) or a weak grasp of style.

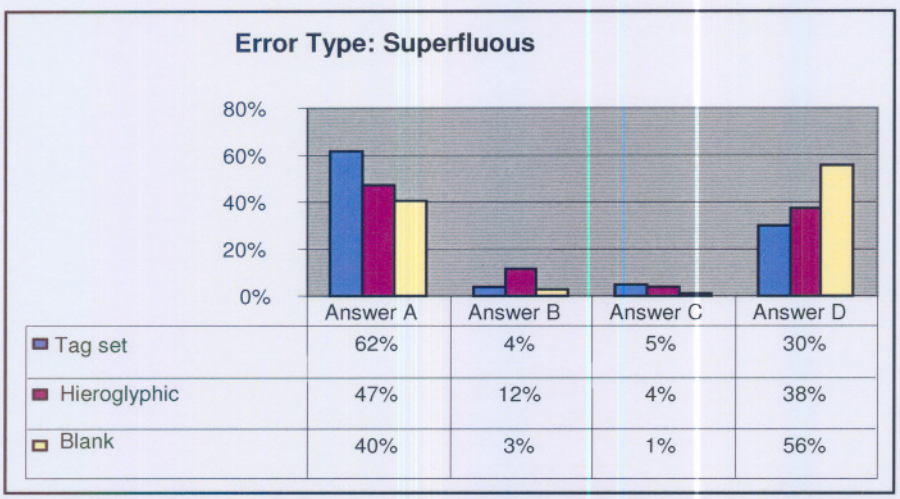
Granted though, when given the test, the students were told to find the "errors" in the essay. If we simply told them to revise, it could have been a different picture, but that will have to be researched separately to be certain. The effect of one command is debatable and the fact that the results of the tag set and hieroglyphics are very similar here, and the Blank test much worse, makes that argument weak. It seems much more likely that students simply look for "wrong" versus "right".



	Answer A	Answer B	Answer C	Answer D
Tag set	64%	8%	2%	26%
Hieroglyphic	60%	13%	3%	25%
Blank	41%	7%	1%	52%

Table 6.2: Error Type Distribution for Error Type: Incorrect

The error type "Incorrect" refers to errors that were obviously incorrect. An example may be a common spelling mistake or capitalization or punctuation error. Once again, the results are similar to those in table 6.1. The students did more correct revisions and had fewer failed attempts at correction in the tag set, than with the other two tests. Students with the tag set test also left out marginally more revisions than was the case for the hieroglyphic test. Once again, it is worrisome to see students' inability to find incorrect language use on their own (reflected by the large number of D-type answers in the Blank test). It is apparent that students would much rather leave out a revision than attempt it. One would prefer to see more B-type answers (or even C-type answers) than simply D-type answers. This is troublesome, since language learning and writing is an exploratory process and without attempting something, the opportunity for learning is wasted.



	Answer A	Answer B	Answer C	Answer D
Tag set	62%	4%	5%	30%
Hieroglyphic	47%	12%	4%	38%
Blank	40%	3%	1%	56%

Table 6.3: Error Type Distribution for Error Type: Superfluous

The error type "superfluous" (table 6.3) refers to all instances where the errors are caused by the over inclusion of words, punctuation or other grammatical information.

In this case, the tag set did much better than the other two tests, with 15% more A-type answers and 8% fewer D-type answers than the Hieroglyphics. The tag set al.so had fewer B-type answers, illustrating more accurate revisions.

This is good news for the tag set, but is disquieting in itself, since it seems to indicate that learners are not able to distinguish by themselves when and where they overuse constructions, such as synonyms. This points to a lack of reading and comprehension skills, as well as a weak lexicon. This is not breaking news, since the poor quality of reading and comprehension skills is frequently reported by researchers (see Troya and Graham, 2003:77 and Pretorius, 2002). Students do not seem to know the value of the correct word for a specific situation, or do not seem worried enough about it to take the trouble to find a better word. This could point to a limited vocabulary or a lack of skills to attempt such corrections.

This lack of distinction in word choice may not be noticed as a big problem while reading and marking texts. However, cumulatively it is a problem which needs to be addressed. The counting function of a computer will be able to find these patterns of repeating errors, making it easier to address the problem. It is also a possibility to automatically identify overused words or constructions to focus the markers' attention on it.

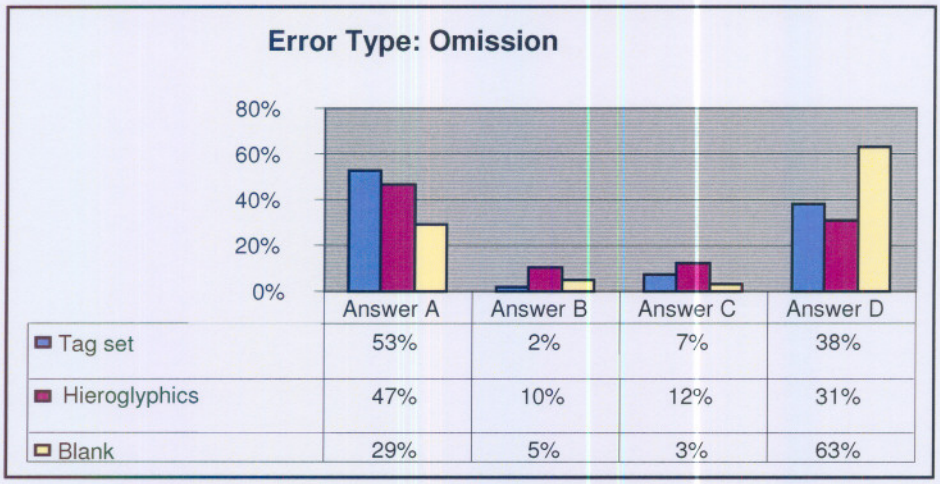


Table 6.4: Error Type Distribution for Error Type: Omission

"Omission" refers to any instance where a learner omitted something from the text. This may be a punctuation mark or a word. It could also refer to supporting facts for an argument. Once again, the tag set outperforms the other two tests, but loses out to the Hieroglyphics test on D-type answers. The D-type answers may be caused by the large number of D-type answers in instances where the revision was not simply a surface element of the text, and was therefore left out by the learners.

To test the suspicion that students prefer "correcting" over "revising", the following three tables were drawn to see where students managed to produce more A-type answers. From the tables, it is clear that no matter what marking system used, students managed more A-type answers for straightforward incorrect errors, than for errors where they had to try and make the text better.

TAG SET	*A-Type Answers*	*B-Type Answers*	*C-Type Answers*	*D-Type Answers*
inaptness	*54%*	*5%*	*7%*	*34%*
incorrect	*64%*	*8%*	*2%*	*26%*
superfluous	*62%*	*8%*	*30%*	*30%*
omission	*53%*	*4%*	*7%*	*38%*

Table 6.5: Error Type Distribution for Tag set marking technique

HIEROGLYPHICS	*A-Type Answers*	*B-Type Answers*	*C-Type Answers*	*D-Type Answers*
inaptness	*52%*	*13%*	*8%*	*28%*
incorrect	*60%*	*13%*	*3%*	*25%*
superfluous	*47%*	*12%*	*4%*	*38%*
omission	*47%*	*10%*	*12%*	*31%*

Table 6.6: Error Type Distribution for Hieroglyphics marking technique

BLANKS	A-Type Answers	B-Type Answers	C-Type Answers	D-Type Answers
inaptness	28%	4%	1%	66%
incorrect	41%	7%	1%	52%
superfluous	40%	3%	1%	56%
omission	29%	5%	3%	63%

Table 6.7: Error Type Distribution for Blanks marking technique

6.1.1 Conclusion: Error Types

An analysis of the error types leads to the conclusion that students manage to do revisions better with the tag set than with the other two tests, and their attempts at revision are more accurate using the tag set than either of the other two. Unfortunately, students seem to take revision as simply a surface element task and seldom look beyond grammatical functions. They prefer to correct things that are "definitely wrong" and do not try to make the text better. This could be linked to the findings in chapter four, where I indicated that the markers seem to prefer focusing on the surface elements. Students may therefore simply be more used to revising only the surface elements and may not be sensitive for other revisions.

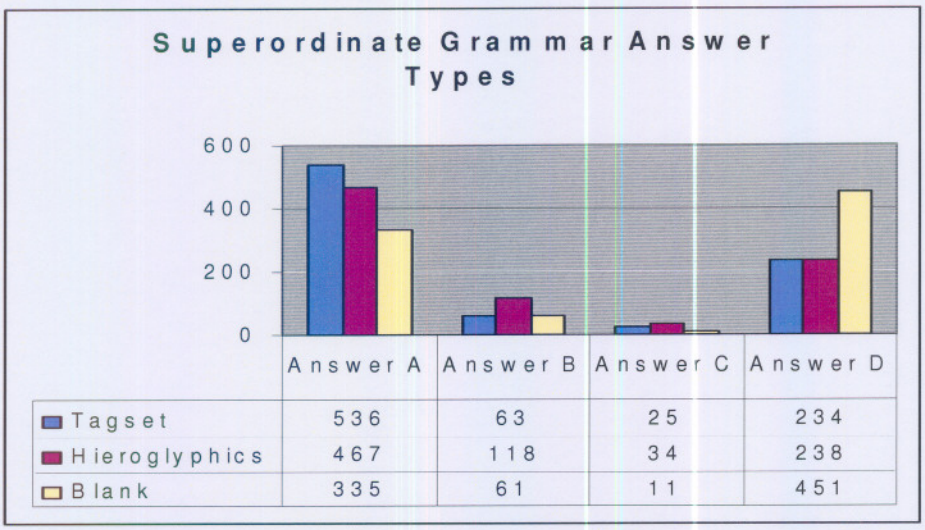
The structure of the test could also have influenced these results, as well as the style of teaching, the teaching technique and the teaching context. In addition, students were not working on their own writing. It could be that they saw this as just a silly class exercise to finish as quickly as possible. It is therefore possible that students will pay more attention to the content of the text if it is their own. This can only be tested once the marking system is up and running, since creating individualized tests for every student in the test group would be an enormous task.

6.2 Discussion of the Superordinate Answer Types

The tag set deals with mainly three superordinate types of errors: Grammar, Discourse and Presentation. An analysis of these three gives an overall picture of the performance of the marking techniques. The numbers in the following graphs were not reworked to

percentage in order to show the huge difference between the number of marked errors in the three different superordinates. The percentages are presented in table form.

6.2.1 Superordinate Grammar Answer Types



Graph 6.8 Superordinate Grammar Answer Types

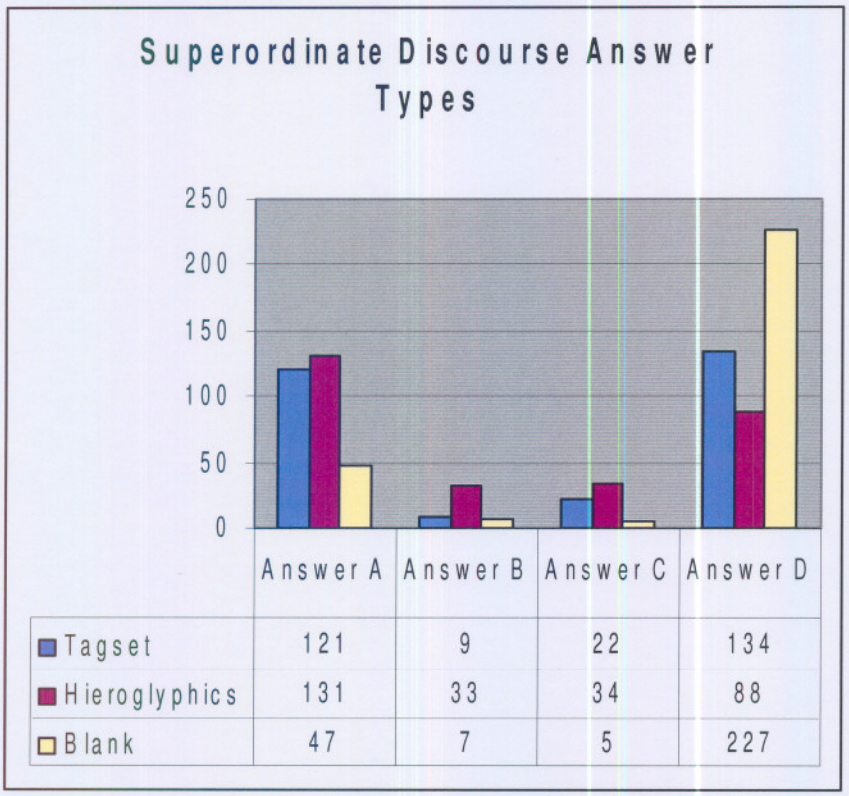
	Tagset	Hieroglyphics	Blank
Answer A	62%	54%	39%
Answer B	7%	14%	7%
Answer C	3%	4%	1%
Answer D	27%	28%	53%

Table 6.9 Percentages for Superordinate Grammar Answer Types

From the above graph and table, it is clear that the tag set resulted in more A-type answers than the Hieroglyphics and Blank tests for the superordinate grammar. The tag set resulted in fewer B-type and C-type answers than either of the others. There is also fewer D-type answers, indicating that students left out fewer revisions, suggesting that

they knew what to do, in more cases than with either of the other two tests. This indicates that the tag set is more effective as far as grammar is concerned, than the other two methods of marking. Students simply got more revisions right, without taking many erroneous guesses.

6.2.2 Superordinate Discourse Answer Types



Graph 6.10 Superordinate Discourse Answer Types

	Tagset	Hieroglyphics	Blank
Answer A	42%	46%	16%
Answer B	3%	12%	2%
Answer C	8%	12%	2%
Answer D	47%	31%	79%

Table 6.11 Percentages for Superordinate Discourse Answer Types

As can be seen here, there were fewer instances of discourse errors marked in the test. The Hieroglyphics test fared slightly better with A-type answers in the superordinate of discourse. A possible reason for this may be that it is easier to show graphically with a pencil mark that two paragraphs should be linked, than it is to describe it with words. The fact that so many students seemed to understand the single word "harsh" better than the description of "grey areas" (see test question number 45) could account for most of the differences. The overall drop in A-type answers across the board confirms suspicion about a focus on surface corrections.

