

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/235772407>

# Development of an automated tool for detecting errors in Tenses

Article in *Gema Online Journal of Language Studies* · May 2012

---

CITATION

1

---

READS

165

5 authors, including:



[Siti Hamin Stapa](#)

National University of Malaysia

37 PUBLICATIONS 51 CITATIONS

[SEE PROFILE](#)



[Normah Binti Omar](#)

Universiti Teknologi MARA

142 PUBLICATIONS 204 CITATIONS

[SEE PROFILE](#)



[Mohd Juzaidin Ab Aziz](#)

National University of Malaysia

41 PUBLICATIONS 113 CITATIONS

[SEE PROFILE](#)



[Saadiyah Darus](#)

National University of Malaysia

49 PUBLICATIONS 172 CITATIONS

[SEE PROFILE](#)

## Development Of An Automated Tool For Detecting Errors In Tenses

*Tengku Nor Rizan Tengku Mohd Maasum*  
[tntm@ukm.my](mailto:tntm@ukm.my)

*School of Language Studies and Linguistics  
Faculty of Social Sciences and Humanities  
Universiti Kebangsaan Malaysia*

*Siti Hamin Stapa*  
[sitihami@ukm.my](mailto:sitihami@ukm.my)

*School of Language Studies and Linguistics  
Faculty of Social Sciences and Humanities  
Universiti Kebangsaan Malaysia*

*Nazlia Omar*  
[no@ftsm.ukm.my](mailto:no@ftsm.ukm.my)

*Faculty of Information Science and Technology  
Universiti Kebangsaan Malaysia*

*Mohd Juzaidin Ab Aziz*  
[din@ftsm.ukm.my](mailto:din@ftsm.ukm.my)

*Faculty of Information Science and Technology  
Universiti Kebangsaan Malaysia*

*Saadiah Darus*  
[adi@ukm.my](mailto:adi@ukm.my)

*School of Language Studies and Linguistics  
Faculty of Social Sciences and Humanities  
Universiti Kebangsaan Malaysia*

### Abstract

The rapid growth of computer technologies creates a plethora of ways in which technology can be integrated into one of the alternatives to facilitate essay marking. Automated essay marking systems developed from the late 1960s have attempted to prove that computers can evaluate essays as competently as human expert. Several computer-based essay marking (CBEM) systems have been developed to mark students' essays and they can be divided into semi-automated and automated systems. This paper illustrates the development of an Automated Tool for Detecting Errors in Tenses (ATDEiT™). The first phase analysed the errors found in 400 essays written by 112 English as second language (ESL) learners at tertiary level using Markin 3.1 software. The results showed that the most common errors were found in tenses. This finding led to the second phase of the research, which was the design of an automated marking tool. Consequently, the techniques and algorithm for error analysis marking tool for ESL

learners were developed. An initial testing was conducted to evaluate the results of the marking tool using 50 essays. Findings showed that ATDEiT™ achieved a high level (93.5%) of recall and an average level (78.8%) of precision. This proves that ATDEiT™ has the potential to be used as an automated tool for detecting errors in tenses for ESL learners.

**Keywords:** automated essay marking, automated tool, error analysis, grammatical errors, English as a second language.

## **Introduction**

English is an important language for communication worldwide. In Malaysia, English is a strong second language, especially in the context of education and international relations. The education system in Malaysia stipulates that the formal teaching and learning of English begins early from primary to tertiary level (Foo & Richards, 2004). The four language skills incorporated in the teaching and learning of English are listening, speaking, reading and writing. Therefore Malaysians learners should be competent in reading, speaking, writing and understanding the language.

One of the language skills which is important for second language (L2) learners is writing. According to Siti Hamin Stapa, Tengku Nor Rizan Tg Mohd Maasum, Rosniah Mustaffa and Saadiyah Darus (2008), learners should be able to write effectively to meet the standard of English language communication. Writing enables learners to document their understanding, voices and also to get the message across.

Research has demonstrated that ESL learners wrote differently compared to native English speakers. It was found that native English speakers wrote essays in a linear development in contrast to the ones written by non-native speakers. As early as 1966, Kaplan studied essays which were written by his foreign students. [Kaplan \(1966\)](#) discovered that Chinese, Thai and Korean students were more inclined to use an indirect approach while Russians demonstrated some degree of digression in writing which was considered to be quite excessive to an English speaker. [Connor \(1996\)](#) discovered that writers from different cultures had their own rhetorical tendencies and that ESL writers transferred such rhetorical patterns from their first language (L1) into their writing in English.

