國 立 交 通 大 學

理學院網路學習學程

碩 士 論 文

以 CLEC 語料庫為基準來設計實作一個英語作文錯誤檢查系統

The Design and Implementation of an Error-Checking System for English Composition Based on Chinese Learner English Corpus

研 究 生：顏佐宇

指導教授：陳登吉　教授

中 華 民 國 九 十 六 年 六 月

以 CLEC 語料庫為基準來設計實作一個英語作文錯誤檢查系統
The Design and Implementation of an Error-Checking System for English
Composition Based on Chinese Learner English Corpus

研 究 生：顏佐宇　　　　　Student：Tzuo-Yu Yan

指導教授：陳登吉　　　　　Advisor：Deng-Jyi Chen

國 立 交 通 大 學
理學院網路學習學程
碩 士 論 文

A Thesis

Submitted to Degree Program of E-Learning

College of Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Degree Program of E-Learning
June 2007
Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 六 年 六 月

以CLEC語料庫為基準來設計實作一個英語作文錯誤檢查系統

學生：顏佐宇　　　　　　　　　　指導教授：陳登吉　博士

國立交通大學

理學院網路學習碩士在職專班

摘　　要

隨著英語課程的進入小學，以及英語檢定的普及，學習英語已蔚為風行。因應這樣的發展，學生們已開始在小學接受英語教學，以及相關的測驗，而在這種情形之下，在外補習或是自學的人數也隨之增長；但是，以學習英語作文為例，當學習者在完成一篇作文的時候，只能尋求人工的方式批改，不但耗費時間，評分標準也可能不一。不同於有著固定答案的選擇題或是非題，早就以電腦閱卷的方式取代人工批閱，即便是有少數系統可以批改，正確率也有待提升。

因此，本研究試著去提昇實驗室中的英語作文的錯誤檢查系統之正確率，希冀以CLEC語料庫中分類清晰，且具代表性的錯誤類型為驗證藍本，撰寫相對應的作文錯誤檢查模組。期望以一致的角度，先將作文當中的拼字、文法錯誤，或是文不對題的情形清楚的標示出來，以提供閱卷者在閱卷時評分的參考。

同時，由於英語對我們華人來說，畢竟屬於外來的語言，其中的文法和語態與中文也不盡相同，導致學習上的困難。因為如此，我們希望建立起一套系統，除了可提供學習者做相關的事前練習之外，也可作為閱卷者評分的參考；更重要的是，系統以CLEC語料庫為驗證背景，整併及撰寫新的文法檢查模組，使錯誤類型之分類更為完整，並提高系統之錯誤檢查率。此外，也可讓受試者在不需要他人協助的情況之下，就可以自行操作練習，並不會受到外在環境的條件限制，也充分的運用了資訊設備，真正享受網路所帶給我們的便利性。

關鍵字：語料庫、文法檢查、英語、作文、學習者。

The Design and Implementation of an Error-Checking System for English Composition Based on Chinese Learner English Corpus

Student：Tzuo-Yu Yan   Advisor：Dr. Deng-Jyi Chen

Degree Program of E-Learning

College of Science

National Chiao Tung University

Abstract

As the elementary school starts teaching English and the popularity of GEPT (General English Proficiency Test), many people consider learning English indispensable. Following this trend, many students start to receive English courses and take English tests. Therefore, the number of students who go to English institutions or appeal to self-study has increased as well. However, taking learning English composition as an example, when the learners finish their compositions, they must seek someone to check their works. This is a time-consuming process and, furthermore, the grading standard may vary. Multiple-choice and true and false questions have fixed answers that can be checked through Optical Mark Recognition, replacing the manual grading method. Even though English composition can be graded by a few systems, its accuracy may be varied due to the difference of learning background.

As a result, this research intends to improve the error checking system for Chinese learner's English composition. The explicitly categorized and the representative error types in CLEC will be the basis to write the error checking system for English composition. By using the coordinate approaches, the misspellings, grammatical errors, and the irrelevance sentences can be marked clearly and served as references for the graders.

Since English is still foreign to Chinese, the differences among voice and tenses have caused difficulties in learning. Thus, we must design a system not only providing the learners to do relevant practice beforehand but also functioning as references for the graders. More importantly, our system is based upon CLEC and integrates many grammar checking modules to cover the error types as many as we can. For a learner, it can be used

to his own English composition conveniently. For a grader, it can be used to uncover all possible errors before they do the final semantic checks. Thus, the grading basis for all testers' English compositions is the same.

Numeric English composition samples (list in appendix) are used to demonstrate the feasibility and applicability of the proposed system.