The number of B-type and C-type answers for the tag set is positive. It indicates that students mostly knew how to do a revision if it was pointed out to them, whereas the students with the Hieroglyphics did not manage to revise that accurately. It is interesting to note the difference in the number of D-type answers. The tag set resulted in much more ignored advice than the Hieroglyphics. A possible reason for this may be that the tag set was designed more with revision in mind than hieroglyphic marks might suggest. For example, a hieroglyphic may simply state "passive voice", seeming to imply that the passive voice is wrong, whereas the tag set would prompt the student to "carefully consider whether the active voice would not work better for the specific

context". It is debatable whether the tag set should be adapted, based on these findings. Students are entitled to disagree with the comments made by the lecturers.

The number of D-type answers is, however, still disturbing especially since (as could be expected) the Blank test once again resulted in the largest number of D-type answers.

6.2.3 Superordinate Presentation Answer Types Distribution

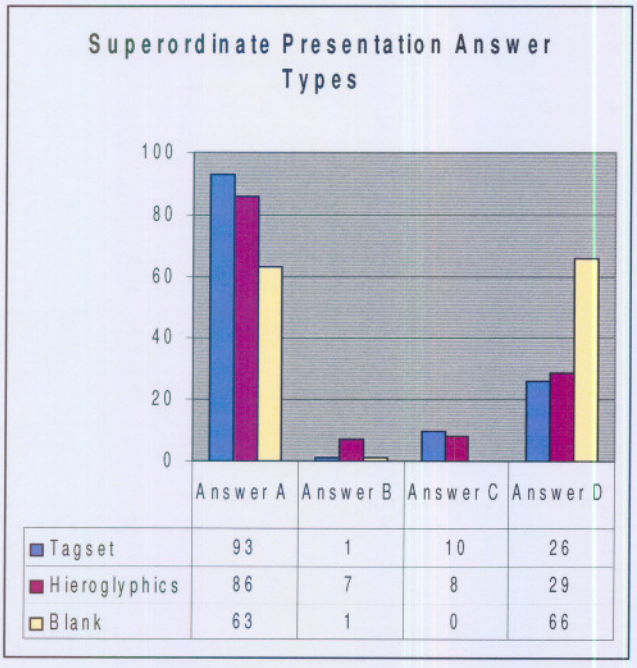


Table 6.12 Superordinate Presentation Answer Types

	Tagset	Hieroglyphics	Blank
Answer A	72%	66%	48%
Answer B	1%	5%	1%
Answer C	8%	6%	0%
Answer D	20%	22%	51%

Table 6.13 Percentages for Superordinate Presentation Answer Types

The superordinate of Presentation saw the tag set performing well once again. The Hieroglyphics is not far behind and even the Blank test did better than expected. Unfortunately, these results refer mostly to issues of capitalization and punctuation, which are still surface level skills. Note, however, that the superordinate presentation refers to the most obvious surface level. It also had the highest percentage A-type answers. This shows that students are very adapt at correcting surface level mistakes.

To summarize the findings on superordinate level: The tag set is more effective for the revision of grammar, resulting in fewer erroneous guesses from students. The tag set was effective for the superordinate discourse, but it seems that students sometimes did not agree with the comments and therefore ignored them. The surface level revisions of the superordinate presentation seem to be no problem for students, but students did fare better using the tag set.

6.3 Discussion of Domain Performance

To take a closer look at where students had trouble overall with utilizing the information from the feedback, an analysis of the performance of the different domains of the tag set could be useful. The following table gives a summary of the results, with the exceptional areas highlighted.

Respondent	Answer	Domain Lexis	Domain Style/register	Domain Syntax	Domain Factual correctness	Domain Capitalization	Domain Cohesion	Domain Punctuation error	Domain Coherence	Domain Layout	Domain Morphology
Tag set	a	65%	88%	60%	37%	88%	15%	88%	15%	4%	77%
Hierogl	a	58%	83%	50%	44%	83%	12%	73%	19%	19%	73%
Blank	a	42%	29%	36%	16%	81%	4%	40%	8%	0%	46%
Tag set	b	9%	4%	6%	2%	2%	12%	0%	4%	0%	4%
Hierogl	b	16%	4%	12%	15%	0%	4%	10%	8%	8%	0%
Blank	b	9%	2%	6%	3%	2%	0%	0%	0%	0%	0%
Tag set	c	1%	2%	5%	10%	2%	0%	0%	8%	35%	0%
Hierogl	c	2%	4%	6%	15%	0%	19%	4%	0%	23%	0%
Blank	c	1%	0%	2%	2%	0%	0%	0%	4%	0%	0%
Tag set	d	25%	6%	30%	51%	8%	73%	12%	73%	62%	19%
Hierogl	d	24%	10%	31%	26%	17%	65%	13%	73%	50%	27%
Blank	d	49%	69%	56%	79%	17%	96%	60%	88%	100%	54%

Table 6.14 Domain performance

It is obvious from the table that the tag set performs consistently better on all counts except for "factual correctness", "coherence" and "layout". However, independent of the marking technique used, students performed weak overall in three specific categories: "Factual correctness", "Cohesion" and "Coherence". These three are all related to each other in that they have to do with the organization and content of the text. Students simply left out up to 73% of these revisions. The reasons for the poor performance may be:

- Students did not understand the markings on the test.
- Students are not used to revise content, but prefer to look at surface elements.
- Students are used to correct texts, but do not attempt to improve the texts.
- Students have trouble reading for content (compare Pretorius 2002).
- Students did not want to revise the content during the test because it takes a lot of time.
- Students have trouble linking and developing a text coherently and cohesively.
- Students are completely unaware of cohesive devices – what they are or how to use them.

6.4 Answer types by labels

Springing from the discussion of the Domain performance, a closer analysis of the problem areas is needed. The following table therefore gives the answer types for the different tests for all the labels.

Respondent	Tag set	Hierogl	Blank	Tag set	Hierogl	Blank	Tag set	Hierogl	Blank	Tag set	Hierogl	Blank
Answer	A	A	A	B	B	B	C	C	C	D	D	D
Style: register wrong	76.9%	68.6%	73.1%	7.7%	19.6%	1.9%	0.0%	0.0%	1.9%	15.4%	11.8%	23.1%
Word: word form wrong	57.7%	48.1%	40.4%	3.8%	13.5%	3.8%	3.8%	3.8%	1.9%	34.6%	34.6%	53.8%
Bias	88.5%	88.5%	34.6%	3.8%	3.8%	0.0%	0.0%	3.8%	0.0%	7.7%	3.8%	65.4%
Word: word form wrong - adjective	69.2%	76.9%	57.7%	0.0%	3.8%	3.8%	0.0%	0.0%	0.0%	30.8%	19.2%	38.5%
Word: Overinclusion general	61.5%	51.3%	52.6%	5.1%	10.3%	3.8%	5.1%	3.8%	1.3%	28.2%	34.6%	42.3%
Word: Omission for punctuation	73.1%	50.0%	19.2%	0.0%	9.6%	3.8%	0.0%	7.7%	1.9%	26.9%	32.7%	75.0%
vagueness: unspecific	76.9%	67.3%	34.6%	0.0%	11.5%	7.7%	1.9%	5.8%	1.9%	21.2%	15.4%	55.8%
Word: Better word	58.5%	46.2%	33.8%	13.1%	19.2%	15.4%	0.0%	3.8%	0.8%	28.5%	30.8%	50.0%
Style: Inconsistent	63.5%	63.5%	32.7%	11.5%	11.5%	7.7%	0.0%	7.7%	0.0%	25.0%	17.3%	59.6%
Style: Contractions	88.5%	76.9%	23.1%	3.8%	3.8%	3.8%	3.8%	3.8%	0.0%	3.8%	15.4%	73.1%
Fact: unsupported argument	25.6%	30.8%	9.0%	1.3%	19.2%	2.6%	12.8%	24.4%	2.6%	60.3%	25.6%	85.9%
Capitalization	88.5%	82.7%	80.8%	1.9%	0.0%	1.9%	1.9%	0.0%	0.0%	7.7%	17.3%	17.3%
Word: omission verb	80.8%	84.6%	84.6%	3.8%	0.0%	0.0%	3.8%	7.7%	0.0%	11.5%	7.7%	15.4%
Spelling/ typing error	76.9%	73.1%	26.9%	0.0%	10.3%	0.0%	1.3%	0.0%	0.0%	21.8%	16.7%	73.1%
sentence: construction/word order	50.0%	46.2%	40.4%	17.3%	17.3%	5.8%	7.7%	5.8%	1.9%	25.0%	30.8%	51.9%
Paragraphing: Relate or Move	15.4%	11.5%	3.8%	11.5%	3.8%	0.0%	0.0%	19.2%	0.0%	73.1%	65.4%	96.2%
Word: Omission general	78.8%	63.5%	50.0%	1.9%	5.8%	9.6%	3.8%	1.9%	5.8%	15.4%	28.8%	34.6%
Punctuation: missing	88.5%	73.1%	40.4%	0.0%	9.6%	0.0%	0.0%	3.8%	0.0%	11.5%	13.5%	59.6%
Sentence: run-on	65.4%	53.8%	26.9%	3.8%	9.6%	1.9%	5.8%	1.9%	0.0%	25.0%	34.6%	71.2%
Word: Overinclusion plural marker	61.5%	34.6%	3.8%	0.0%	15.4%	0.0%	3.8%	3.8%	0.0%	34.6%	46.2%	96.2%
Facts: wrong	3.8%	42.3%	11.5%	3.8%	15.4%	0.0%	0.0%	3.8%	0.0%	92.3%	38.5%	88.5%
Sentence: passive voice	15.4%	15.4%	19.2%	3.8%	50.0%	23.1%	7.7%	19.2%	3.8%	73.1%	15.4%	53.8%
Coherence: counterproductivity	15.4%	19.2%	7.7%	3.8%	7.7%	0.0%	7.7%	0.0%	3.8%	73.1%	73.1%	88.5%
Coherence: unfinished thought	11.5%	15.4%	11.5%	3.8%	3.8%	11.5%	15.4%	7.7%	3.8%	69.2%	73.1%	73.1%
Layout error	3.8%	19.2%	0.0%	0.0%	7.7%	0.0%	34.6%	23.1%	0.0%	61.5%	50.0%	100.0%
Spelling: Ambiguity - homophone/ homonym	30.8%	38.5%	19.2%	42.3%	38.5%	26.9%	3.8%	3.8%	0.0%	23.1%	19.2%	53.8%
Style: Tautological	23.1%	38.5%	3.8%	3.8%	11.5%	0.0%	30.8%	15.4%	3.8%	42.3%	34.6%	92.3%
Concord	76.9%	73.1%	46.2%	3.8%	0.0%	0.0%	0.0%	0.0%	0.0%	19.2%	26.9%	53.8%
Collocation or fixed expression error	76.9%	69.2%	73.1%	11.5%	11.5%	11.5%	0.0%	0.0%	0.0%	11.5%	19.2%	15.4%

Table 6.15 Label answer type

The analysis of the domains pointed out that the biggest problems are with coherence, cohesion and factual correctness. A closer analysis of the above table points more specifically to:

- Wrong word forms and the inability to distinguish between homophones
- The inability to give a better word (weak lexicon)
- Inconsistent style

- Problems with sentence construction, and the inability to correct and identify run-on sentences
- Unsupported and unfinished arguments or wrong facts and the inability to support or finish those arguments.
- Problems with identifying tautological and counterproductive statements.

The objection can be raised that students did not know enough about the essays in the test to be realistically expected to be able to notice and revise problems with style, unsupported arguments or wrong facts and tautological statements. These may be considered subjective areas, and I do not have a way to measure them. However, unsupported arguments, style and tautological statements are obvious enough that I strongly believe the students were given enough context to be able to identify and correct the problems.

The fact that students have trouble with these errors in both the Hieroglyphic and Tag set tests (not to mention the Blank test) leaves us with three possibilities:

- Students know the concepts, but have trouble overall in correcting them.
- Students do not know the concepts.
- Students have problems in their reading and comprehension skills, resulting in the inability to create cohesion and understand the flow of an argument (compare Pretorius, 2002:187,191; Van Wyk, 2002).

It seems that in order to address these problems, it should be necessary to give the students even more assistance. It is debatable whether this assistance should be the responsibility of the tag set or computer system. It is preferable to have separate classes

on the topics of coherence and cohesion, sentence construction, and argumentation with the specific aim of preparing students on how to use it in their writing.

6.4.1 Label effectiveness

To summarize the effectiveness of the labels, the following table (6.20) gives the A-type answers for the different labels in the Hieroglyphic and Tag set tests. (The Blank test did not perform well at all, so a short individual discussion will follow later.) In the table, the yellow fill highlights labels where the Hieroglyphics did better than the tag set, whereas the green fill highlights where the tag set did better than the Hieroglyphics. (The red fill indicates problem areas in the table below.)

	Tag set	Hieroglyphics
Style: register wrong	76.9%	68.6%
Word: word form wrong	57.7%	48.1%
Bias	88.5%	88.5%
Word: word form wrong - adjective	69.2%	76.9%
Word: Overinclusion general	61.5%	51.3%
Word: Omission for punctuation	73.1%	50.0%
vagueness: unspecific	76.9%	67.3%
Word: Better word	58.5%	46.2%
Style: Inconsistent	63.5%	63.5%
Style: Contractions	88.5%	76.9%
Facts: unsupported argument	25.6%	30.8%
Capitalization	88.5%	82.7%
Word: omission verb	80.8%	84.6%
Spelling/typing error	76.9%	73.1%
sentence: construction/word order	50.0%	46.2%
Paragraphing: Relate or Move	15.4%	11.5%
Word: Omission general	78.8%	63.5%
Punctuation: missing	88.5%	73.1%
Sentence: run-on	65.4%	53.8%
Word: Overinclusion plural marker	61.5%	34.6%
Facts: wrong	3.8%	42.3%
Sentence: passive voice	15.4%	15.4%
Coherence: counterproductivity	15.4%	19.2%
Coherence: unfinished thought	11.5%	15.4%
Layout error	3.8%	19.2%
Spelling: Ambiguity - homophone/homonym	30.8%	38.5%
Style: Tautological	23.1%	38.5%
Concord	76.9%	73.1%
Collocation or fixed expression error	76.9%	69.2%

Table 6.16 Label effectiveness

I find that the tag set is more effective than the Hieroglyphics for most grammatical categories, but if the red flagged labels are left out of consideration, the tag set outperforms the Hieroglyphics on all but one count. The labels marked in red, are where students did not manage to correct errors. These all have A-type answers in fewer than 50% of all cases. Marking these can be considered a waste of the markers' time. The inability of students to correctly revise errors of these labels can not be attributed to the marking technique, since no marking technique had any success. Students should be given additional instruction on these first, before any specific marking technique can be tested effectively.