At the tertiary level of education, writing is regarded as a crucial skill to stimulate learning and critical thinking. Silva and Matsuda (2002) consider writing as one of the most difficult skills to master since it is a complex, recursive and creative process. In addition, it requires the development of an efficient and effective composing process. In writing, the composing process includes aspects such as planning, drafting, editing, changing and getting feedback. [Myles \(2002\)](#) adds that factors such as idea development, clarity, cohesion and grammar need to be taken into consideration in essay writing.

One of the crucial aspects raised by many researchers (Truscott, 1996; Ferris, 2002) in SLA is the presence of errors. These errors arise from two perspectives; the result of SLA and learning a TL. From these two perspectives, the written errors made by adult L2 learners are often quite different from those made by native speakers.

From the perspective of SLA, Ferris (2002) has pointed out several generalizations. Firstly, it takes a significant amount of time to acquire an L2 and even more when the learner is attempting to use the language for academic purposes. Secondly, depending on learner characteristics and age of first exposure to L2, some acquirers may never attain native-like control of various aspects of the L2. Thirdly, SLA occurs in stages. Phonetics/phonology (pronunciation), syntax (the construction of sentences), morphology (the internal structure of words), lexicon (vocabulary) and discourse (the communicative use that sentences are put to) may all represent separately occurring stages of acquisition. Fourthly, as learners go through the various stages of acquisition of different elements of L2, they will make errors reflective of their SLA processes. These errors may be caused by inappropriate transference of the L1 patterns and/or by incomplete knowledge of the L2.

Ferris (2002) notes that L2 writers need (a) a focus on different linguistic issues or error patterns than native speakers; (b) feedback or error correction that is tailored to their linguistic knowledge and experience, and (c) instruction that is sensitive to their unique linguistic deficits and needs for strategy training. Common ESL writing errors that are adapted from Ferris (2002) are presented in Table 1.

Table1: Common ESL writing errors (Ferris, 2002)

Type of Error	Example	
Morphological errors	Verb: - Tense - Form - Subject verb agreement	Noun: - Articles/Determiners - Noun endings (plural/possessive)
Lexical errors	Word choice, word form, informal usage, idiom error and pronoun error	
Syntactic errors	Sentence structure, run on and fragments	
Mechanical errors	Punctuation and spelling	

According to Ellis (1985), the level of SLA for each individual learner is not the same. As such, a tool to help ESL learners identify their writing errors would facilitate their essay writing so that these errors could be minimized. It would be more beneficial if the tool could be used by the learners themselves to check for errors before they actually submit their essays to the lecturers.

In addition, lecturers also need to understand that different learners may make distinct types of errors. Truscott (1996) mentions that different types of errors may need varying treatment in terms of error correction. For this reason, lecturers should understand the need to prioritize error feedback for individual learners. This could be identified by

looking at global errors versus local errors, frequent errors and structures elicited by the assignment that has been discussed in class.

From the point of view of learning a TL, the level of language a particular learner of a TL is operating falls into the following areas namely; substance, text and discourse (James, 1998). Substance, text and discourse relate to medium, usage and use respectively. According to James (1998), if a learner is operating on the phonological or the graphological substance system, that is spelling or pronunciation, she has produced an encoding or decoding error. However, if she is operating upon the lexico-grammatical system of the TL to produce or process text, these errors are referred to as composing or understanding errors. If she is operating on the discourse level, they are known as misformulation or misprocessing errors.

At tertiary level study, learners are required to submit numerous written tasks as part of their course requirements and assessments. Saadiyah Darus, Siti Hamin Stapa, Supyan Hussin and Koo Yew Lie (2000) state that learners prefer to receive feedback on essays regarding errors especially their type, coherence of the text and organization of ideas. Identifying the types of errors when giving feedback to the learners could be a time consuming process. The overwhelming and time consuming task of marking or grading written essays will be even more emphasized if teachers have large class sizes (Page, 1994).

With the expansion of the application of information technology in education, computer-based essay marking (CBEM) systems could be used to help lecturers with the essay grading tasks. With the help of an error marking tool, the lecturer's massive task in grading essays could be facilitated. In addition, learners could employ the marking tool themselves to check for language errors and edit their essays before submitting them to the lecturers.

The main purpose of this paper is to report the findings from two phases of this study. The first phase is the analysis of the common errors found in 400 essays written by 112 tertiary level ESL learners using Markin 3.1 software. In the second phase, the development of ATDEiT™ and its evaluation is illustrated. The screen shot of the prototype ATDEiT™ is available and accessible online at <http://research6977.com/markintool/index.php?m3=1&n3=2> (Appendix A).