Keywords: corpus、grammar check、English、composition、learner

誌　　謝

目　　　錄

表　目　錄

圖　　目　　錄

一、 緒論

1.1 研究背景與動機

由於國際間的貿易、經濟、社會文化等往來頻繁，英語已經日漸成為國家之間溝通往來的溝通工具；學習英語已經成為刻不容緩的一件事情，具備基礎的英語素養，也已成為地球村每一份子的基本能力。

因此教育部為了提升國民之英語能力，強化我國國際競爭力，便積極著手進行相關的英語教學，更於九十四學年度宣布，英語教學提前至小學三年級開始實施[17]，財政充裕的縣市(如台北縣、市)更於低年級即開始實施英語教學，相信不久的將來，能夠見到此政策之成效。

「全民英語能力分級檢定測驗」即是教育部推動學習英語學習的要項，教育部核准財團法人語言訓練測驗中心辦理該測驗，自民國 89 年開辦以來，受到社會各界的重視與採認，至 95 年總報考人數已逾 220 萬人次[20]。

在目前這樣的情形下，許多學習者為了增強自己的英語能力，平常就勤加練習，以其中的英語作文為例，當學習者在進行相關的練習之後，通常都需要以人工的方式來閱卷，但是人工閱卷較為耗時費力，且無法有著相同的閱卷標準。再者，現有之電腦閱卷系統大多只能對具有標準答案的填充題、選擇題進行批閱，少數能夠批閱作文的系統正確率也有待提升，更也有些系統是源自於歐美語系的國家，文法檢查方式與字彙運用都不甚相同。因此我們希望能夠實作出符合華人語言使用習慣，且正確率較高的英語作文錯誤檢查系統，能夠提供給學習者練習使用。

所以，一個完整的學習者語料庫-Chinese Learner English Corpus[19]對我們來說就很重要了，此語料庫中針對學習者作文當中的錯誤類型加以分類，列出大量的使用者錯誤的句型，讓我們藉助此語料庫當中的錯誤類型作為系統之驗證藍本，提升目前英語答題驗證系統的錯誤檢查率；同時我們希望了解目前系統的錯誤檢查比率是多少，並且針對其文法檢查不足的向度，設計文法模組，同時依循語料庫中分類完整且清晰的錯誤類型，整併文法模組，以達到我們提升正確率之目的。

1.2 研究目的與範圍

本研究主要希望達到的目的如下：

1. 基於 CLEC 語料庫的錯誤類型為文法模組驗證基準[19]，建構一個符合華人使用習慣的英語作文錯誤檢查系統。

2. 系統可以將作文之文法錯誤、拼字錯誤、文不對題的情況標示出來，以立即回饋的方式反應給學習者。

3. 以 CLEC 語料庫的文法分類[19]，設計系統文法模組功能，並建立標準一致的作文錯誤檢查標準。。

4. 提供學習者一個可以自我學習的作文練習環境。

此研究的範圍主要是在英語測驗的作文部分，由於本實驗欲研究之作文錯誤檢查系統，在文法模組的設計上，需要完整而清晰的分類；同時，也需要大量的英語作文測試樣本，以驗證文法模組，這也是 CLEC 對我們來說最重要的目的。

1.3 研究方法與步驟

我們欲先確認系統可檢查出來的錯誤數量及類型，並且運用 CLEC 語料庫加以驗證，了解程式可找出錯誤的比率，以設計文法模組，同時提升系統可找出錯誤的比率。我們使用的方法如下：

1. 了解系統所能檢查出的文法、拼字錯誤類型以及文不對題的檢查方式。
2. 分析系統模組可檢查出 CLEC 語料庫的錯誤種類及數量。
3. 以 CLEC 語料庫分類方式為樣本，設計文法模組。
4. 以系統之文法模組，與 CLEC 語料庫相互驗證。
5. 統計與 CLEC 語料庫相互驗證之後的錯誤種類及數量。