A closer look at the red flagged labels indicates that most of them deal with the categories of context or semantic content as identified by Givón – either organization or the statements made in the text. As stated already, students have trouble linking paragraphs coherently, revising incomplete arguments or spotting irrelevant/counterproductive statements. Another possibility is that the test was simply not a very effective way of testing restructuring. Some students did rewrite large parts of the test to indicate how they would restructure, but most did not go through all that trouble. It is also difficult (and presumably seems silly to students) to think up a supporting fact or reference out of the blue.

The two exceptions are "Layout error" and the homophone/homonym spelling errors. The layout error was a double space, so it was difficult to spot on a typed test. If it was a more visible error, students would probably have seen it much better. With regards to the spelling error, students seem to have a problem with words that sound

alike. Most of the students did not have dictionaries with them when doing the revisions, so they were not able to look up the confusing words. The raw data for this specific error suggests that students did try to fix these errors, but did not know how to. It is obvious therefore that the tag set and Hieroglyphics are transparent enough, but a greater awareness of how to correct these errors should be taught.

It is interesting to note that the Hieroglyphics are more effective than the tag set when the adjective form of the word is used incorrectly. (The Hieroglyphics scored 76.9%, whereas the tag set only scored 69.2%.) A possible reason for this is that the tag set specifically tells the student that the adjective is wrong, and it could be that all the students are not sure what "adjective" means (sadly). They are, however, familiar enough with how the language should sound (i.e. their inherent competence is developed far enough), to deduct the right word form if they are not worried about what "adjective" means. This is simply speculation though and it should be researched to see if their is any truth in the speculation.

The Hieroglyphics are also more effective when pointing out errors of coherence and incorrect facts (although they are still not effective enough to count as an effective means of feedback). Simply judging from their, it seems that students are more at ease with sarcastic notes scribbled on the page, (such as "Really? Have you been everywhere?") or short statements or questions about the text (such as "harsh" or "Is this true?"). If one simply looks at the numbers, one can therefore argue that all marking comments should be in question form to make the students think about the writing. However, this type of questions cannot be standardised, so it should be inserted with the comment function in collaboration with the standardised tag. (The

marking system is to have a standardised feedback tag, as well as the opportunity for the lecturer to insert his/her own comments. As seen in the tag set, some tags **specifically advise** the marker to insert his/her own comments. The tag, "This piece of writing is well-structured", with the XML tag "POS" is an example of this.)

An interesting problem lies in the "correction" of the passive voice. The tag set says "Carefully CONSIDER whether the active voice would not work better." It could be that the students considered it and thought the sentence was better in the passive voice. Students do not seem to experience passive voice sentences as a problem. Once again, the choice between active and passive voice lies deeper than simply the surface structure of the text. Passive and active voice may be used to emphasize certain statements and depending on the topic of the essay, it could be a conscious choice to use the one or the other. If students did not revise it, it could simply mean that they did not agree with the marker. On the other hand, it is always possible that students are not aware what the difference between active and passive voice is, or how to make a choice for one or the other.

With regards to the tautological statement[6] made by the one writer, it seems that students are unfamiliar with the term "grey area". The Hieroglyphics simply had the word "harsh" written next to the statement, and that seemed to be more effective feedback. The tag set was therefore changed to include the word "harsh" since it seemed to be clearer to the students. It is possible that by monitoring the tag set over time, some more refinements will be possible. The fact remains that students had trouble with the concept of a tautological statement. This could be due once again to

[6] "If a game is not won, they don't get paid" on page 5 of the tests.

the test format. On the other hand, it could be due to an inability to read and interpret the text, which is much more serious a problem. One way to test if reading is in fact the problem, is to monitor the effectiveness of the tag when students revise their own writing. If the tag is effective on their own writing, then reading may be the problem (assuming that they know exactly what their own statement was intended to mean).

6.4.2 Implications of the effectiveness of the Blank test

The Blank test proved that students need some support in identifying errors and doing revision. Students were only able to identify capitalization errors on their own, as well as picking up missing verbs. They were also to a limited extent able to identify a register problem and an incorrect fixed expression. ALL of the rest of the errors were either missed, ignored or misunderstood. This could be due to an inability by students to identify problems in writing, or simply a reluctance to try.

6.4.3 Conclusion on label effectiveness

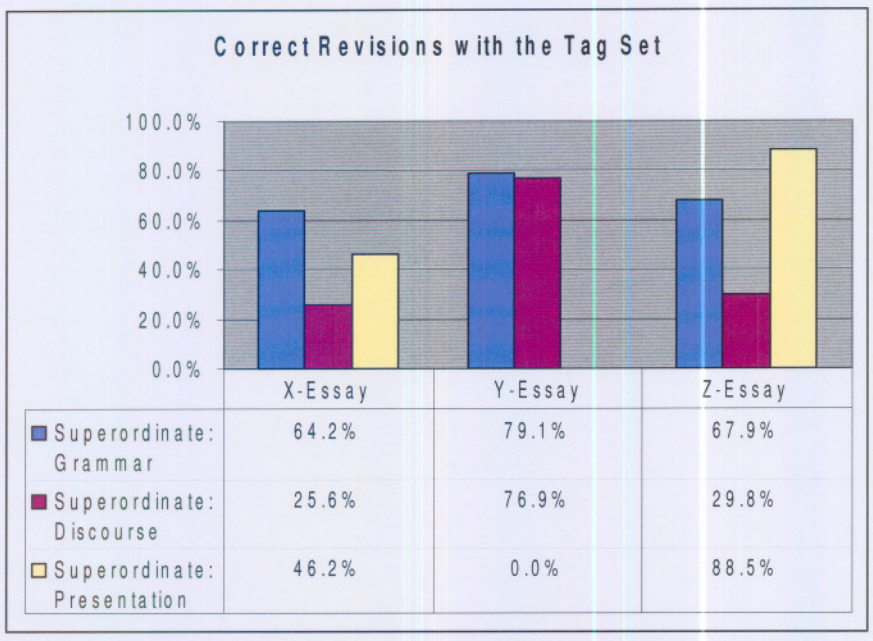
The labels for the tag set will still be refined during the testing phase of the project. However, the experiment proved that it is better already than the normal method of marking and much better than the "no marking" technique. Spencer (1998:120;163) similarly found no concrete evidence that minimal marking (Blank marking) resulted in improved proficiency over time.

6.5 Do students fare better correcting poor essays?

While working through the data, it was noticed that students seemed to be more capable in their revisions on poor texts than with the slightly better ones. I have therefore labelled the three essays used as X, Y and Z (in declining quality) and investigated the success of the revisions. In this section, a breakdown of the data will be presented.

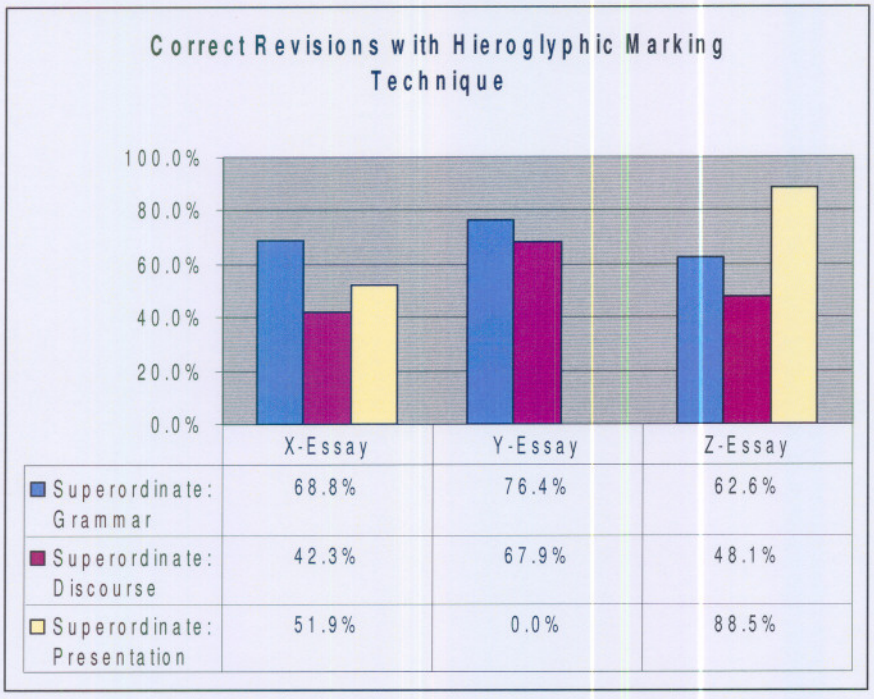
6.5.1 Correct Vs Incorrect Revisions In Declining Quality Of Essays

To get a bird's eye view of the effectiveness of revisions between the Y, X and Z essays, the data was simply broken down into correct versus incorrect revision. This means that if a student attempted the right kind of revision, the revision was counted as correct. Therefore, the following graphs indicate the combined effects of the A- and the B-type answers.



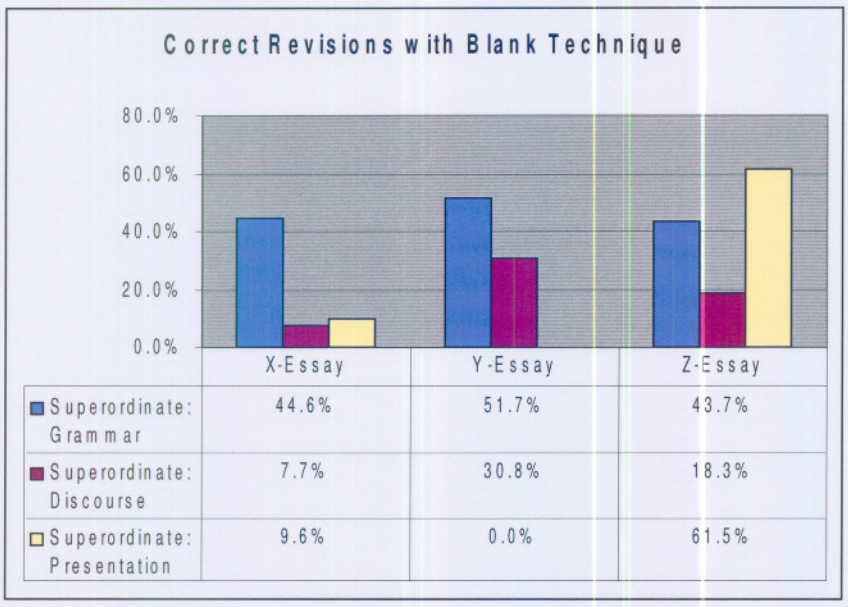
	X-Essay	Y-Essay	Z-Essay
Superordinate: Grammar	64.2%	79.1%	67.9%
Superordinate: Discourse	25.6%	76.9%	29.8%
Superordinate: Presentation	46.2%	0.0%	88.5%

Table 6.17 Correct Revisions with Tag Set



	X-Essay	Y-Essay	Z-Essay
Superordinate: Grammar	68.8%	76.4%	62.6%
Superordinate: Discourse	42.3%	67.9%	48.1%
Superordinate: Presentation	51.9%	0.0%	88.5%

Table 6.18 Correct Revisions with Hieroglyphic Marking Technique



	X-Essay	Y-Essay	Z-Essay
Superordinate: Grammar	44.6%	51.7%	43.7%
Superordinate: Discourse	7.7%	30.8%	18.3%
Superordinate: Presentation	9.6%	0.0%	61.5%

Table 6.19 Correct Revisions with Blank Marking Technique

In the above three graphs, one can see that the students did better overall in revising the Y-Essay. This implies that:

- Students cannot effectively revise writing which is of a moderate quality.
- Students struggle to revise writing which is very bad (just as teachers/lecturers struggle to mark it).
- Overall, students seem to be more comfortable revising Grammar than Discourse or Presentation.

The exception seems to be with the Hieroglyphic marking system, where students had marginally more success in revising the grammar of the X-essay than the Y-essay. The elements of discourse and presentation were true to the pattern. This tells us that students need more practise in revision, or possibly even training in how to do effective revision. If students have trouble revising very bad essays or essays slightly better than the norm, they are not effective proofreaders of their own work or the work of others. This is an important part of writing and should therefore be addressed.

6.6 Conclusion

During the test, students were also asked to make a short note of what they considered the most effective marking technique to be. Most of them ignored this request, but the few (seven or so) that did answer, indicated that the tag set was in their opinion a better technique than the other two. Although this is the opinion of but a few, it is an indication that students would not complain about getting more detailed feedback.

The tag set proved to be more effective than the other two marking techniques. This justifies its existence and use. The average semester mark for the two classes on which this experiment was conducted, is 52%. Using the tag set, these same students managed to correctly revise problems in a text to an accuracy of up to 88%. Combined with the further possibilities created by computer-assisted marking, this tag set and system could make writing education more effective. This is especially true if one keeps in mind that this system aims to improve something that is not very functional at present, to a level which is *more* functional. Until writing education and artificial intelligence have evolved much further, a 100% automatic and accurate computer marker is not a realistic aim, but the database generated by a consistent use of this system, could go a long way towards a better understanding and more accurate analysis of learner problems.

The findings of the experiment did not only prove that the tag set and marking technique can be effective. It also raised a few suspicions about the current state of writing and teaching. It seems that learners sometimes do not know what is expected of them – what should they do with feedback. It may also be that they decided against using the advice from the feedback, but that is not consistent with the findings in section 2.2.2.2.4. There also seems to be a problem with reading and comprehension skills (visible especially in errors to do with style and organisation) and the psychological barrier of seeing "error fixing" as the sole purpose of revision. Students seem to be more likely to leave out a revision than attempt it, and are not experienced in finding incorrect language use on their own. Students are not able to identify overuse of constructions and do not seem to be able to (or do not care to) try and enhance a text. They focus on surface elements (which they are quite capable of

correcting independently) and do not seem to know the concepts of supported arguments and cohesion. Alternatively, students do know the concepts, but do not know how to practically utilize it. Furthermore, chapter four found that lecturers focus on surface elements, so students may simply be used to looking only at that.

From the data it appeared as if the tag set and the marking technique used, work well enough to implement without major changes. Small changes were made to the final version of the tag set, but these mostly had to do with word choice. In some cases a note was added, imploring the marker to add a comment of his/her own. The tag set will of course be monitored to ensure that it answers to the needs of both the lecturers using the marking system, and the learners. The final version (in context of this study) of the tag set is presented in chapter 7.