## **Literature Review**

The rapid development of information technology in the field of education has led to the introduction of several CBEM systems to facilitate the marking of learners' essay. The systems can be sub-divided into semi-automated and automated systems (Saadiyah Darus, 1999).

Semi-automated CBEM systems require lecturers to read essays using the system and award marks. Some examples of the semi-automated systems are Methodical Assessment

of Reports by Computer (MARC) and Markin32. Automated marking systems, on the other hand, mark essays automatically by allocating a score as well as providing feedback within seconds. Some examples of the systems include Project Essay Grader, Intelligent Essay Assessor, e-rater, Criterion Online Writing Evaluation, Intellimetric Scholar, SEAR and Intelligent Essay Marking System (Saadiyah Darus, 2005).

A number of automated tools such as Expert System for the Teaching of English, Easy English, and A New Grammar Checker for English have been proposed to facilitate the evaluation of students' essays. Nevertheless, these tools are not specifically developed for Malaysian ESL learners. Since the written errors made by adult ESL learners are quite different from those by native English speakers, a tailor-made marking tool that could be utilised in ESL writing is deemed necessary. In addition, current researches in error analysis of Malaysian ESL students' writing are quite a number; Ang Leng Hong, Hajar Abdul Rahim, Tan Kim Hua and Khazriyati Salehuddin (2011), Siti Hamin Stapa and Mohd Mustafa Izahar (2010), Saadiyah Darus and Khor Hei Ching (2009), Saadiyah Darus and Kaladevi Subramaniam (2009), Marlyna Maros, Tan Kim Hua and Khazriyati Salehuddin (2007), Marlyna Maros, and Khazriyati Salehuddin and Tan Kim Hua (2007) to name a few.

### **Automated essay marking systems**

A number of researchers have attempted to develop automated essay marking systems. Ying and Chih (2003) have developed the Two Distance Applications Support English Distance Learning using a multimedia database and Internet technologies known as English multimedia corpus. English articles, dialogs, and videos are included in the system. Semantic query and "Link grammar" are applied to construct the English multimedia corpus system. The main function of this system is to query the English sentence pattern through keywords from the English multimedia corpus. The other function is to detect grammatical errors in written English. According to Ying and Chih (2003), the system not only teaches English grammar, but also allows teachers to understand the most frequent mistakes.

Rule-Based Style and Grammar Checker is another automated essay marking system which is developed by Naber (2003). The system has an open source style and grammar checker for the English language. The system generates a text and returns a list of possible errors. Each word in the text is assigned to its part-of-speech tag and each sentence is split into chunks, e.g. noun phrase. The parsing results are processed further by matching all the checker's predefined error rules.

Amoah, Lupiana and Ghemri (2006) use a Sentence Analyzer and Viewer for Detecting Grammatical Errors system, which aims to identify grammatical errors in a sentence. The tool uses the stages in natural language processing which access lexical and syntactic analysis to examine the learners' input. This system is able to detect agreement error in sentence and display helpful messages to learners.

For English Grammar Checking, Kumar and Nair (2007) use Artificial Immune System (AIS) based techniques as an approach for grammar checking. The motivation of the study comes from the human immune system, which identify external harmful entities from the self cells in the human body. It can detect any grammatical construct outside the corpus and identify it as an error.

The Grammar Diagnostic Expert System (GRADES) introduced by Fox and Bowden (2002) detects and explains grammatical errors of non-native speakers. It is designed to diagnose errors by native Japanese adults who are learning English as a second language. The system contains a small lexicon of words which can be expanded easily. GRADES's diagnosis is conducted through a classification process whereby an error category is considered and pattern matching rules are used. An explanation is also generated to help the user learn why the sentence is ungrammatical.

Although computers and artificial intelligence have been proposed as tools to facilitate the evaluation of students' essays, they are not specifically developed for Malaysian ESL learners. A marking tool which is specifically developed to analyze errors in Malaysian ESL writing is very much needed.

## **Methodology**

The study is divided into two phases. The first phase is aimed at identifying the common errors in students' essays. The second phase of the research is the development of techniques and algorithm for detecting and analysing these errors in students' essays.

### **Phase one: Identifying the common errors**

In phase one, essays of a group of undergraduates at Universiti Kebangsaan Malaysia were collected. These essays were part of students' written assignments submitted for the course that they had registered. 400 essays were collected and analysed. Since all students entering Malaysian universities come from similar backgrounds, these students formed a representative sample of other Malaysian students.