1.4　章節概要

本研究主要分成六個章節來介紹，依序如下：

第一章　敘明本研究的動機、目的、以及研究的方法及步驟。

第二章　系統技術及相關文獻探討

第三章　系統架構與系統模組設計

第四章　系統效能與比較。

第五章　系統畫面展示與限制說明。

第六章　說明本研究之結果與相關未來系統可以延伸開發之功能。

二、 系統技術探討

　　本研究欲研究的部份，是建構一個以 CLEC 語料庫為基準的作文錯誤檢查系統，此系統中擁有的相關檢查功能為：拼字檢查、文法檢查、同義字檢查。

　　以下將逐一介紹拼字檢查、文法檢查、同義字檢查功能所使用的相關技術。

2.1　GNU Aspell

　　在作文錯誤檢查當中，拼字的正確與否，會影響到文法的判別；因此拼字檢查雖然容易，但卻是系統中不可或缺的部份，影響的層面也可見一斑。

　　系統使用的軟體是 GNU Aspell[22]，GNU Aspell 是一個被設計用來替換 Ispell[6]的拼字檢查程式，原本 Ispell 是一個支援很多歐洲語系的 Unix 程式。它可以使用在 Emacs[3]編輯器下，但是大部分的人發現，在發現錯誤後的建議字上並不十分理想，所以 Kevin Atkinson 才開始致力研究於 Aspell[4]。

　　Aspell 的使用方式有兩種：一種是當作函式庫，被程式呼叫使用；另一種則是作為一個獨立的拼字檢查程式。而它主要比其他拼字檢查程式優越之處，在於兩點[2]：

1. 它可以建議拼錯的字該如何改正，也就是列出一些可能的建議字在旁邊，以提供參考之用。

2. Aspell 優於 Ispell,的主要一點，在於 Aspell 不需要使用特定的辭典就可以檢查用 UTF-8 編碼的文件。

3. 它還擁有一些 Ispell 所沒有的功能，例如：可使用多本字典和當其他使用者在使用時，每個使用者都能擁有一部個人字典，而不會互相影響彼此。

　　很多的軟體或是系統目前都使用 Aspell 當作拼字錯誤檢查程式，如知名的瀏覽器 Opera[14]或是作業系統 Linux 或 Unix，只要安裝至少一部 Aspell 可使用的辭典，馬上都可以用以上的方式來操作 Aspell，把他當作單字的錯誤檢查程式。

圖 2-1 所示的是 Aspell 檢查錯誤單字後的畫面[22]。
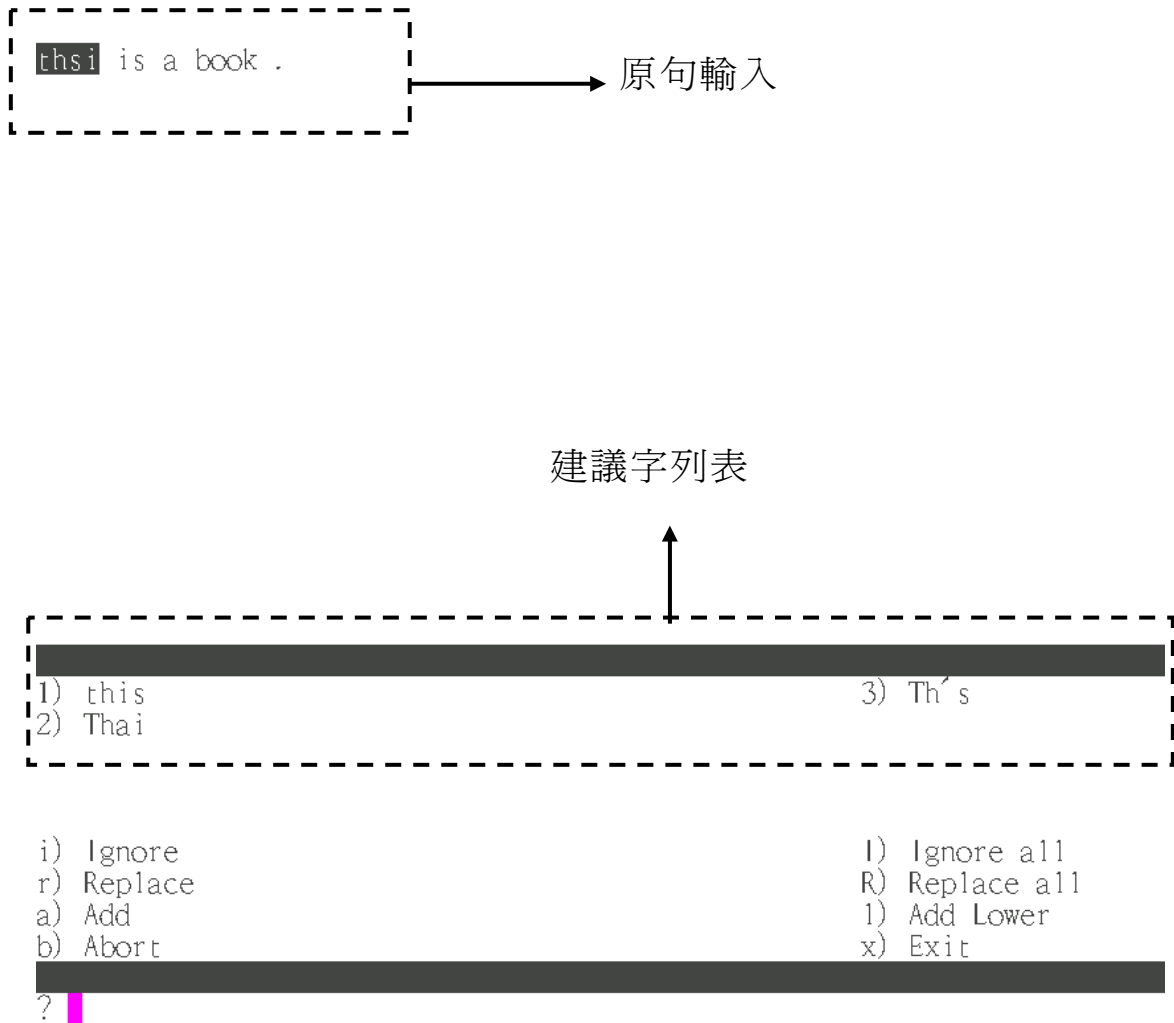
圖 2-1　　系統中呼叫 Aspell 檢查錯誤