7. FINAL VERSION OF THE TAG SET

The final version of the tag set is presented in this chapter, in the form it was incorporated into the software interface. It is still a work in progress and will be refined as testing continues throughout the next few years. The tag set has been aranged alphabetically according to the tag. The superordinate and domain is not relevant for implementation purposes, but for testing.

SUPERORDINATE	DOMAIN	TAG	TYPE	FEEDBACK	EXAMPLE/EXPLANATION	XML LABEL
Grammar	Lexis	Ambiguity: False Friend	incorrect	Make sure that this is the word you wanted to use and not one sounding the same as in your first language.	Using a word that looks like a word from their native language, but means something different. Example: lemoen (Afrikaans) = Orange (English) and not lemon.	**FMF**
Grammar	Syntax	Ambiguous word order	unclear/inaptness	Your sentence is ambiguous. Change the word order and/or punctuation.	Use this when the sentence is ambiguous due to the word order or the punctuation. Example: Men, who have poor eyesight, make poor pilots.	**SU**
Discourse	Style/register	Be impersonal	inaptness	In academic writing, try to be formal and impersonal. For example, do not use words such as "I" or "I think".	Use this when you feel that the learner should rather be impersonal. In academic writing, it is customary to be impersonal.	**S**
Discourse	Style/register	Bias	inaptness	Be careful of gender bias.	Use this when a learner makes use of language which could be considered sexist.	**SB**

Presentation	Capitalization	Capitalization	incorrect	Your use of capital letters or lower case letters is incorrect.	There are plenty of Banks in my country. I am going to visit my friend in the usa.	PC
Discourse	Coherence	Coherence: counterproductivity	inaptness	This statement is going against your main argument. If you want to show the other side of the argument, make sure you indicate it as such.	Use this when a learner makes a statement that goes against his/her main argument.	DC
Discourse	Syntax	Coherence: unfinished thought	omission	This sentence is too vague. It looks like an unfinished thought. You should complete the sentence to indicate that you have finished the thought. Add some more concrete information to support your statement.	Use this when a learner did not complete a thought, but left it hanging, for example when "etc." is used without good cause or when a statement is left without support or explanation.	DCU
Discourse	Syntax	Cohesion: repetition of statement	inaptness	You repeat basically the same statement using different words. If you see a distinction, make it clearer.	Use this when a learner writes the same thing, but tries to disguise it as a new statement. This may indicate a weak argument or maybe weak sentence structure.	DCR
Grammar	Lexis	Collocation or fixed expression error	incorrect	There is a specific way of writing this. Make sure you use the more usual way.	Use this when a learner should have used a fixed expression or standardised collocation such as "strawberries and cream", but instead used another (uncreative) distortion such as "cream and strawberries".	**WO**
Grammar	Morphology	Concord	incorrect	The form of your verb should agree with the	The verb must agree with the subject in number.	**GNN**

				subject it refers to.		
Discourse	Factual correctness	Fact: unsupported argument	omission	Support your argument with an example or a fact.	Use this when a learner leaves a statement without explanation. Example: "There are many stupid rules in the bank."	FU
Discourse	Factual correctness	Facts: Opinion	inaptness	Do not use opinions as facts, unless you indicate them as opinions.	Use this when a learner tries to pose his opinion as a solid truth, e.g. "Robbie Williams is the best singer in the world."	FO
Discourse	Factual correctness	Facts: wrong	incorrect	Make sure your facts are correct.	Use this when a learner uses the wrong facts in a piece of writing, e.g. "Charlie Brown has a dog called Peanuts."	FW
Discourse	POSITIVE COMMENTS	Good Reasoning	Good	You link the ideas very well.	Use this to give praise to the student when he/she manages to link ideas together in a novel way.	POG
Discourse	POSITIVE COMMENTS	Good structure	Good	This piece of writing is well-structured.	Use this when a paragraph is well-structured. It is advisable to add a comment explaining what you found good.	POS
Discourse	POSITIVE COMMENTS	Interesting point	Good	You have a very interesting point here. Good thinking.	Use this to give praise to the student for good insight and valid arguments.	POI
Discourse	Structure	Introduction: weak	inaptness	The introduction to the essay should be stronger. See the note added by the lecturer.	The introduction simply restates the topic without indicating what will be discussed in the essay.	DI

Presentation	Layout	Layout error	inaptness	Your layout is not right or could have been better.	Use this when a learner has to write something in a specific layout format, such as an official letter. It would be wise to add a comment explaining what the problem is.	PL
Presentation	Layout	Layout inhibits reading	inaptness	Your layout makes it difficult to read this text.	Use this when the layout of a text makes it difficult to read, like when there are no spaces between paragraphs.	PLR
Grammar	Syntax	Negation incorrect	incorrect	You should make sure how to make this sentence negative.	Sometimes our teacher not allow us.	GN
Discourse	Cohesion	Paragraph: jumbled	inaptness	This paragraph does not deal with only one thought.	Use this when a learner tries to cram too many thoughts into one paragraph. You could use this with the tag "Start new paragraph" to help the student decide where to break the paragraph.	DCPJ
Discourse	Cohesion	Paragraph: start new	inaptness	You should start with a new paragraph here, since you start with a new idea.	Use this when it is obvious that a learner should have started a new paragraph, but neglected to do so.	DCPN
Discourse	Cohesion	Paragraph: weak opening sentence.	inaptness	You could have used a stronger opening sentence for this paragraph. Find one that is more relevant to the rest of the paragraph.	Use this when a paragraph starts with a sentence that is not relevant to (or indicative of) the content of the rest of the paragraph.	DCPWO

Discourse	Cohesion	Paragraphing: Link	inaptness	This paragraph is continuing the thought of the previous one. Why don't you make it one paragraph?	Use this when a student just randomly inserted a paragraph marker, without regard for the theme of the paragraph.	DCPL
Discourse	Cohesion	Paragraphing: Relate or Move	inaptness	This paragraph does not link with the one before it. Show why you place it here and how it relates to the previous or place it somewhere else.	In some cases a paragraph is out of place and should rather be placed somewhere else. Another possibility is that a paragraph is just not linked clearly enough to the one before or after it.	DCPR
Grammar	Punctuation error	Punctuation: missing	incorrect	You need to insert a punctuation mark here.	Use this when a learner left out a comma, full stop or other punctuation mark.	**QM**
Grammar	Punctuation error	Punctuation: unnecessary	incorrect	You do not need a punctuation mark here.	Use this when a learner inserted an unnecessary punctuation mark, or placed a punctuation mark in the wrong place. This tag could be used together with "punctuation missing" if a punctuation mark is inserted in the wrong place.	**QR**
Grammar	Punctuation error	punctuation: wrong	incorrect	Your punctuation is wrong. There is another punctuation mark that will be better.	The student's use of the punctuation mark(s) is wrong. There should be punctuation, but the punctuation used is not correct . Example: They saw lions. elephants. and rhino.	**QC**
Discourse	Coherence	Reasoning incoherent	inaptness	How do you come to this conclusion with the facts you mentioned? Your argument is not structured clearly enough for me to	Use this when you cannot see a logical link between a statement and the deduction the student makes from that.	DCR

				see the logical development.		
Discourse	Factual correctness	Reference omitted/wrong	inaptness	You should indicate where you got this information from. Beware of plagiarism. Use the correct formatting to indicate the source.	Use this when a learner uses a direct quote or mentions an author, but do not indicate the source. This could also be attributed to style problems.	FR
Discourse	Cohesion	Relevance to topic	inaptness	This statement is not relevant to the topic or the connection to the topic is not clear enough.	Use this when a statement is not directly relevant to the topic. Example: A learner attempts to persuade with an emotional argument where he/she should have used cold, hard facts.	DCRT
Grammar	Syntax	Sentence cohesion	incorrect	The way you link this two sentences pr two thoughts, does not make sense or is incorrect.	Example: two sentences that have nothing to do with each other cannot be linked by "and".	YSC
Grammar	Syntax	Sentence incomplete	incorrect	This is not a complete sentence.	In the movie, "Patch Adams".	YSI
Grammar	Syntax	Sentence: run-on	inaptness	This sentence is too long. Break it up. Use punctuation.	Long sentences are difficult to read. Students should break it up into two or more coherent sentences, each with their own correct punctuation and word order.	YSR
Grammar	Syntax	sentence: construction/word order	incorrect	Your word order is incorrect or could be better. Change it.	Use this when the word order in a sentence is incorrect.	YWO
Discourse	Syntax	Sentence: active voice	inaptness	Carefully consider whether the passive voice will not work better for this sentence.	Use this where the student used the active voice, but in your opinion, the passive voice would have worked better.	YP

Discourse	Syntax	Sentence: passive voice	inaptness	Carefully consider whether the active voice will not work better for this sentence.	Use this where the student used the passive voice, but in your opinion, the active voice would have worked better.	YA
Grammar	Lexis	Spelling/typing error	incorrect	This word is spelled incorrectly.	Use this only for spelling or typing errors. "Aslo" for "also" and "beeutiful" for "beautiful".	FS
Grammar	Lexis	Spelling: Ambiguity - homophone/homonym	incorrect	Homophone/homonym error: Make sure you used the correct word or the correct spelling of the word.	Confusing two words that sound or look the same e.g. "there" and "their".	FSF
Discourse	Style	Style: Construction overuse	inaptness	You use the same construction too much. This makes your reading monotonous. Try using different words together to say what you want to say.	Use this when you realise that a learner is overusing a construction. Add a note to explain.	SCU
Discourse	Style/register	Style: Contractions	inaptness	Rather do not use contractions in academic writing.	Use this when a learner used words like "won't" or "isn't" in academic writing.	SCT
Discourse	Syntax	Style: Inconsistent	inaptness	Be consistent in your style	Use this when a learner keeps switching between "one", "you", "a person" or if you see other inconsistencies in style. Maybe add a note.	YI
Grammar	Lexis	Style: register wrong	inaptness	Use the right register. Make sure you are not too informal or too formal for the context of your writing.	Use this when a learner used an informal word in a formal piece of writing or a formal word in an informal piece of writing.	SR

Grammar	Syntax	Style: sentence vague	omission	You leave out too much information for this sentence to make sense. Be more detailed.	Use this when a student is too cryptic to make sense, e.g. "They are both accounts..." vs. "They are both savings accounts..."or "It is as if **they** are trying to suppress **their** feelings". Who is suppressing whose feelings? Also use this when a learner uses constructions like *"everyone knows" or "they say that"* or when the learner uses "it" when a more specific identifying word should have been used.	YV
Discourse	Factual correctness	Style: Tautological	inaptness	You allow no room for any "grey areas". You should consider that there are two sides to most things. Be careful of a harsh statement.	"Soldiers are always bloodthirsty killers."	DFT
Grammar	Lexis	Style: Verbosity	inaptness	You can say this in fewer words.	Using too many words to say what needs to be said.	GLV
Discourse	Factual correctness	Vagueness: unspecific	inaptness	Be more specific.	Use this when a learner refers to "they" or "them" without specifying who it is. Also use this when a learner uses "etc." instead of completing a list in a neat manner. Examples: "They say that..." "They should go to hospitals, clinics, etc." instead of "Students should go to hospitals, clinics and other medical fascilities."	DFV

Grammar	Articles	Word: article wrong	incorrect	The article should have been in a different form. Distinguish between definite and indefinite articles.	They asked her to describe a missing child. They asked her to describe **the** missing child.	**GA**
Grammar	Lexis	Word: Bad word	incorrect	This word is racist/sexist or derogative and should rather not be used.	Use this tag when a word is racist or sexist or in any way derogative and should not have been used in the context of the writing.	GLB
Grammar	Lexis	Word: Better word	incorrect	There are more appropriate words available with which you can express this idea.	Use this tag when a learner should have used a better word, like when a learner uses a near-equivalent form to the one intended (e.g. "not happy" for "sad") or when the learner describes something because he don't know the exact word for it.	GLBW
Grammar	Lexis	Word: conditional wrong	incorrect	This is not the best (or right) way to formulate a condition.	"They could probably study better **when** they get electricity" vs. "They could probably study better **if** they get electricity."	GLC
Grammar	Syntax	Word: Omission for punctuation	omission	You should rather insert a linking word here than a punctuation mark.	Use this when a learner uses punctuation instead of a word, e.g.. "She gave me peaches, nuts, apples in a basket to take home." Also use this when a learner starts a new sentence when he/she should have continued with the previous.	GYQ

Grammar	Syntax	Word: Omission general	omission	You need to insert a word here.	Example: "When you open an account they ask you **(a)** few questions." or "...only the last year or two..." vs. "...only during the last year or two..."	GYWM
Grammar	Morphology	Word: omission plural marker	incorrect	You need to insert a plural marker here.	Use this when a learner did not indicate a plural where he/she should have. E.g. "their customer" or "...they want one of your family member to come..."	GMW
Grammar	Lexis	Word: Omission possession marker	incorrect	Make sure you use the correct possession marker or mark the possession correctly.	If an apostrophe is used to indicate possession, make sure that it is placed in the right position. E.g., Students' vs. Student's	GLWMP
Grammar	Lexis	Word: wrong word preposition	incorrect	You use the wrong preposition here.	"days **in** which" vs. "days **on** which"	GLWWP
Grammar	Syntax	Word: Omission preposition	omission	You need a preposition here.	The man went (**into**) the bank.	GYWMP
Grammar	Syntax	Word: Omission tense marker	omission	Use the correct time word. You can say this in much less words then.	Use this when a learner uses a description of a tense, instead of just using the correct time-word.	GYWMT
Grammar	Syntax	Word: omission verb	omission	You need to insert a verb here.	"Only in the last year or two they begin..." vs. "Only in the last year or two **do** they begin..."	GYWMV