The research adopted the error classification scheme that was originally developed by [Lim \(1974\)](#), consisting of 13 types of errors. An addition of 4 other types of grammatical errors was found. (This was based on the research team's experience of teaching writing for more than 15 years). The final error classification scheme comprised 17 types of errors which were: tenses, articles, subject verb agreement, other agreement errors, infinitive, gerunds, pronouns, possessive and attributive structures, word order, incomplete structures, negative constructions, prepositions, mechanics, miscellaneous unclassifiable errors, word choice, word form, and the verb to be.

In analyzing the errors in students' essays, we employed the Markin 3.1 software (Holmes, 2007) in an attempt to facilitate the process. All the essays were word-processed and converted into rtf files so that they can be analysed using Markin 3.1.

Markin 3.1 is a semi-automated tool that allows instructors to mark written work submitted by students in the form of electronic documents. It provides five marking facilities; annotation buttons, add feedback, add comment, add a grade, and compile error statistics. This software is used to ensure that the classification of errors and statistical analysis of errors were made accurately and consistently. In doing so, the annotation buttons in the software were first customized based on the classification scheme of 17 types of errors. Three subject-matter experts were hired to evaluate the essays and classify the errors according to the classification scheme. These subject matter experts were experienced language teachers who taught English for at least ten years at the institutions of higher learning in Malaysia.

The results of the error analysis carried out semi-automatically by the subject-matter experts on 400 essays using Markin 3.1 software is presented in Table 2. It shows the frequency of occurrences and the average number of the types of errors. The average was calculated by dividing the number of errors with the total numbers of essays. Findings showed that there were six most common errors made by learners: tenses, prepositions, articles, word choice, mechanics and the verb 'to be'. The analysis showed that the most common errors made by these students when they wrote their essays was in tenses. Hence it was justified that error in tenses be given priority in the development of an automated marking tool.

Table 2: Types of errors in learners' essays

Errors	No. of Errors	Average
Tenses	1,595	3.99
Articles	1,204	3.01
Subject verb agreement	631	1.58
Other agreement errors	520	1.30
Infinitive	145	0.36
Gerunds	292	0.73
Pronouns	696	1.74
Possessive and attributive	232	0.58
Word order	194	0.49
Incomplete structures	253	0.63
Negative construction	56	0.14
Prepositions	1,468	3.67
Word choice	1,123	2.81
Word form	629	1.57
Mechanics	1,105	2.76
Verb to be	820	2.05
Miscellaneous unclassifiable errors	1,123	2.79

### **Phase two: Developing techniques and algorithm for detecting and analysing the most common errors**

The automated marking tool comprised two stages; the natural language stage and the logical stage. The natural language stage consisted of parsing the natural language input



(in the form of tagged essays) into the system whereby the system will read a plain text file containing sentences in English. A parser was used to classify the sentences and identify their part-of-speech (POS) tags. The parser used for this process was the CST's (Centre for Sprogetechnology) Part of Speech Tagger, which is a Memory-Based Shallow Parser. POS tags assigned each word in an input sentence to its suitable part of speech, such as noun, verb and determiner. The process was done to identify the word's syntactic category. Then the parsed text was 'fed' into the system to detect errors in tenses. Next, the types of tense errors were identified for the natural language input. Finally, the natural language input is produced and it can be used as feedback for students to revise and edit their essay.

Figure 1 illustrates the process in automated marking tool for ESL (Nazlia Omar, Nur Asma Razali & Saadiyah Darus, 2009, p. 477).