圖 2-1 是系統利用程式呼叫 Aspell 的操作畫面，可以看到輸入的句子是「thsi is a book」，經過檢查之後，發現錯誤的單字是「thsi」，並且將他反白起來，表示這部份發生錯誤。另外，對於發生錯誤的單字，給予一些可能的建議字，如：this、Thai…等，提供使用者參考。

2.2　Apple Pie Parser

在進行文法檢查之前，我們必須要知道句子的結構與單字的詞性，但結構與詞性必須仰賴 parser 才能將其解譯出來，解譯出來之後，提供詞性與句構給文法模組作文法檢查，這也是本研究主要著重的部份；目前的語料庫研究中，著重的部份首推 parser 的研究與詞性的標記（POS(part of speech) tagging），而目前系統所使用的 Apple Pie Parser，就是基於 Pennsylvania 大學所開發的 Penn Treebank 中的詞性標記[12]。下圖即為 Penn Treebank 的詞性標記表 [11]。

1.	CC	Coordinating conjunction	25.	TO	*to*
2.	CD	Cardinal number	26.	UH	Interjection
3.	DT	Determiner	27.	VB	Verb, base form
4.	EX	Existential *there*	28.	VBD	Verb, past tense
5.	FW	Foreign word	29.	VBG	Verb, gerund/present participle
6.	IN	Preposition/subord. conjunction	30.	VBN	Verb, past participle
7.	JJ	Adjective	31.	VBP	Verb, non-3rd ps. sing. present
8.	JJR	Adjective, comparative	32.	VBZ	Verb, 3rd ps. sing. present
9.	JJS	Adjective, superlative	33.	WDT	*wh*-determiner
10.	LS	List item marker	34.	WP	*wh*-pronoun
11.	MD	Modal	35.	WP$	Possessive *wh*-pronoun
12.	NN	Noun, singular or mass	36.	WRB	*wh*-adverb
13.	NNS	Noun, plural	37.	#	Pound sign
14.	NNP	Proper noun, singular	38.	$	Dollar sign
15.	NNPS	Proper noun, plural	39.	.	Sentence-final punctuation
16.	PDT	Predeterminer	40.	,	Comma
17.	POS	Possessive ending	41.	:	Colon, semi-colon
18.	PRP	Personal pronoun	42.	(	Left bracket character
19.	PP$	Possessive pronoun	43.	)	Right bracket character
20.	RB	Adverb	44.	"	Straight double quote
21.	RBR	Adverb, comparative	45.	'	Left open single quote
22.	RBS	Adverb, superlative	46.	"	Left open double quote
23.	RP	Particle	47.	'	Right close single quote
24.	SYM	Symbol (mathematical or scientific)	48.	"	Right close double quote

圖 2-2　　Penn Treebank 的詞性標記表

Apple Pie Parser(APP)是一個利用 bottom-up probabilistic chart 統計方式，以及使用 best-first search 演算法，計算最佳得分之後，產生 parse tree。也就是說 APP 是以 Penn Treebank 中所統計的文法規則以及字典，利用演算法所找出來的最佳解析詞性。下圖為 Apple Pie Parser 執行畫面。

```
Apple Pie Parser (5.9)    April.4.1997  S.Sekine(NYU)
-----
Type `*help' for help, or type your sentence after prompt
Now loading dictionary and grammar...
-----
>> Mary is a beautiful girl.
(S (NPL Mary) (VP is (NPL a beautiful girl)) -PERIOD-)
>>
```