Grammar	Syntax	Word: Overinclusion general	superfluous	This word(s) does not need to be here. Make sure that you do not repeat yourself.	I am very utterly pleased.	GYWR
Grammar	Syntax	Word: Overinclusion plural marker	superfluous	There is no need to indicate that the word is a plural.	all the **plenty** teachers	GYWPL
Grammar	Syntax	Word: Overinclusion tense marker	superfluous	You do not have to indicate the tense here.	Use this when a student indicates a tense twice or unnecessarily.	GYWPT
Grammar	Syntax	Word: Overinclusion unnecessary pronoun	superfluous	You do not need to indicate the subject or object of the sentence twice.	The man **he** chooses the bank he likes.	GYWPP
Grammar	Lexis	Word: Preposition wrong	incorrect	This is the wrong preposition. Find out which would have been the correct one to use.	Use this when a student used the wrong preposition e.g.. It is not so easy as it looks.	GLWP
Grammar	Lexis	Word: Pronoun wrong	incorrect	You should have used a different pronoun here. Find out which one.	Use this when the learner used the wrong pronoun.	GLWPR
Grammar	Lexis	Word: Repetition	inaptness	You use the same words repeatedly. Find different words that may convey your message more clearly.	Use this tag when you realise that a student keeps on using the same word. E.g.. If a student use the word "Good" to mean "excellent" and "strong" and "hard" and "pretty" etc. This will be context sensitive. If you have to use the "Word: better word" tag a lot for the same word, rather start using the "Word: repetition" tag.	GLWR

Grammar	Lexis	Word: word choice obscuring meaning	incorrect	This word is not clear enough. Find a better word to say what you want to say.	Use this when another word would make the intended meaning much clearer. E.g.. "Only third year students were **able/allowed** to go." All were able to go, but all were not allowed to go.	GLWC
Grammar	Lexis	Word: word form wrong	incorrect	This word should have been in a different form for this context.	Use this for words in the wrong form, not covered by the other labels below.	GLWF
Grammar	Lexis	Word: word form wrong - adjective	incorrect	Wrong word form: Use the adjective form of the word.	Use this when the learner should have used the adjective form of the word, e.g.. I was too frightened to move.	FGADJ
Grammar	Morphology	Word: word form wrong - past participle	incorrect	This the wrong form of the word. Use the "-ed" form of the word.	Use this tag when the learner did not use the past participle e.g. "I was too frighten to move."	GMWFP
Grammar	Morphology	Word: word form wrong - present participle	incorrect	This the wrong form of the word. Use the "-ing" form of the word.	I am busy work in the garden.	GMWFPP
Grammar	Lexis	Word: wrong word	incorrect	This is the wrong word. Find and use the correct word for the context.	Use this when a student should have used another word instead e.g.. "Students should be learned (**taught**) to..." or "injury" (damage) to property." Property cannot hurt.	GLWW
Discourse	Morphology	Word: wrong word - modal	incorrect	This is the wrong word. Distinguish between permission, possibility and obligation.	To enter the concert, you may (**must**) bring your ticket.	GLWWM

Grammar	Lexis	Word: Wrong word - Time/temporal adverbial	incorrect	This time-word does not fit the rest of the essay.	Use this when a student uses e.g. a word in the past-tense when the whole essay is written in the present tense.	GADV

8. CONCLUSION AND RECOMMENDATION FOR FURTHER RESEARCH

This final chapter will present a summary of the major findings of the dissertation, indicating if and how the research objectives have been met. It will also indicate which new areas of research have been identified and present my opinion on how to conduct further research in those fields.

8.1 Have the objectives of the study been met?

The problem statement in section 1.5 stated that the dissertation will attempt to answer the questions:

1. Is it possible to create standardized written feedback on second language writing?
2. Can standardization of feedback ensure (a) clarity and (b) consistency?

The answer to both these questions is a definite "yes", but there are limitations, not least of which is the fact that the answer to question one depends on the fulfilment of question two. In order to answer question two, the objectives of this dissertation were to use best practice in feedback to create a set of tags for use in a computer marking interface. These tags had to be tested to ensure that the feedback tag set is clear and consistent.

This resulted in the findings of chapter two, in which problems with feedback as practice were distinguished from problems with feedback as concept. I argued that although the concept is sound, it may still fail as pedagogic tool due to ineffective practice. To counter that, a checklist of "best practice" in feedback was created that could be used to qualify the computerised marking technique advocated here, as effective or ineffective.

The methodology used to create and test the tag set, was described in chapter three. The methods used are sufficiently documented and transparent enough that comparative studies can be conducted to verify the present study. Possible problems with the testing technique used, were identified but the conclusion was reached that they will not influence the findings relevant to the objectives of this particular study.

In creating the tag set, an additional opportunity presented itself to establish what feedback practice is ineffective. Chapter four reports on what lecturers typically focus on when marking student texts. The findings indicated that lecturers are more likely to focus on surface elements in the writing. This corresponds to findings on the international front. Possible reasons for the phenomenon are indicated.

The data generated by the first experiment, and using insights gained from the literature review, was used to create an initial feedback tag set. The reasoning behind the initial tag set is presented in chapter five. The relevance and use of XML and error categorization is explained and the tag set is evaluated according to the check list for ideal feedback as established in chapter two. According to the checklist criteria, the feedback tag set marking technique should be very effective. This inference was duly

tested and the resulting findings were discussed in chapter six. The test confirmed that the tag set marking technique is more effective than the conventional "hieroglyphic" marking technique, and much more effective than minimal marking. Only a few minor changes were therefore made to the final tag set as presented in chapter seven.

This dissertation therefore proves that it is possible to create standardized feedback which is clear and consistent. However, a computer marking interface will definitely be needed for it otherwise it would take too much time. The only other variable in the equation (the person doing the marking) still poses a problem seeing as they cannot be completely objective. Teacher training will therefore still remain a priority in order to ensure standardization. For the purposes of this dissertation, however, the objectives had been met.

8.2 Where to from now?

The tag set has been tested and refined and it is working. It answers to most of the criteria established for effective feedback and proved that standardization of feedback (to an extent) is possible. Now we have to implement the tag set in the software created for this purpose. This will happen during the course of the coming year, during which further testing will be done. In this section, I want to raise some issues for further investigation and further development.

8.2.1 Immediate implementation

Simultaneous with this paper, the requirements for the user and student interfaces were written and the specifications given to programmers. The interface is therefore already programmed and will be implemented for testing in 2006. We hope to test it at more than one campus/university, in order to get as broad a testing field as possible.

The implementation will need a short but thorough training session for all the staff members who will use the system. It is expected that all members of staff in the department of English will use the system so that we can get as much data as possible from them. Staff members will first be introduced to the qualities of effective feedback, explaining how this system will assist them in providing it. They will then be introduced to the system in a hands-on workshop.

8.2.2 Further refinements – tag set and interface

During the course of the year, questionnaires will regularly be sent to the lecturers using the system. The questionnaires will be aimed at finding the shortcomings of both the tag set and the user interface. We want to make the system as user friendly as possible; therefore the first year will be very important for finding the "little foxes" that irritate the users. It is of course impossible to foresee all possible problems with the tag set and user interface, so the importance of this stage of development should not be underestimated. The same applies for the user interface on the learner side. One has to be prepared for learners who know very little about computers and who will be completely in the dark with this new system, simply because it is on the computer.

One also has to be prepared for lecturers who are complete technophobes and would not want to use the computer.

Also during this stage, we will be able to gather data on the types of errors that lecturers comment on. This may be used for further marker training, as well as for the positioning of the buttons and user options.

The tag set will of course also be in need of refinement. Lecturers could comment on "missing tags" for something the tag set fails to make provision. Learners could also indicate that they really do not understand the error labels in their texts.

8.2.3 Creation of the exercise database

The idea is that, based on the individual learner's problems, he/she will get exercises automatically from the computer. We will need a very large database of exercises for this purpose. The exercises have not yet been created, so it will have to be finished as soon as possible.

The exercises will have to fit a set of criteria in order to make them as functional as possible. There are hundreds of different ways one can tackle the design of exercises, but the overarching theme here would probably be on the lines of "comprehensible input" rather than simply formal instruction. The aim is two-way communication. This is a whole different field of study and the details will have to be sorted out by the researchers working on the exercises.

8.2.4 Creation of the marked essay database

The marked essays will automatically be archived in a corpus of student writing. We will therefore start creating a big database, which could be very useful for longitudinal studies. With the creation of any database, there is always a lot of administration in order to make the database useful.

Students would also be asked to sign a form in which they state that their essays may be used for academic purposes. This is just an ethical guideline that will ensure that the data in the database may be used for research. The focus of further research will determine how thoroughly the corpus will be annotated.

Of course, with the creation of a system such as this, there is a lot of programming to be done. If the system is to be optimally useful, it should include a user management function for managing the files, the student marks, the different classes and so on. The idea is also that students' essay mark will not be accepted before they have done their revision exercises. This type of administration can only be done on a hacker-safe computer and that will take a lot of programming to achieve. Once this is done, it could of course assist in the longitudinal studies that would be possible with this system.

8.2.5 Writing algorithms for automatic error detection

The marking system has the aim of saving the lecturers' time when marking. The logical next step would be to have a screening mechanism – a type of advanced grammar checker – that will screen the text for errors before the lecturer starts marking. This will result in much less time being spent on surface errors. The lecturer can then pay more attention to the organizational structure of the text, which is still much more difficult to analyze by computer. The ultimate goal is of course to assist the lecturer with that too. Steps have been taken to start creating this purpose-built grammar and spelling checker and the first version has been tested already. It may be implemented in the marking system as a screening element, early in 2006.

8.2.6 Letting the computer monitor the marker

This is an important area of study. From the data gathered for this dissertation, it was quite clear that markers focused on the surface structure, and especially on the "wrong" pieces of the surface structure. Since the system aims at bettering feedback as a total system, this problem has to be addressed as well.

The proposal is therefore that once a lecturer has finished marking, the computer will count the types of errors marked. If the lecturer failed to comment on structure and organization, the computer should prompt the lecturer to investigate those areas as well. However, this poses a problem. When did a lecturer comment on structure and organization enough? Is there any ratio to go by? There does not appear to be any research of the sort, so this is an area wide open for investigation. To further

complicate the matter, one should keep in mind that a paragraph with a central thought, can only be in the wrong place (regarding the structure of the text) once. Inside that paragraph however, we can have numerous erroneous sentences, containing in their turn, numerous grammatical and lexical errors. To establish a ratio whereby the computer can judge if or if not a lecturer paid sufficient attention to organization would be difficult.

As mentioned in chapter 5, lecturers will in some cases be forced to insert more than one tag for some errors. These may then need additional explanation in the form of comments. It could be necessary to remind the lecturers of these comments in some cases. Where and how exactly, will need to be researched. Monitoring the additional comments that the lecturers insert without prompting from the computer, may also prove insightful into the refining of the tag set.

8.2.7 Thinking about computers

The other extreme of technophobes, are people have the notion that "if it is on the computer, I don't have to do a thing." Evidence of this is for example the blind trust placed in spelling checkers. It is thus important to make it absolutely clear for the lecturers and students alike, that this system is a tool. It is not a magic formula to take all their responsibilities away, but it is a tool that will relieve some of the burden of that responsibilities.

8.2.8 Examples

In section 5.1.3 it was mentioned that students would be able to ask for additional information on errors. This information would be in the form of examples. Due to constraints on time, funding and availability of research partners, this feature could not be programmed into the initial system. To ensure that these examples are effective and clear will need plenty of research.

8.3 Closing thought

There is a saying: "If it ain't broke, don't fix it". Some engineering friends of mine altered it to: "If it ain't broke, it don't have enough functions."

Feedback it seems has too many functions. Therefore it is "broke". We therefore try to make something with too many functions simply better than it is now. It will still have little "broken" pieces, but at least some of the functions will be working better.

9 List of sources

ASKEW, S and LODGE, C. 2000. Gifts, ping-pong and loops – linking feedback and learning. (*In* Askew, S., *ed.* Feedback for learning. London: Routledge Falmer. p.1-17)

BACHA, N. 2001. Writing evaluation: what can analytic versus holistic essay scoring tell us? *System.* 29: 371-383.

BAX, S. 2003. CALL – past, present and future, *System,* 31:13-28.

CAGNEAUX, E, DENNIS, S. and GRANGER, S. 1998. Computer-aided error analysis. *System,* 26:163-174.

CHAPELLE, C.A. 2003. English language learning and technology. Amsterdam: John Benjamins.

CHAUDRON, C. 1983. Evaluating writing: Effects of feedback on revision. Revised version of a paper presented at the 17th annual TESOL convention, Toronto, Canada (ERIC document No. ED 227 706).

CHO, Y. 2003. Assessing Writing: are we bound by only one method? *Assessing Writing,* 8:165-191.

CHRYSTAL, D. 1992. An Encyclopaedic dictionary of language and languages. Oxford: Blackwell.

CLARKE, S. 2000. Getting it right – distance marking as accessible and effective feedback in the primary classroom. (*In* Askew, S., *ed.* Feedback for learning. London: RouledgeFalmer. p. 32-45.).

COETZEE-VAN ROOY, A.S. 2000. Cultural identity and acquisition planning for English as a second language in South Africa. Vanderbijlpark: Potchefstroom University for CHE, South Africa (Thesis: PhD.)

COETZEE-VAN ROOY, A.S. and VERHOEF, M.M. 2000. Perceptions of English proficiency – views from Southern Sotho speakers. *South African Journal of Linguistics,* Supplement 38: 63-186, December.

COHEN, A.D. and CAVALCANTI, M.C. 1990. Feedback on compositions: teacher and student verbal reports. (*In* Kroll, B., *ed.* Second language writing – research insights for the classroom. Cambridge. Cambridge University Press. p. 11-23.)

CORDER, S.P. 1973. Introducing applied linguistics. Harmondsworth, Middlesex: Penguin.

CONNOR, U. and FARMER, M. 1990. The teaching of topical structure analysis as a revision strategy for ESL writers. (*In* Kroll, B., *ed.* Second language writing – research insights for the classroom. Cambridge University Press. p. 126-139.)

CRYSTAL, D. 2003. A dictionary of linguistics and phonetics. 3rd ed. Oxford: Blackwell.

EGGINS, S. 2000. An introduction to systemic functional linguistics. London: Continuum.

ELLIS, R. 1992. Second language acquisition and language pedagogy. Clevedon: Multilingual matters.

ELLIS, R. 1994. Understanding second language acquisition. Oxford: Oxford University Press.

ELLIS, R. 1996. The study of second language acquisition. Oxford University Press.

FATHMAN, A.K. 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: Cambridge University Press. p. 11-23.)

FERRIES, D. and ROBBERTS, B. 1996. Error correction in L2 writing classes: How explicit does it need to be? [online] Available: http://www.JSLW.

FERRIS, D. 2003. Responding to writing. (*In* Kroll, B. *ed.* Exploring the dynamics of second language writing. Cambridge University Press. p. 119-140)

FERRIS, D.L. 2002. Treatment of error in second language student writing. The University of Michigan Press.