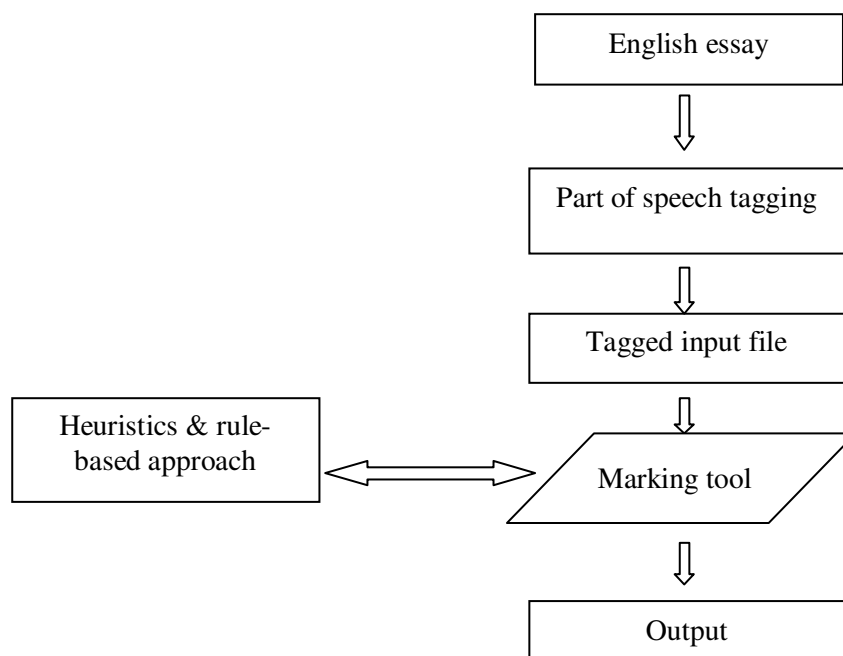


Figure 1: Process in automated marking tool for ESL writing

For the purpose of this research, a set of heuristics was developed based on a corpus of ESL essays. The heuristics enable the automated marking tool to detect grammatical errors in tenses. According to [Zanakis and Evans \(1981\)](#) heuristics represent an indefinite assumption, often guided by common sense to provide good but not necessarily optimal solutions to difficult problems, easily and quickly. The following example illustrates the application of heuristics which are based on syntax and context dependent meaning (Nazlia Omar, Nur Asma Mohd Razali & Saadiyah Darus, 2009, p. 478). In general, context-dependent meaning or properties are activated only by relevant contexts in which the word appears. Context-dependent properties are a source of semantic encoding variability.

For example,

1. Heuristic to identify error in Simple Future Tense:

If noun or pronoun is in the set of heuristic Future Tense followed by modal, check the verb after modal. If verb is tagged as VB, the sentence has no grammatical error.

Example sentence: Aziz will go to the market.

Tagged sentence: Aziz/NNP will/MD go/VB to/TO the/DT market/NN.

2. Heuristic to identify error in Present Progressive Tense:

If noun or pronoun is in the set of heuristic Present Progressive followed by 'is', check verb after 'is'. If verb is tagged as VBG, the sentence has no grammatical error.

Example sentence: Mary is reading the newspaper.

Tagged sentence: Mary/NNP is/VBZ reading/VBG the/DT newspaper/NN.

In the following section, the computer programming language process is described. The simple present tense is used as an example to illustrate the process. Initially, the algorithm which applies to the heuristics to detect errors in Simple Present Tense is constructed. Algorithm is a sequence of unambiguous instructions in a programming language. In this algorithm, 'check' refers firstly to the identification of a match, which meets the heuristic's (indefinite assumptions) condition. Heuristics are applied to any relevant words in the sentences that meet the heuristics' criteria. Figure 3 illustrates an extract from the algorithm for heuristics to detect error in Simple Present Tense.

```
1. FOR each sentence in parsed and tagged text file, DO;
2. FOR each sentence, DO;
  2.1 IF word is of type 'noun' DO;

  IF noun is in the set of heuristic Simple Present 1,
  IF noun is tagged as NNS, check the immediate verb after the noun.
  IF verb is tagged as VBP or VB, print the following sentence: The sentence has no
  grammatical error.
  IF noun is tagged as PRPplural, check the immediate verb after the noun.
  IF verb is tagged as VBP or VB, print the following sentence: The sentence has no
  grammatical error.

  IF noun is in the set of heuristic Simple Present 2,
  IF noun is tagged as NN, check the immediate verb after the noun.
  IF verb is tagged as VBZ, print the following sentence: The sentence has no
  grammatical error.
```

Figure 3: Extract from the algorithm for heuristics to detect error in Simple Present Tense  
(Nazlia Omar, Nur Asma Razali & Saadiyah Darus, 2009, p. 478)

## Results

Once the development of ATDEiT™ was completed, an initial testing was conducted to evaluate the marking tool. The initial testing produced three possible outcomes. The three possible outcomes were as follows: whether the errors are correctly detected, incorrectly detected and undetected. The approach used for the evaluation is based on Information Extraction systems using recall and precision (Jurafsky & Martin, 2000). Recall represents the percentage of all the possible correct answers produced by the system whereas precision is the percentage of answers that are correctly identified by the system. To some extent, the evaluation will reflect the accuracy of the system in obtaining the correct result. Nazlia Omar, Nur Asma Mohd Razali and Saadiyah Darus (2009) have provided a detailed description of the calculation procedure for calculating recall and precision of the marking tool.