圖 2-3   Apple Pie Parser 執行畫面

圖 2-3 中輸入的句子為「Mary is a beautiful girl」，經過 parser 解析之後，得到的結果是(S(NPL Mary)(VP is(NPL a beautiful girl))-PERIOD-)，也就是將輸入的句子用括號的形式輸出，每個括號中前面的字為詞性標記，如：NPL、VP…等，括號後方的為原輸入單字。

要注意的是，無論句子中的拼字以及文法是否正確，APP 都會將句子做上標記，再交由系統的文法檢查模組進行檢驗的。

2.2.1 Apple Pie Parser Parser

在經過 Apple Pie Parser 對於輸入的英文句子進行標記後，接著進行的就是 Apple Pie Parser Parser(APPP)的標記，也就是句子結構性的標記；為了系統可以清楚的辨別各個詞性與句構,我們將剛才 APP 輸出的結果,再經過 APPP 標記一次，把我們想要的資料結構加註進去。

我們以 APP 切出 token 後，以 lex&yacc 的方式，也就是 APPP 中輸出成具有 XML 語法標記的樹狀結構標記句型[22]：

1. LEAF：tree 的末端，格式為(tag　單字)

   例：(NN book)

2. NODE：可以是 LEAF 或(tag　NODE_LIST)

   例：(NPL(DT This)(NN book))

   　　(tag (　NODE_LIST　))

3. NODE_LIST：NODE 或 NODE_LIST+NODE

   例：(DT This)(NN book)

4. TREE：最頂端的結構。

   例：(S)

我們也可以藉由樹狀圖的表示方式來說明上列的結構，以及相關的標記方式，由下兩圖可以更清楚的呈現此部分定義的結構。

```
01 <Tree>
02 <CNode tag="S" > （簡單敘述子句）
03   <CNode tag="NPL" > （最小名詞片語）
04    <CLeaf tag="NNP" word="Mary"  index="0" /> （單數專有名詞）
05   </CNode>
06   <CNode tag="VP" > （動詞片語）
07    <CLeaf tag="VBZ" word="is"  index="1" /> （第三人稱單數動詞）
08    <CNode tag="NPL" > （最小名詞片語）
09       <CLeaf tag="DT" word="a"  index="2" /> （限定詞）
10       <CLeaf tag="JJ" word="beautiful"  index="3" /> （形容詞）
11       <CLeaf tag="NN" word="girl"  index="4" /> （單數名詞）
12    </CNode>
13   </CNode>
14   <CLeaf tag="-PERIOD-" word="."  index="5" /> （句子的結束）
15 </CNode>
16 </Tree>
```

圖 2-4   Apple Pie Parser Parser 結果輸出
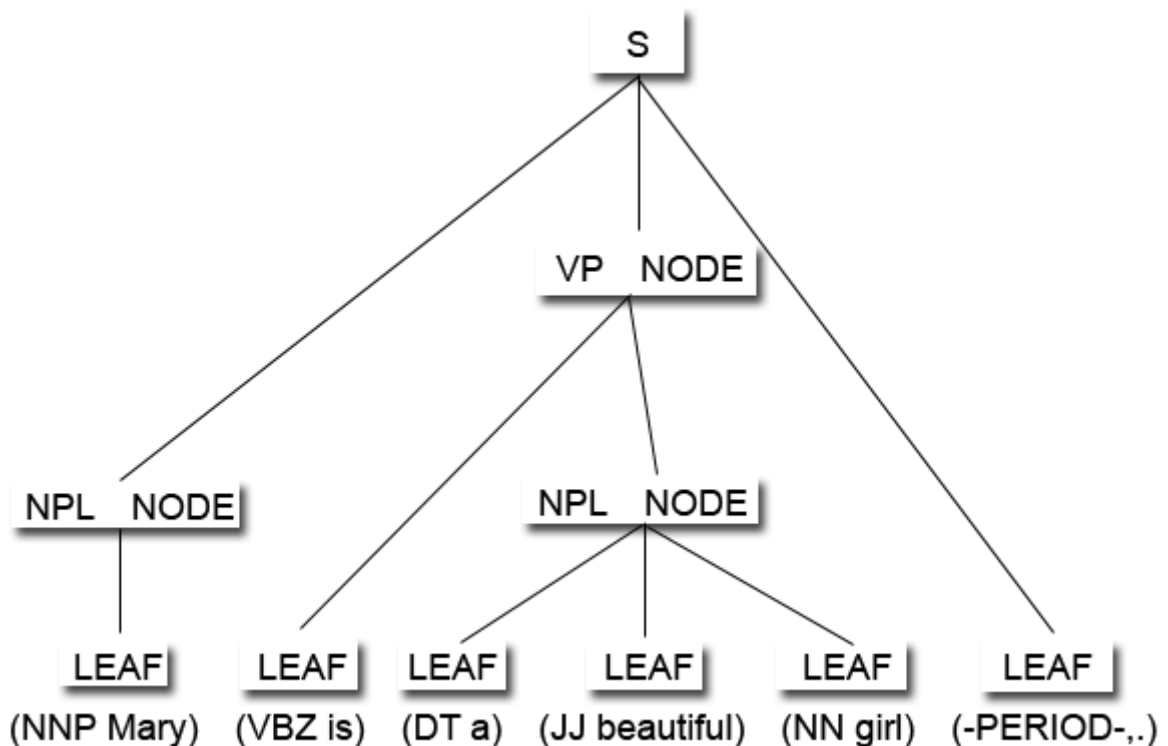
圖 2-5   Apple Pie Parser Parser 結構