FRODESEN, J. and HOLTEN, C. 2003. Grammar and the ESL writing class. (*In* Kroll, B. *ed.* Exploring the dynamics of second language writing. Cambridge: Cambridge University Press. p. 141-161.)

GENNRICH, D.1997. Teaching writing for 2005: new models of assessment and feedback. *NAETE Journal*, 12:16-31.

GESLIN, N. and WADE, R. 2000. Teaching a discourse grammar in a multimedia program. *Journal for language teaching*, 34(2): 157-169.

GILLER, A. 2000. Te err is human, to correct is...undecided erred corrections: The correctness of correcting errors. [Online] Available: http//www.lwc.edu/staff/goet/span400research.htm.

GIVÓN, T. 1989. Mind, code and context – essays in pragmatics. London: Lawrence Erlbaum Associates.

GRANGER, S. 2002. A bird's-eye view of learner corpus research. (*In*: Granger, S. Hung, J and Petch-Tyson, S. *eds*. Computer learner corpora, second language acquisition and foreign language teaching. Amsterdam: John Benjamins. p. 3-33.)

GRANGER, S. and MEUNIER, F. 2003. Error tagging project – revised guidelines. (Unpublished document).

HAINES, C. 2004. Assessing students' written work – marking essays and reports. London: RoutledgeFalmer.

HALLIDAY, M.A.K. and MATTHIESSEN, C.M.I.M. 2004. An introduction to functional grammar. 3rd ed. London: Arnold.

HALLIDAY, M.A.K. and HASSAN, R. 1990. Language, context, and text: aspects of language in a social-semiotic perspective. Oxford: Oxford University Press.

HARGREAVES, E. MCCALLUM, B. and GIPPS, C. 2000. Teacher feedback strategies in primary classrooms – new evidence. (*In* Askew, S., *ed*. Feedback for learning. London: RouledgeFalmer. p.1-17)

HARTMANN, R.R.K. and STORK, F.C. Dictionary of language and linguistics. London: Applied Science.

HASAN, R. 2002. Workshop on systemic functional linguistics. Presented at LSSA conference, Pietermaritzburg. July.

HATTINGH, K. 2004. The syntactic development of grade 12 ESL learners. Potchefstroom: North-West University, South Africa. (Thesis – MA).

HEDGECOCK and LEFKOWITZ 1994. Feedback on feedback: assessing learner receptivity to teacher response in L2 composing. *Journal of Second Language Writing*, 3(2):141-163.

HONEY, J. 1997. Language is power. The story of standard English and its enemies. London: Faber and Faber.

HYLAND, F. 2003. Focusing on form: student engagement with teacher feedback. *System*, 31:217-230.

HYLAND, F. 1998. The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, 7(3): 255-286.

HYLAND, K. 1990. Providing productive feedback. *ELT Journal*, 44(4):279-285.

JACOBS, G.M., CURTIS, A., BRAINE, G. and HUANG, S. 1998. Feedback on student writing: taking the middle path. *Journal of Second Language Writing*, (7)3:307-317.

JAMES, C. 1997. Errors in language learning and use – exploring error analysis. Edinburgh Gate: Longman.

JARVIS, H. 2004. Investigating the classroom applications of computers on EFL courses at higher education institutions in UK. *Journal of English for Academic Purposes,* 3: XXX.

JOHNSON, K. and JOHNSON, H. (eds.) 1998. An encyclopaedic dictionary of applied linguistics. Oxford: Blackwell.

KATZNELSON, H, PERPIGNAN, H and RUBIN, B. 2001. What develops along with the development of second language writing? Exploring the "by-products". *Journal of Second Language Writing,* 10(3):141-159.

KRASHEN, S.D. 1985. The input hypothesis: issues and implications. London: Longman.

KRAPELS, A.R. 1990. An overview of second language writing process research. (*In* Kroll, B., *ed.* Second language writing – research insights for the classroom. Cambridge: Cambridge University Press. p. 37-56.)

KROLL, B. 1990. What does time buy? ESL student performance on home versus class compositions. (*In* Kroll, B., *ed.* Second language writing – research insights for the classroom. Cambridge. Cambridge University Press. p. 140-155.)

LEE, I. 2002. Helping students develop coherence in writing. *English teaching forum,* 40(3):32-39. July.

LEKI, I. 2003. A challenge to second language writing professionals: Is writing overrated? (*In* Kroll, B. *ed.* Exploring the dynamics of second language writing. Cambridge University Press. p. 315-331.)

LEKI, I. 1990. Coaching from the margins: issues in written response. (*In* Kroll, B., *ed.* Second language writing – research insights for the classroom. Cambridge. Cambridge University Press. p. 57-69.)

LEKI, I. 1991. The preferences of ESL students for error correction in college-level writing classes. *Foreign Language Annals,* 24(3): 203-218.

LIGHTBOWN, P.N and SPADA, N. 1990. Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition,*12: 429-448.

LIGHTBOWN, P.M. and SPADA, N. 1999. How languages are learned – revised edition. Oxford: Oxford University Press.

LYSTER, R. and RANTA, L. 1997. Corrective feedback and learner uptake. *Studies in Second Language Acquisition,* 19(1):37-66.

MATSUDA, P.K. 2003. Second language writing in the twentieth century: a situated historical perspective. (*In* Kroll, B. *ed.* Exploring the dynamics of second language writing. Cambridge University Press. p.15-34.)

MILROY, J and MILROY, L. 1992. Authority in language. Investigating language prescription and standardisation. London: Routledge.

MOLETSANE, J.R. 2002. Selective error correction in ESL narrative compositions. Potchefstroom: PU for CHE. (Mini-dissertation – Magister Artium).

MONYAKI, B.S. 2001. A comparison of teachers' and learners' views on feedback strategies used in the writing process. Potchefstroom: PU for CHE. (Mini-dissertation – Magister Artium).

MUNCIE, J. 2000. Using written teacher feedback in EFL composition classes. *ELT Journal,* 54(1):47-53.

NAGATA, N. 1993. Intelligent computer feedback for second language instruction. *The Modern Language Journal,* 77(3):330-339.

NEL, C., DREYER, C. and CARSTENS, W.A.M. 2001. Educational technologies: a classification and evaluation. *Journal for Language Teaching,* 35(4):238-258, Dec.

NWAILA, C. 1996. Black English and education in South Africa – an investigation. Pretoria: University of Pretoria. (Thesis – D.Litt.)

PARKINSON, J. 2001. Explicit teaching of grammar and improvement in the grammar of student writing. *Journal for Language Teaching,* 35(4):278-292.

PAULUS, T.M. 1999. The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing,* 8(3): 256-289.

PENNINGTON, M.C. 2003. The impact of the computer in second language writing. (*In* Kroll, B. *ed.* Exploring the dynamics of second language writing. Cambridge University Press. p. 287-310.)

PORTE, G. 1996. When writing fails: how academic context and past learning experiences shape revision. *System,* 24 (1): 107-116

PRETORIUS, E.J. 2002. Reading ability and academic performance in South Africa: are we fiddling while Rome is burning? *Language Matters,* 33:169-196.

REID, J. 1994. Responding to ESL students' texts: the myths of appropriation. *TESOL Quarterly,* 28(2), Summer.

SELINKER, L. and LAMENDELLA, J.T. 1979. The role of extrinsic feedback in interlanguage fossilization – a discussion of "rule fossilization: a tentative model". *Language Learning,* 29(2): 363-375.

SILVA, T. 1990. Second language composition instruction: developments, issues, and directions in ESL. (*In* Kroll, B., *ed.* Second language writing – research insights for the classroom. Cambridge. Cambridge University Press. p. 11-23.)

SPEICHER, B.L. and BIELANSKI, J.R. 2000. Critical thoughts on teaching standard English. *Curriculum Inquiry*, 30(2):147-169

SPENCER, B. 1998. Responding to student writing: strategies for a distance-teaching context. Pretoria: University of South Africa. (Thesis – D.Litt.)

SPINGIES, C. The marking and assessment of children's writing. *Crux*, 24(14):18-32.

STORCH, N. and TAPPER, J. 1997. Student annotations: what NNS and NS university students say about their own writing. *Journal of Second Language Writing*, 6(3):245-264.

SULLIVAN, N. and PRATT, E. 1996. A comparative study of two ESL writing environments: a computer-assisted classroom and a traditional oral classroom. *System*, 29(4):491-501.

SUSSER, B. 1998. The mysterious disappearance of word processing. *Computers and Composition*, 15(3):347-371.

TAYLOR, B.P. 1981. Content and written form: a two-way street. *TESOL Quarterly*, 15(1):5-13.

TRUSCOTT, J. 1996. The case against grammar correction in L2 writing classes. *Journal of Second Language Writing*, 46(2):327-369.

TURTON, N.D. 1995. ABC of common grammatical errors: for learners and teachers of English. London: Macmillan Educational.

VAN DER WALT, J.L., VAN DER WALT, I. and DREYER, C. 1994. The effect of error correction on the grammatical competence of ESL students. *Journal for Language Teaching*, 28(1):7-17

VAN WYK, A, 2002. A university bridging course focusing on academic reading and writing skills. *Journal for Language Teaching*, 36(3/4):220-232, Dec.

WEBB, V. 2002. English as a second language in South Africa's tertiary institutions: a case study at the University of Pretoria. *World Englishes*, 21(1):49-62. March.

WEIGLE, S.C. 2002. Assessing writing. Cambridge: Cambridge University Press.

WIBLE, D., KUO, C., CHIEN, F., LIU, A. and TSAO, N. 2001. A Web-based EFL writing environment: integrating information for learners, teachers, and researchers. *Computers and Education*. 37(3-4):297-315.

ZAMEL, V. 1985. Responding to student writing. *Tesol Quarterly*. (19)1:79-101. March.

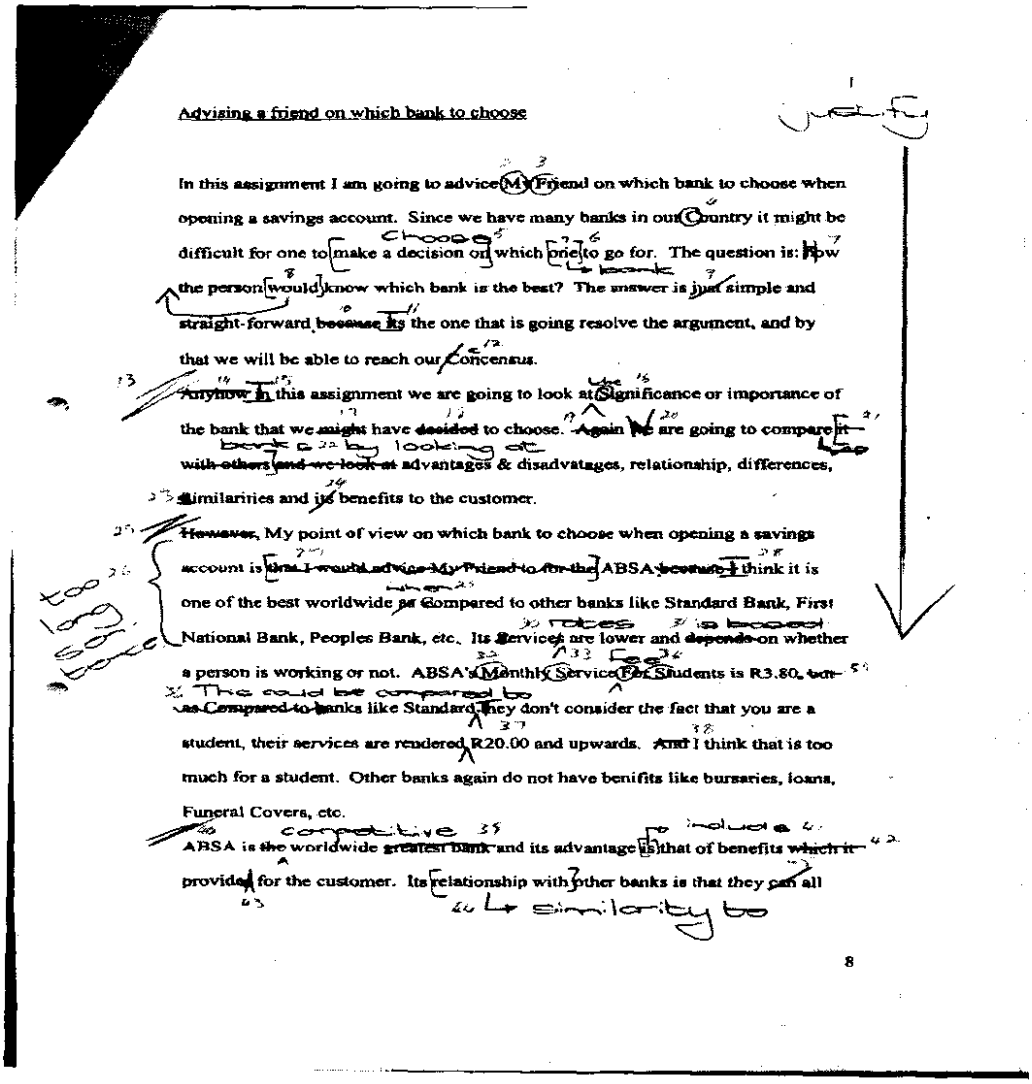
9.1 Sources on XML and SGML

EAGLES: 2003. Introduction and background to the EAGLES project. [Web:] http://www.ilc.cnr.it/EAGLES96/home.html.

Corpus Encoding Standard - Document CES 1. Version 1.5. Last modified 20 March [Web:] http://www.cs.vassar.edu/CES/. Accessed 06/12/2004 at 11:43

http://lamp.infosys.deakin.edu.au/sgml/sgmlfaq.txt Accessed 06/12/2005 at 12:10

http://xml.coverpages.org/FAQv21-200201.html Accessed 06/12/2005 at 12:12
http://www.isgmlug.org/sgmlhelp/g-index.htm Accessed 06/12/2005 at 12:32
http://xml.coverpages.org/sgml.html Accessed 06/12/2005 at 12:30
http://www.w3.org/TR/REC-xml/ Accessed 06/12/2005 at 12:26
http://xml.coverpages.org/sgmlhist0.html Accessed 06/12/2005 at 12:24
http://xml.coverpages.org/general.html Accessed 06/12/2005 at 12:22
http://xml.coverpages.org/EnglishFAQ-NOT-20020401.txt Accessed 06/12/2005 at 12:19
http://www.w3.org/TR/NOTE-sgml-xml.html Accessed 06/12/2005 at 12:18
http://www.w3.org/XML/ Accessed 06/12/2005 at 12:15



Advising a friend on which bank to choose

In this assignment I am going to advice My Friend on which bank to choose when opening a savings account. Since we have many banks in our Country it might be difficult for one to make a decision on which one to go for. The question is: How the person would know which bank is the best? The answer is just simple and straight-forward because its the one that is going resolve the argument, and by that we will be able to reach our Concensus.