Recall refers to the measure of the percentage of information that is actually found in the system. In this context, recall is the amount of correct information returned by the system. The correct information is then compared with those produced by human analysts or answer keys. The answer keys or Nkey are actually the amount of correct information plus the number of undetected ones. Thus, the following formula is used to calculate recall:

$$\text{Recall} = \frac{\text{Ncorrect}}{\text{Ncorrect} + \text{Nkey}}$$

Precision, on the other hand, is a measure of percentage of correctness of the information produced. It reflects the accuracy of the system in obtaining the correct result. The standard precision formula is as follows:

$$\text{Precision} = \frac{\text{Ncorrect}}{\text{Ncorrect} + \text{Nincorrect}}$$

ATDEiT™ was tested using a test dataset containing 50 essays. The analysis illustrated that the system achieved a high level of recall (93.5%). In terms of precision, the system scored an average level (78.8%). Thus, the results were consistent and encouraging.

## Conclusion

The aim of this article is to discuss a research in the development of a marking tool to detect error in tenses in order to assist ESL writers. The research was divided into two phases. In the first phase, the study investigated the common errors found in Malaysian ESL tertiary level students' essays while in the second phase, illustrated the design and development of an automated marking tool, ATDEiT™.

Based on the analysis of 400 essays, the results showed that the most common errors were found in tenses. This finding led to the second phase of the research, whereby an

automated marking tool for tenses was recommended. As a result, the techniques and algorithm for error analysis marking tool for ESL learners using heuristics and a rule-based approach were developed. Using a test dataset containing 50 essays, the automated marking tool was evaluated. The analysis showed that the system achieved a high level of recall and an average level of precision. The analysis based on the recall and precision scores showed that these heuristics can be refined so as to improve the accuracy of the result. It also indicates that heuristics and rule-based approach has potential use in developing an automated marking tool.

The development of an error analysis marking tool will be useful for ESL learners as it could assist learners in reviewing and proofreading their essays. For now, the automated tool system is only applicable to Active Simple Sentences. For future application, heuristics and rule-based approaches could be extended to other types of sentences, such as passives, complex, compound, and complex-compound sentences. It is hoped that an error marking tool such as this can be further developed and refined to assist ESL learners write better and more effective essays.

In the process of the development of an automated marking tool to detect errors made by ESL writers we found that there were other errors that need to be addressed. For example, subject-verb-agreement, prepositions, word choice, and articles. Further research should be conducted to develop marking tools for these errors in assisting ESL learners when they write in English.

These marking tools can either be used as stand-alone software or can also be incorporated into other systems. For example, it can be included in a computer assisted language learning package. Another option is to incorporate the tool into the Coursework Management System that has been developed by Abdullah Mohd Zin, Saadiyah Darus, Mohd Jan Nordin and Abdul Malik Md. Yusoff (2003).

### **Acknowledgement**

The research team is grateful for the research grant received from e-ScienceFund 01-01-02 SF0092. We would also like to thank the Centre for Sprogteknologi, Faculty of Humanities, University of Copenhagen.

### **References**

Abdullah Mohd Zin, Saadiyah Darus, Mohd Jan Nordin & Abdul Malik Md. Yusoff. (2003). Using a coursework management system in language teaching. *Teaching English with Technology*, 3(1). (Online) Retrieved 5 March 2012, from <http://www.tewtjournal.org/pastissues2003.htm>