8

上述兩圖即可清楚的呈現，句子中的每個字都是一個 LEAF，很多的 LEAF 可以合成 NODE，這些 NODE 也可再跟其餘的 LEAF 或 NODE 合成更上層的 NODE，最後所有的 NODE 合在一起而成為一顆完整的樹(TREE)。

2.3　WordNet

當一篇英語作文在系統的拼字檢查、文法檢查均正確的情況之下，我們就試著檢查出其內容是否有語義上的問題；但語義檢查的範圍十分廣泛，由於目前需檢查的部份為英語作文，因此我們僅檢查該篇作文是否文不對題。

系統在這部份運用到同義字的對照來進行檢查語義的部份，我們藉由 WordNet[15]來達成我們的要求。WordNet 是美國普林斯頓大學的認知科學實驗室，其中 George A. Miller 所帶領研究的英語辭典，內容包含動詞、名詞、形容詞和副詞，以上的詞類組合被組合而成所謂的同義詞組(synsets)，每種詞組都代表一個概念。這些詞組彼此之間再藉著語意的概念與詞彙的關係互相聯繫而互相連結；如此一來，一個有意義的網絡就可以使用瀏覽器在網路上瀏覽了。

圖 2-6　　WordNet 同義字查詢

9

上圖是 WordNet 查詢關鍵字的 Web 介面，在查詢同義字時，系統就會去查詢該字在各種詞性有哪些同義字，以查詢「display」為例，他在名詞方面的同義字有 show、presentation...等，在動詞方面的同義字則有 expose、exhibit...等。

WordNet 可以免費下載與公開使用，而且 WordNet 的結構讓它成為一個研究計算機語言學和自然語言處理的好用工具。多年以來，也有著許多人投注在 WordNet 的發展上[15]。

WordNet 在系統中運作的方式，是以模組來呼叫，也就是說，根據 APPP 所解析出來的詞性標記，再將他放進模組中檢驗。

2.4　語料庫

什麼是語料庫呢？黃希敏教授提到[23][24]：其實自古以來就有「語料語言學」(Corpus Linguistics)的存在，只是過去沒有這個名稱。單從「語料」的定義來看，這個字翻譯自拉丁文 corpus（複數為 corpora），意思是「體」，而從語言學的角度來看，就是「整體性的語言資料」，意思是指人類口語或文字紀錄的彙編。

以上是語料庫之定義，同時黃希敏教授的論文裡也將語料庫依照內容，分為以下四類[23][24]：

1. 單一語料庫：

   為了某種特定的研究所做的語料收集，過程需要非常嚴謹，收入該語料庫的每個字句都必須具有代表性。例如：文言文語料庫。此語料庫中所包含的內容都必須是文言文，其中可能包含文言文的句構、詞的用法...等等，研究者可以從語料庫中分析出一些資訊，如：文言文中不同作者的風格，不同的朝代文風有何改變，諸如此類的資訊都可找到。

2. 複合語料庫：

   此資料庫是將兩種或數種語料「集大成」，不過在收集時需要在類型之間加以區隔，讓研究者自行選取某單一類型來做研究或某幾個跨類型語料來做比較。例如：多語（multilingual）語料庫：包括兩種或兩種以上語言，提供翻譯研究、語言對比研究（contrastive studies），進而可