Anyhow in this assignment we are going to look at Significance or importance of the bank that we might have decided to choose. Again we are going to compare it with others and we look at advantages & disadvatages, relationship, differences, similarities and its benefits to the customer.

However, My point of view on which bank to choose when opening a savings account is that I would advise My Friend to for the ABSA because I think it is one of the best worldwide as Compared to other banks like Standard Bank, First National Bank, Peoples Bank, etc. Its Services are lower and depends on whether a person is working or not. ABSA's Monthly Service Fee Students is R3.80, but As Compared to banks like Standard they don't consider the fact that you are a student, their services are rendered R20.00 and upwards. And I think that is too much for a student. Other banks again do not have benifits like bursaries, loans, Funeral Covers, etc.

ABSA is the worldwide greatest bank and its advantage is that of benefits which it provides for the customer. Its relationship with other banks is that they can all

8

EXERCISE:

The following paragraphs were all written by students. The writers have made errors. We have marked the errors for you. We want you to correct those errors. On your answer sheet, write down the number of the error. Write down:
(A) what you consider the error to be, and
(B) how you would correct that specific error.

The following two paragraphs come from an essay, arguing that university degrees are too theoretical and should be more practical:

Why not *mingle* (better word ①) the theory and **the practical** (②) of each subject every period, then the student will know why he (③ Gender) is learning the theory and can prove it by doing it practical? (④ Part of speech) The first part of class can be used to explain the theory and the other half for class discussions, an act, writing ~~of something~~ (⑤), visiting a hospital, (⑥) court etc. From the beginning of each degree they (⑦ ?) should focus more on the practical **side** (⑧ aspect?) and how you will use your theory to act in a specific work situation.

By focusing more on the practical side of the work, students become (⑨) faster grown-up, because they are more positive in their work. There are facts that one (⑩ Pronouns) must know and learn, but then again everyone has his or hers own opinion and if you can prove your opinion then it will and can work in the

real world. They must teach you how to think for yourself and to be creative. The world outside *doesn't* want old **ideas** but rather new ideas from the younger generation.

The following two paragraphs are from an essay in which the author had to advise a friend on which bank to choose.

Anyhow, in this assignment we are going to look at *Significance* **or importance** of the bank that we might have decided to choose.

Again we are going to compare it with others and we look at advantages & *disadvatages*, **relationship**, differences, similarities and its benefits to the customer.

(TURN OVER)

However, **My** point of view on which bank to choose when opening a savings account, is that I would *advice* **My friend** *to for* ABSA **because** I think it is one of the best worldwide as compared to other banks like Standard Bank, First National Bank, Peoples Bank, etc. Its services are lower and depends on whether a person is working or not. ABSA's monthly *service* for Students is R3.80, but as Compared to banks like Standard they don't consider the fact that you are a student, their services are rendered R20.00 and upwards. And I think that is too much for a student. Other banks do not have *benifits* like bursaries, loans, funeral **covers** etc.

The following two paragraphs are from an essay arguing that soccer players should not be paid more:

In this country rugby players are already being paid too much for being sport stars. When playing rugby they don't have their heart in winning. They know that they will get paid whether they win or loose, so they don't play for the love of the game but for the money. It's said that some rugby players are being paid more than President Mbeki, this means that they are considered to be **worth** more than the most important person in this country. *This is not the example we need to set for the children, proclaiming that the sportstars are the best, even though they seldom win any* ***games.***

When soccer players refuse to play a game that is important not only for the country but for the nation's moral as well, it sets a bad example for the children – don't do something if you're not paid enough. The minister of sport should put his foot down and say that people won't get paid if they don't do their jobs. If a game is not won, they don't get paid. If a

player wants *to go to another country to play for them because he'll get more pay.* let him go. There **are** more than enough talent in our country, especially soccer players, who will be more than willing to play for South Africa *because of love for* **his** country and for the love of the game.

EXERCISE

The following paragraphs were all written by students. The writers have made errors. We have given you an indication of the amount of errors on each page. On your answer sheet, **mark the error and correct it**. Write down:
(A) what you consider the error to be, and
(B) how you would correct that specific error.

The following two paragraphs come from an essay arguing that university degrees are too theoretical and should be more practical:

Why not mingle the theory and the practical of each subject every period, then the student will know why he is learning the theory and can prove it by doing it practical? The first part of class can be used to explain the theory and the other half for class discussions, an act, writing of something, visiting a hospital, court etc. From the beginning of each degree they should focus more on the practical side and how you will use your theory to act in a specific work situation.

By focusing more on the practical side of the work, students become faster grown-up, because they are more positive in their

work. There are facts that one must know and learn, but then again everyone has his or hers own opinion and if you can prove your opinion then it will and can work in the real world. They must teach you how to think for yourself and to be creative. The world outside doesn't want old ideas but rather new ideas from the younger generation.

Comment: 9 errors on this page

The following two paragraphs are from an essay in which the author had to advise a friend on which bank to choose.

Anyhow, in this assignment we are going to look at Significance or importance of the bank that we might have decided to choose. Again we are going to compare it with others and we look at advantages and disadvatages, relationship, differences, similarities and its benefits to the customer.

Comment: 12 errors on this page

(TURN OVER)

However, My point of view on which bank to choose when opening a savings account, is that I would advice My Friend to for ABSA because I think it is one of the best worldwide as compared to other banks like Standard Bank, First National Bank, Peoples Bank, etc. Its services are lower and depends on whether a person is working or not. ABSA's monthly service for Students is R3.80, but as Compared to banks like Standard they don't consider the fact that you are a student, their services are rendered R20.00 and upwards. And I think that is too much for a student. Other banks do not have benifits like bursaries, loans, funeral covers, etc.

Comment: 13 errors on this page

The following two paragraphs are from an essay arguing that soccer players should not be paid more:

In this country rugby players are already being paid too much for being sport stars. When playing rugby they don't have their heart in winning. They know that they will get paid whether they win or loose, so they don't play for the love of the game but for the money. It's said that some rugby players are being paid more than President Mbeki, this means that they are considered to be worth more than the most important person in this country. This is not the example we need to set for the children, proclaiming that the sportstars are the best, even though they seldom win any games.

When soccer players refuse to play a game that is important not only for the country but for the nation's moral as well, it sets a bad example for the children – don't do something if you're not paid enough. The minister of sport should put his foot down and say that people won't get paid if they don't do their jobs. If a game is not won, they don't get paid. If a player wants to go to

Comment: 11 errors on this page

another country to play for them because he'll get more pay, let him go. There are more than enough talent in our country, especially soccer players, who will be more than willing to play for South Africa because of love for his country and for the love of the game.

Comment: 4 errors on this page

ADDENDUM B: Tag set test

EXERCISE

The following paragraphs were all written by students. The writers have made errors. We have marked the errors for you. We want you to correct those errors. On your answer sheet, write down the number of the error. Write down:
(A) what you consider the error to be, and
(B) how you would correct that specific error.

The error and the comment referring to it, will be written in the same way. If the error is written in *italics*, the comment will also be in *italics*. If the error is underlined or **bold**, the comment will also be underlined or **bold**.

The following two paragraphs come from an essay, arguing that university degrees are too theoretical and should be more practical:

Why not *mingle* the theory and **the practical** of each subject every period, then the student will know why he is learning the theory and can prove it by doing it practical? The first part of class can be used to explain the theory and the other half for class discussions, an act, writing of something, visiting a hospital, court etc. From the beginning of each degree *they* should focus more on the practical **side** and how you will use your theory to act in a specific work situation.

Comment: *Use the right register. Make sure you are not too informal or too formal for the context of your writing.*

Comment: **This word should have been in a different form for this context.**

Comment: Be careful of gender bias.

Comment: Wrong word form. Use the adjective form of the word.

Comment: This word does not need to be here. Make sure that you do not repeat yourself.

Comment: You should rather insert a linking word here than a punctuation mark.

Comment: *Be more specific.*

Comment: **There are more appropriate words available with which you can express this idea.**

By focusing more on the practical side of the work, students become faster grown-up, because they are more positive in their work. There are facts that one must know and learn, but then again everyone has his or hers own opinion and if you can prove

Comment: There are more appropriate words available with which you can express this idea.

your opinion then it will and can work in the real world. They must teach you how to think for yourself and to be creative. The world outside *doesn't* want old **ideas** but rather new ideas from the younger generation.

Comment: Be consistent in your style. Use "one" or "you" but not both.

Comment: *Rather do no use contractions in academic writing.*

Comment: **Support your argument with an example or fact.**

The following two paragraphs are from an essay in which the author had to advise a friend on which bank to choose.

Anyhow, in this assignment we are going to look at *Significance* **or importance** of the bank that we might have decided to choose.

Again we are going to compare it with others and **we** look at advantages and *disadvatages*, **relationship**, differences, similarities and its benefits to the customer.

(TURN OVER)

Comment: Use the right register. Make sure you are not too informal or too formal for the context of your writing.

Comment: *Your use of capital letters or lower case letters is incorrect.*

Comment: **This word does not need to be here. Make sure that you do not repeat yourself.**

Comment: This word does not need to be here. Make sure that you do not repeat yourself.

Comment: *There is a better word available.*

Comment: **You need to insert a verb here.**

Comment: *This word is spelled incorrectly.*

Comment: **Be more specific. This word needs to be in the plural.**

Comment: Your word order is incorrect or could be better. Change it.

However, **My** point of view on which bank to choose when opening a savings account, is that I would *advice* **My Friend** *to for* ABSA **because** I think it is one of the best worldwide as compared to other banks like Standard Bank, First National Bank, Peoples Bank, etc. Its services are lower and depends on whether a person is working or not. ABSA's monthly *service* for Students is R3.80, but as Compared to banks like Standard they don't consider the fact that you are a student, their services are rendered R20.00 and upwards. And I think that is too much for a student. Other banks do not have *benifits* like bursaries, loans, funeral **covers**, etc.

Comment: *There is a better word available.*

Comment: **This paragraph does not link with the one before it. Show why you place it here and how it relates to the previous or place it somewhere else.**

Comment: *This word should have been in a different form for this context.*

Comment: **Your use of capital letters or lower case letters is incorrect.**

Comment: *You need to insert a word here.*

Comment: **You need to insert a punctuation mark here.**

Comment: Support your argument with an example or fact.

Comment: You should rather insert a linking word here than a punctuation mark.

Comment: *You need to insert a word here.*

Comment: This sentence is too long. Break it up. Use punctuaion. Your word order is incorrect. Change it.

Comment: *This word is spelled incorrectly.*

Comment: There is no need to indicate that the word is a plural.

Comment: Make sure your facts are correct.

The following two paragraphs are from an essay arguing that soccer players should not be paid more:

In this country rugby players are already being paid too much for being sport stars. When playing rugby they don't have their heart in winning. They know that they will get paid whether they win or loose, so they don't play for the love of the game but for the money. It's said that some rugby players are being paid more than President Mbeki, this means that they are considered to be **worth** more than the most important person in this country. *This is not the example we need to set for the children, proclaiming that the sportstars are the best, even though they seldom win any* ***games.***

When soccer players refuse to play a game that is important not only for the country but for the nation's moral as well, it sets a bad example for the children – don't do something if you're not paid enough. The minister of sport should put his foot down and say that people won't get paid if they don't do their jobs. If a game is not won, they don't get paid. If a player wants *to go to*

Comment: You need to insert a punctuation mark here.

Comment: Carefully consider whether the active voice will not work better for this sentence.

Comment: This word is spelled incorrectly.

Comment: Support your argument with an example or fact. You should always indicate where you got your information from. Beware of plagiarism. Use the correct formatting to indicate the source.

Comment: This sentence is too long. Break it up. Use punctuation.

Comment: **Be more specific. There is a better word available.**

Comment: *This statement is going against your main argument. If you want to show the other side of the argument, make sure you indicate it as such.*

Comment: **This sentence is too vague. It looks like an unfinished thought. You should complete this sentence to indicate that you have finished the thought.**

Comment: Your layout is not right or could have been better.

Comment: Homophone/homonym error: make sure you used the correct word or the correct spelling of the word.

Comment: You allow no room for any "grey areas". You should always consider that there are two sides to anything.

another country to play for them because he'll get more pay, let him go. There **are** more than enough talent in our country, especially soccer players, who will be more than willing to play for South Africa *because of love for* **his** country and for the love of the game.

Comment: *Your word order is incorrect or could be better. Change it.*

Comment: **The form of the verb must agree with the subject it refers to.**

Comment: There is a specific way of writing this. Make sure you use the more usual way.

Comment: **Be consistent in your style.**

ADDENDUM C: Initial tag set without XML

SUPEROR DINATE	DOMAIN	LABEL	TYPE	FEEDBACK TAG	Example/Explanation
Grammar	Lexis	Ambiguity: False Friend	incorrect	Make sure that this is the word you wanted to use and not one sounding the same as in your first language.	Using a word that looks like a word from their native language, but means something different. Example: lemoen (Afrikaans) = Orange (English) and not lemon.
Grammar	Syntax	Ambiguous word order	unclear/	Your sentence is ambiguous. Change the word order and/or punctuation.	Use this when the sentence is ambiguous due to the word order or the punctuation. Example: Men, who have poor eyesight, make poor pilots.
Discourse	Style/register	Be impersonal	inaptness	In academic writing, try to be formal and impersonal. For example, do not use words such as "I" or "I think".	Use this when you feel that the learner should rather be impersonal. In academic writing, it is customary to be impersonal.
Discourse	Style/register	Bias	inaptness	Be careful of gender bias.	Use this when a learner makes use of language which could be considered sexist.
Presentation	Capitalization	Capitalization	incorrect	Your use of capital letters or lower case letters is incorrect.	There are plenty of Banks in my country. I am going to visit my friend in the usa.
Discourse	Coherence	Coherence: counterproductivity	inaptness	This statement is going against your main argument. If you want to show the other side of the argument, make sure you indicate it as such.	Use this when a learner makes a statement that goes against his/her main argument.
Discourse	Syntax	Coherence: unfinished thought	omission	This sentence is too vague. It looks like an unfinished thought. You should complete the sentence to indicate that you have finished the thought. Add some more concrete information to support your statement.	Use this when a learner did not complete a thought, but left it hanging, for example when "etc." is used without good cause or when a statement is left without support or explanation.