- Amoah, P., Lupiana, R., & Ghemri, L. (2006). A sentence analyzer and viewer for detecting grammatical errors. *Texas Southern University Research Week*.
- Ang Leng Hong, Hajar Abdul Rahim, Tan Kim Hua, & Khazriyati Salehuddin. (2011). Collocations in Malaysian English learners' writing: A corpus-based error analysis, *3L: The Southeast Asian Journal of English Language Studies*, 17(Special Issue), 31-44.
- CST's Part-of-Speech Tagger. (Online) Retrieved 21 April 2012, from [http://www.cst.dk/online/pos\\_tagger/uk/index.html](http://www.cst.dk/online/pos_tagger/uk/index.html)
- Connor, U. (1996). Contrastive rhetoric: Cross-culture aspect of second language writing. New York: Cambridge University Press.
- Ellis, R. (1985). Understanding language acquisition. Oxford: Oxford University Press.
- Ferris, D. (2002) Treatment of error in second language student writing. Ann Arbor: The University of Michigan Press.
- Foo, B., & Richards, C. (2004). English in Malaysia. *RELC Journal*, 35(2), 229-240.
- Fox, R., & Bowden, M. (2002). Automated diagnosis of non-native English speaker's natural language, Proceedings of the 14th IEEE International Conference on Tools with Artificial Intelligent (ICTA'02) (pp. 301-306).
- Holmes, M. Markin 3.1 Release 2 Build 7. Creative Technology. 1996-2004. (Online) Retrieved 17 August 2007, from <http://www.cict.co.uk/software/markin/>
- James, C. (1998). Errors in language learning and use – Exploring error analysis. Harlow, Essex: Addison Wesley Longman Limited.
- Jurafsky, D., & Martin, J. H. (2000). Speech and language processing: An introduction to natural language processing, computational linguistics and speech recognition. Upper Saddle River: Prentice Hall, Inc.
- Kaplan, R. B. (1966). Cultural thought patterns in intercultural communication. Language Learning, 16(1 & 2), 1-20.
- Kumar, A., & Nair, S. (2007). An artificial immune system based approach for English grammar checking, ICARIS 2007, Lecture Notes in Computer Science, 4628 (pp. 348-357).
- Lim, H. P. (1974). An error analysis of English compositions written by Malaysian speaking high school students. Unpublished M.A TESL thesis. University of California.

- Marlyna Maros, Tan Kim Hua, & Khazriyati Salehuddin. (2007). Interference in learning English: grammatical errors in English essay writing among rural Malay secondary school students in Malaysia, *Jurnal e-Bangi*, 2(2), 1-15.
- Marlyna Maros, Khazriyati Salehuddin, & Tan Kim Hua. (2007). Negative transfer in the learning of English: A case of grammatical errors in the writings of Malay secondary school students in Malaysia. *Journal of Language and Linguistics*, 25(2), 18-33.
- Myles, J. (2002). Second language writing and research: The writing process and error analysis in student texts. *TESL-EJ Journal*. 6(2), A-1.
- Naber, D. (2003). *A rule based style and grammar checker*. Diplomarbeit Thesis, Universitat Bielefeld.
- Nazlia Omar, Nur Asma Razali & Saadiyah Darus (2009). Automated grammar checking of tenses for ESL writing. *Lecture Notes in Computer Science (LNCS) 5589*, 475-482.
- Page, E. B. (1994). Computer grading of student prose - Using modern concept and software. *Journal of Experimental Education*. 62,127-142.
- Saadiyah Darus (1999). A prospect of automatic essay marking. Paper presented at the International Seminar on Language in the Global Context: Implications for the Language Classroom, RELC. Singapore.
- Saadiyah Darus. (2005). *A framework of a CBEM system for ESL writing at institutions of higher learning in Malaysia*. Unpublished Ph.D thesis. Universiti Kebangsaan Malaysia.
- Saadiyah Darus & Kaladevi Subramaniam. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European Journal of Social Sciences*, 8(3), 483-495.
- Saadiyah Darus & Khor Hei Ching. (2009). Common errors in written English Essays of Form One Chinese students: A Case Study. *European Journal of Social Sciences*, 10(2), 242-253.
- Saadiyah Darus, Siti Hamin Stapa, Supyan Hussin & Koo Yew Lie. (2000). A survey of computer-based essay marking (CBEM) systems. *Proceedings of the International Conference on Education and ICT in the New Millennium*, Universiti Putra Malaysia: Faculty of Educational Studies (pp. 519-529).
- Silva, T., & Matsuda, P. K. (2002). Writing. In N. Schmit (Ed.), *An introduction to applied linguistics*. London: Oxford University Press.

Siti Hamin Stapa & Mohd Mustafa Izahar. (2010). Analysis of errors in subject-verb agreement among Malaysian ESL learners. *3L: Language, Linguistics, Literature: The Southeast Asian Journal of English Language Studies*, 16 (1), 21-26.

Siti Hamin Stapa, Tengku Nor Rizan Tg Mohd Maasum, Rosniah Mustaffa & Saadiyah Darus (2008). Workplace written literacy and its effect on the curriculum. *GEMA Online™ Journal of Language Studies*, 8(1), 23-33.