以協助電腦翻譯與語言教學法的研發。

3. 開放性語料庫：

國內中研院製作的現代漢語開放性語料庫「中央研究院平衡語料庫」
(Sinica Corpus)，可不斷擴張，提供從事語言文字分析的人士使用。

4. 學習者語料庫：

廣泛蒐集學習者的語料，提供外語教學界研究。

對於系統而言，文法錯誤的檢查模組必須以大量的錯誤樣本測試其正確性
及完備性。在這樣的情況之下，上述之語料庫就是一個最好的樣本，可以提供
給我們系統作為驗證測試之用，以便在撰寫程式碼的時候有所遵循。

2.4.1 學習者語料庫

李珮瑛曾在文章中提到[18]：近年來，語料庫的發展已經為語言學習者提
供了寶貴的學習資源，其中學習者語料庫也是如此，而翻閱其歷史，最早的學
習者語料庫是八零年代末期所建立的朗曼學習者語料庫(Longman Learners'
Corpus)[10]。九零年代中期，比利時魯汶大學 Centre for English Corpus
Linguistics[1]的 Sylvaine Granger 建立了國際學習者英語語料庫(International
Corpus of Learner English, ICLE)[5]，該語料庫是一廣泛國際合作的計畫，現
存有超過二百萬詞，存有十四種不同母語背景的英文學習者語料。

而 Leech[9]也曾在其論文中提到，建立學習者語料庫的目的是：

1. 比較學習者語料庫 LC（Learner Corpus）和欲學習語言為母語的語料
   庫 ECNS（English Corpus of Native Speakers），比較其中使用過多或
   過少的語詞。

2. 學習者的母語對於在使用欲學習語言時的影響程度。

3. 學習者使用新語言時，在哪些方面能夠達到所欲表達的目標，在哪些方
   面無法達成。

4. 學習者在哪些方面無法達到所欲表達的目標，而需要幫助。

對我們的研究來說，需要找到一個適當的學習者語料庫，且語料內容的範圍限制在中國學習者之英語作文，將是最適合本研究的學習者語料庫，關於這部份，我們在下一節將會有更詳細的說明。

2.4.2 Chinese Learner English Corpus

在前面我們已經提到，語料庫可說是一種資料庫，運用了資訊科技，將大量和語言相關的資料，使用具流通性的 XML 標籤語言，將之標記之後，存到我們的電腦當中。

目前語料庫種類繁多，如果我們僅針對學習者語料庫，而且將範圍限制在英語學習者語料庫來看，具代表性的大概有以下幾個[21]：

表 2-1　英語學習者語料庫列表

語料庫名稱	詞數	研究單位
ICLE（International Corpus of Learner English）[5]	200 萬	比利時 Louvain-La-Neuve 大學
JEFLL (Japanese EFL Learner) Corpus[8]	50 萬	日本明海大學
中國學習者語料庫（CLEC）[19]	100 萬	廣東外貿大學、上海交通大學
大學英語學習者口語語料庫（COLSEC）[25]	5 萬	上海交通大學
香港科技大學學習者語料庫（HKUST Learner Corpus）[21]	2500 萬	香港科技大學
中國英語專業語料庫（CEME)[21]	148 萬	南京大學
中國英語學習者口語語料庫（SECCL）[21]	100 萬	南京大學
國際外語學習者英語口語語料庫中國部分（LINSEI-China）[21]	10 萬	華南師大
碩士寫作語料庫（MWC）[21]	12 萬	華中科技大學

除此之外，尚有許多學習者英語語料庫，如：Longman Corpus of Learner's English (LCLE)、美國的英語學習者語料庫—MELD、英國的商業性學習者語料

庫—CLC 和 LLC...等。

但是我們在選擇語料庫,並希望以他作為作文錯誤檢查樣本時,我們評估的項目包括:

1. 語料庫製作背景:
   希望是華人所製作的,其中語料也較具代表性。
2. 內容之關聯性:
   由於我們的系統的範圍限制在英語作文錯誤檢查系統,因此語料內容也須符合我們的需求。
3. 詞數須適中:
   詞數過多或過少都會讓我們分類不易。

綜合以上條件,上表中,Chinese Learner English Corpus 可說是最適合我們需求的英語學習者語料庫,因為它的內容就是作文,並非其他口語、寫作或是其他目的之語料庫,且詞數適中,較容易分析,因此我們選定 CLEC 為作文的錯誤檢查驗證語料庫。