Discourse	Syntax	Cohesion: repetition of statement	inaptness	You repeat basically the same statement using different words. If you see a distinction, make it clearer.	Use this when a learner writes the same thing, but tries to disguise it as a new statement. This may indicate a weak argument or maybe weak sentence structure.
Grammar	Lexis	Collocation or fixed expression error	incorrect	There is a specific way of writing this. Make sure you use the more usual way.	Use this when a learner should have used a fixed expression or standardised collocation such as "strawberries and cream", but instead used another (uncreative) distortion such as "cream and strawberries".
Grammar	Morphology	Concord	incorrect	The form of your verb should agree with the subject it refers to.	The verb must agree with the subject in number.
Discourse	Factual correctness	Fact: unsupported argument	omission	Support your argument with an example or a fact.	Use this when a learner leaves a statement without explanation. Example: "There are many stupid rules in the bank."
Discourse	Factual correctness	Facts: Opinion	inaptness	Do not use opinions as facts, unless you indicate them as opinions.	Use this when a learner tries to pose his opinion as a solid truth, e.g. "Robbie Williams is the best singer in the world."
Discourse	Factual correctness	Facts: wrong	incorrect	Make sure your facts are correct.	Use this when a learner uses the wrong facts in a piece of writing, e.g. "Charlie Brown has a dog called Peanuts."
Discourse	POSITIVE COMMENTS	Good Reasoning	Good	You link the ideas very well.	Use this to give praise to the student when he/she manages to link ideas together in a novel way.
Discourse	POSITIVE COMMENTS	Good structure	Good	This piece of writing is well-structured.	Use this when a paragraph is well-structured. It is advisable to add a comment explaining what you found good.
Discourse	POSITIVE COMMENTS	Interesting point	Good	You have a very interesting point here. Good thinking.	Use this to give praise to the student for good insight and valid arguments.
Discourse	Structure	Introduction: weak	inaptness	The introduction to the essay should be stronger. See the note added by the lecturer.	The introduction simply restates the topic without indicating what will be discussed in the essay.

Presentation	Layout	Layout error	inaptness	Your layout is not right or could have been better.	Use this when a learner has to write something in a specific layout format, such as an official letter. It would be wise to add a comment explaining what the problem is.
Presentation	Layout	Layout inhibits reading	inaptness	Your layout makes it difficult to read this text.	Use this when the layout of a text makes it difficult to read, like when there are no spaces between paragraphs.
Grammar	Syntax	Negation incorrect	incorrect	You should make sure how to make this sentence negative.	Sometimes our teacher not allow us.
Discourse	Cohesion	Paragraph: jumbled	inaptness	This paragraph does not deal with only one thought.	Use this when a learner tries to cram too many thoughts into one paragraph. You could use this with the tag "Start new paragraph" to help the student decide where to break the paragraph.
Discourse	Cohesion	Paragraph: start new	inaptness	You should start with a new paragraph here, since you start with a new idea.	Use this when it is obvious that a learner should have started a new paragraph, but neglected to do so.
Discourse	Cohesion	Paragraph: weak opening sentence.	inaptness	You could have used a stronger opening sentence for this paragraph. Find one that is more relevant to the rest of the paragraph.	Use this when a paragraph starts with a sentence that is not relevant to (or indicative of) the content of the rest of the paragraph.
Discourse	Cohesion	Paragraphing: link	inaptness	This paragraph is continuing the thought of the previous one. Why don't you make it one paragraph?	Use this when a student just randomly inserted a paragraph marker, without regard for the theme of the paragraph.
Discourse	Cohesion	Paragraphing: relate or move	inaptness	This paragraph does not link with the one before it. Show why you place it here and how it relates to the previous or place it somewhere else.	In some cases a paragraph is out of place and should rather be placed somewhere else. Another possibility is that a paragraph is just not linked clearly enough to the one before or after it.
Grammar	Punctuation error	Punctuation: missing	incorrect	You need to insert a punctuation mark here.	Use this when a learner left out a comma, full stop or other punctuation mark.

Grammar	Punctuation error	Punctuation: unnecessary	incorrect	You do not need a punctuation mark here.	Use this when a learner inserted an unnecessary punctuation mark, or placed a punctuation mark in the wrong place. This tag could be used together with "punctuation missing" if a punctuation mark is inserted in the wrong place.
Grammar	Punctuation error	punctuation: wrong	incorrect	Your punctuation is wrong. There is another punctuation mark that will be better.	The student's use of the punctuation mark(s) is wrong. There should be punctuation, but the punctuation used is not correct . Example: They saw lions. elephants. and rhino.
Discourse	Coherence	Reasoning incoherent	inaptness	How do you come to this conclusion with the facts you mentioned? Your argument is not structured clearly enough for me to see the logical development.	Use this when you cannot see a logical link between a statement and the deduction the student makes from that.
Discourse	Factual correctness	Reference omitted/wrong	inaptness	You should indicate where you got this information from. Beware of plagiarism. Use the correct formatting to indicate the source.	Use this when a learner uses a direct quote or mentions an author, but do not indicate the source. This could also be attributed to style problems.
Discourse	Cohesion	Relevance to topic	inaptness	This statement is not relevant to the topic or the connection to the topic is not clear enough.	Use this when a statement is not directly relevant to the topic. Example: A learner attempts to persuade with an emotional argument where he/she should have used cold, hard facts.
Grammar	Syntax	Sentence cohesion	incorrect	The way you link this two sentences pr two thoughts, does not make sense or is incorrect.	Example: two sentences that have nothing to do with each other cannot be linked by "and".
Grammar	Syntax	Sentence incomplete	incorrect	This is not a complete sentence.	In the movie, "Patch Adams".
Grammar	Syntax	Sentence: run-on	inaptness	This sentence is too long. Break it up. Use punctuation.	Long sentences are difficult to read. Students should break it up into two or more coherent sentences, each with their own correct punctuation and word order.
Grammar	Syntax	sentence: construction/word	incorrect	Your word order is incorrect or could	Use this when the word order in a sentence is

		order		be better. Change it.	incorrect.
Discourse	Syntax	Sentence: active voice	inaptness	Carefully consider whether the passive voice will not work better for this sentence.	Use this where the student used the active voice, but in your opinion, the passive voice would have worked better.
Discourse	Syntax	Sentence: passive voice	inaptness	Carefully consider whether the active voice will not work better for this sentence.	Use this where the student used the passive voice, but in your opinion, the active voice would have worked better.
Grammar	Lexis	Spelling/typing error	incorrect	This word is spelled incorrectly.	Use this only for spelling or typing errors. "Aslo" for "also" and "beeutiful" for "beautiful".
Grammar	Lexis	Spelling: Ambiguity - homophone/homonym	incorrect	Homophone/homonym error: Make sure you used the correct word or the correct spelling of the word.	Confusing two words that sound or look the same e.g. "there" and "their".
Discourse	Style	Style: Construction overuse	inaptness	You use the same construction too much. This makes your reading monotonous. Try using different words together to say what you want to say.	Use this when you realise that a learner is overusing a construction. Add a note to explain.
Discourse	Style/register	Style: Contractions	inaptness	Rather do not use contractions in academic writing.	Use this when a learner used words like "won't" or "isn't" in academic writing.
Discourse	Syntax	Style: Inconsistent	inaptness	Be consistent in your style	Use this when a learner keeps switching between "one", "you", "a person" or if you see other inconsistencies in style. Maybe add a note.
Grammar	Lexis	Style: register wrong	inaptness	Use the right register. Make sure you are not too informal or too formal for the context of your writing.	Use this when a learner used an informal word in a formal piece of writing or a formal word in an informal piece of writing.

Grammar	Syntax	Style: sentence vague	omission	You leave out too much information for this sentence to make sense. Be more detailed.	Use this when a student is too cryptic to make sense, e.g. "They are both accounts..." vs. "They are both savings accounts..."or "It is as if **they** are trying to suppress **their** feelings." Who is suppressing whose feelings? Also use this when a learner uses constructions like "everyone knows" or "they say that" or when the learner uses "it" when a more specific identifying word should have been used.
Discourse	Factual correctness	Style: Tautological	inaptness	You allow no room for any "grey areas". You should consider that there are two sides to most things. Be careful of a harsh statement.	"Soldiers are always bloodthirsty killers."
Grammar	Lexis	Style: Verbosity	inaptness	You can say this in fewer words.	Using too many words to say what needs to be said.
Discourse	Factual correctness	vagueness: unspecific	inaptness	Be more specific.	Use this when a learner refers to "they" or "them" without specifying who it is. Also use this when a learner uses "etc." instead of completing a list in a neat manner. Examples: "They say that..." "They should go to hospitals, clinics, etc." instead of "Students should go to hospitals, clinics and other medical fascilities."
Grammar	Articles	Word: article wrong	incorrect	The article should have been in a different form. Distinguish between definite and indefinite articles.	They asked her to describe a missing child. They asked her to describe **the** missing child.
Grammar	Lexis	Word: Bad word	incorrect	This word is racist/sexist or derogative and should rather not be used.	Use this tag when a word is racist or sexist or in any way derogative and should not have been used in the context of the writing.

Grammar	Lexis	Word: Better word	incorrect	There are more appropriate words available with which you can express this idea.	Use this tag when a learner should have used a better word, like when a learner uses a near-equivalent form to the one intended (e.g. "not happy" for "sad") or when the learner describes something because he don't know the exact word for it.
Grammar	Lexis	Word: conditional wrong	incorrect	This is not the best (or right) way to formulate a condition.	"They could probably study better **when** they get electricity" vs. "They could probably study better **if** they get electricity."
Grammar	Syntax	Word: Omission for punctuation	omission	You should rather insert a linking word here than a punctuation mark.	Use this when a learner uses punctuation instead of a word, e.g.. "She gave me peaches, nuts, apples in a basket to take home." Also use this when a learner starts a new sentence when he/she should have continued with the previous.
Grammar	Syntax	Word: Omission general	omission	You need to insert a word here.	Example: "When you open an account they ask you **(a)** few questions." or "...only the last year or two..." vs. "...only during the last year or two..."
Grammar	Morphology	Word: omission plural marker	incorrect	You need to insert a plural marker here.	Use this when a learner did not indicate a plural where he/she should have. E.g. "their customer" or "...they want one of your family member to come..."
Grammar	Lexis	Word: Omission possession marker	incorrect	Make sure you use the correct possession marker or mark the possession correctly.	If an apostrophe is used to indicate possession, make sure that it is placed in the right position. E.g.. Students' vs. Student's
Grammar	Lexis	Word: wrong word preposition	incorrect	You use the wrong preposition here.	"days **in** which" vs. "days **on** which"
Grammar	Syntax	Word: Omission preposition	omission	You need a preposition here.	The man went **(into)** the bank.
Grammar	Syntax	Word: Omission tense marker	omission	Use the correct time word. You can say this in much less words then.	Use this when a learner uses a description of a tense, instead of just using the correct time-word.

Grammar	Syntax	Word: omission verb	omission	You need to insert a verb here.	"Only in the last year or two they begin..." vs. "Only in the last year or two **do** they begin..."
Grammar	Syntax	Word: Overinclusion general	superfluous	This word(s) does not need to be here. Make sure that you do not repeat yourself.	I am very utterly pleased.
Grammar	Syntax	Word: Overinclusion plural marker	superfluous	There is no need to indicate that the word is a plural.	all the **plenty** teachers
Grammar	Syntax	Word: Overinclusion tense marker	superfluous	You do not have to indicate the tense here.	Use this when a student indicates a tense twice or unnecessarily.
Grammar	Syntax	Word: Overinclusion unnecessary pronoun	superfluous	You do not need to indicate the subject or object of the sentence twice.	The man **he** chooses the bank he likes.
Grammar	Lexis	Word: Preposition wrong	incorrect	This is the wrong preposition. Find out which would have been the correct one to use.	Use this when a student used the wrong preposition e.g.. It is not so easy as it looks.
Grammar	Lexis	Word: Pronoun wrong	incorrect	You should have used a different pronoun here. Find out which one.	Use this when the learner used the wrong pronoun.
Grammar	Lexis	Word: Repetition	inaptness	You use the same words repeatedly. Find different words that may convey your message more clearly.	Use this tag when you realise that a student keeps on using the same word. E.g.. If a student use the word "Good" to mean "excellent" and "strong" and "hard" and "pretty" etc. This will be context sensitive. If you have to use the "Word: better word" tag a lot for the same word, rather start using the "Word: repetition" tag.
Grammar	Lexis	Word: word choice obscuring meaning	incorrect	This word is not clear enough. Find a better word to say what you want to say.	Use this when another word would make the intended meaning much clearer. E.g.. "Only third year students were **able/allowed** to go." All were able to go, but all were not allowed to go.
Grammar	Lexis	Word: word form wrong	incorrect	This word should have been in a different form for this context.	Use this for words in the wrong form, not covered by the other labels below.
Grammar	Lexis	Word: word form wrong - adjective	incorrect	Wrong word form: Use the adjective form of the word.	Use this when the learner should have used the adjective form of the word, e.g.. I was too frightened to move.

Grammar	Morphology	Word: word form wrong - past participle	incorrect	This the wrong form of the word. Use the "-ed" form of the word.	Use this tag when the learner did not use the past participle e.g. "I was too frighten to move."
Grammar	Morphology	Word: word form wrong - present participle	incorrect	This the wrong form of the word. Use the "-ing" form of the word.	I am busy work in the garden.
Grammar	Lexis	Word: wrong word	incorrect	This is the wrong word. Find and use the correct word for the context.	Use this when a student should have used another word instead e.g.. "Students should be learned (**taught**) to..." or "injury" (damage) to property." Property cannot hurt.
Discourse	Morphology	Word: wrong word - modal	incorrect	This is the wrong word. Distinguish between permission, possibility and obligation.	To enter the concert, you may (**must**) bring your ticket.
Grammar	Lexis	Word: Wrong word - Time/temporal adverbial	incorrect	This time-word does not fit the rest of the essay.	Use this when a student uses e.g. a word in the past-tense when the whole essay is written in the present tense.