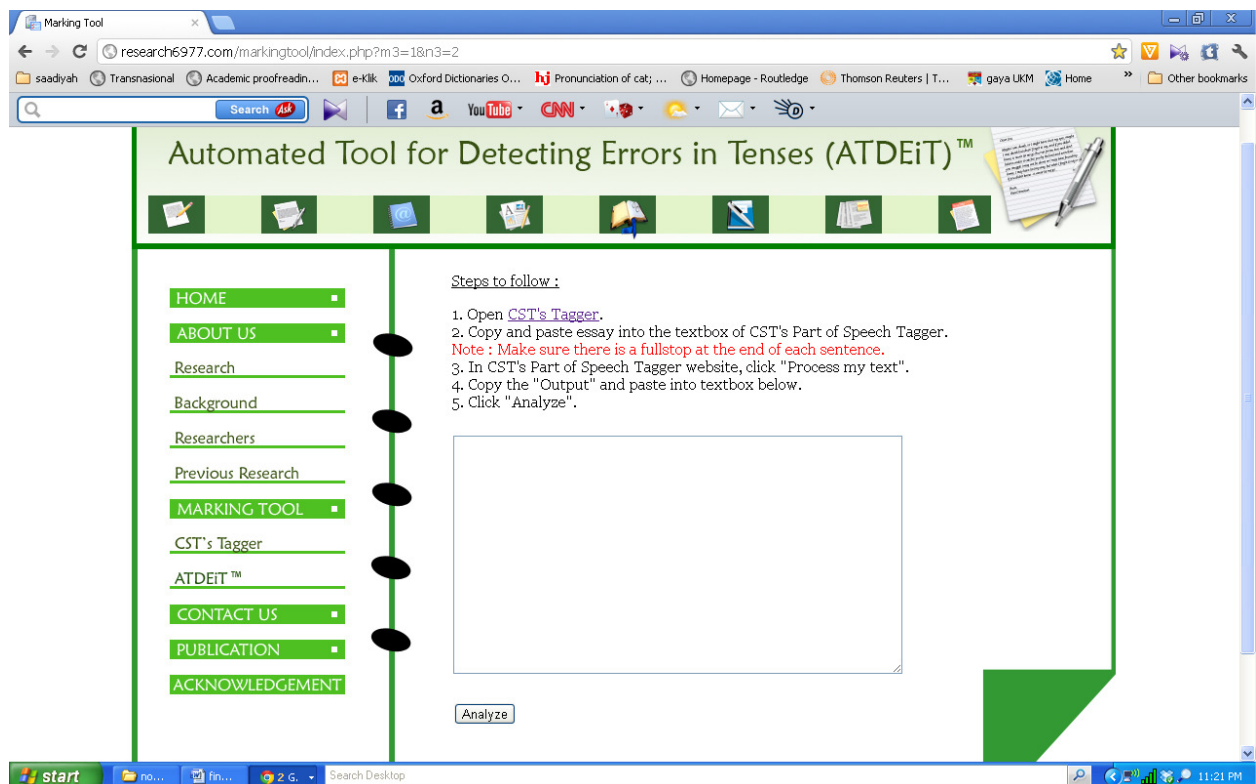
Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.

Ying, H.W., & Chih, H. L. (2003). Two intelligent applications support English distance learning, *Proceedings of the 23rd International Conference on Distributed Computing Systems Workshop*. IEEE Computer Society, Washington, D. C., USA.

Zanakis, S. H. & Evans, J. R. (1981). Heuristics 'optimization': Why, when and how to use it. *Interfaces*. 11(5), 84-91.

## Appendix A

### Screenshot of ATDEiT™



### **About the authors**

Tengku Nor Rizan Tengku Mohd Maasum (Ph.D) is an Associate Professor of Applied Linguistics at the School of Language Studies and Linguistics, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia. Her research interests include cooperative learning, multiculturalism, workplace and IT literacy.

Siti Hamin Stapa (Ph.D) is an Associate Professor of Applied Linguistics at the School of Language Studies and Linguistics, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia. She has conducted research in the area of Language and ICT, Language and Culture, IT Literacy and Gender Studies.

Nazlia Omar (Ph.D) is an Associate Professor at the School of Computer Science, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. She holds her Ph.D in Computer Science from the University of Ulster, U.K. Her main research interest is in the area of Natural Language Processing, Database and Information Systems.

Mohd Juzaidin Ab Aziz (Ph.D) is an Associate Professor at the School of Computer Science. He is currently the Head of Computer Science Program. His research interests are Natural Language Processing and the Science of Programming.

Saadiah Darus (Ph.D) is an Associate Professor at the School of Language Studies and Linguistics, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia. She was the lead researcher of e-ScienceFund 01-01-02 SF0092. Her research interest is computer-aided writing specialising in writing at the workplace. She has published a number of journal articles both locally and internationally.