Chinese Learner English Corpus(簡稱 CLEC 語料庫)為上海交通大學楊惠中、廣東外貿大學桂詩春兩位教授合力編著,內容記載著各種統計資料和列表,而語料庫的詞數為一百萬詞,也許大家會認為以今天的技術來說,建立個幾百、幾千萬詞並非難事,但是語料庫必須對語言失誤按照統一的語言失誤表進行標記,而失誤標記是頗耗費人力與時間的。如果由少數人來做,則易於統一標準,但需較多時間;如果由多數人來做則反之;因此將語料庫定在一百萬詞是較便於操作的,又可以累積經驗,以便往後的擴充[19]。

下面我們將介紹 CLEC 的樣本內容以及建立方式,以便之後運用到文法模組的設計與撰寫。

2.4.3　CLEC 的建立

在介紹 CLEC 如何建立前,必須說明什麼是 LC(Learner Corpus),他和 ECNS(English Corpus of Native Speakers)有何不同。

由於英語並非我們的母語,所以英語在生活環境當中,是需要學習的一門科目,除了課程需要,一般學生也幾乎很少在日常生活中運用。對於學習者來說,無論是聽、說、讀、寫各方面,所用的英語詞彙大多都和學習環境與校園

生活有關，從 CLEC（LC）的統計表中就可以看出端倪。

而對於英語系國家的人來說，由於英語已經是母語，需要學習的是語言的使用技巧，以及文章體裁、結構…等方面，因而在 ECNS 中，所使用的詞彙也與日常生活有關，是與英語學習者所使用的詞彙不同；於是我們在建立 CLEC（LC）時，樣本必須來自於不同發展階段的學習者；而制定 ECNS 則考慮文體類型而非語言能力。學習者的寫作能力只是停留在"一般的"英語（例如我們不能期望學習者去寫小說、社論、科技文章，而這些類型是一般 ECNS 都有的樣本）[19]。

1. 樣本的選定

CLEC 的語料分布，定為 5 個階段，如表 2-2 所示：

表 2-2　　CLEC 語料分布表

類型	詞次
ST2	208088
ST3	209043
ST4	212855
ST5	214510
ST6	226106
總計	1070602

1. ST2：中學階段，主要是高中生。
2. ST3：大學英語 4 級，大學 1～2 年級非英語科系學習者。
3. ST4：大學英語 6 級，大學 3～4 年級非英語科系學習者。
4. ST5：英語專業科系 1～2 年級學習者。
5. ST6：英語專業科系 3～4 年級學習者。[19]

2. 言語失誤分類表的制定

為了要以學習者語料庫作為系統文法模組的驗證依據，我們希望他是擁有清楚而明確的錯誤類型標示的；CLEC 中的編制原則正是如此，這也是他的特點。他的編寫原則如下[19]：

1. 簡單合理，易於系統操作。

   採兩級分類，先分為十一大類（如：[vp]），每大類再細分為小類（如：[vp6]），總共六十一小類（見表 2-3）[19]。

2. 分類表的類別數量適中。

   分類太粗不易看出失誤為何；分類太細，容易將同種失誤歸到不同類別。目前六十一類屬於中等規模的分類表。

3. 提供足夠的錯誤信息。

   CLEC 當中的句子提供了相當足夠的錯誤訊息標示。
   例如：In the past, people are [vp6, 4－] kind to each other...，使用方括號標示錯誤，其中的 vp6 表示是動詞第六種錯誤，也就是時態的問題，至於 4-表示錯誤發生的範圍，－代表錯誤的位置，4 代表失誤的字前有四個導致錯誤的字，要根據這四個字才能判斷 are 這個詞用錯了。

4. 開放性。

   可定義更細分類的錯誤類型，如[vp6]可細分為[vp61]、[vp62]...等。

5. 對語體誤用或失誤原因暫不做標記。

   目前只針對錯誤處進行標記，至於導致使用者使用錯誤的原因，由於牽涉到標記者的主觀意見，因此暫不標記。

表 2-3 為其中整理的語言失誤分類表，詳細內容請參考[19]。

表 2-3 　言語失誤分類表

詞形		動詞短語		名詞短語		代詞	
碼	類型	碼	類型	碼	類型	碼	類型
fm1	Spelling	vp1	pattern	np1	pattern	pr1	reference
fm2	word building	vp2	set phrase	np2	set phrase	pr2	anticipatory it
fm3	Capitalization	vp3	agreement	np3	agreement	pr3	agreement
		vp4	finite/non-finite	np4	case	pr4	case
		vp5	non-finite	np5	countability	pr5	wh-
		vp6	tense	np6	number	pr6	indefinite
		vp7	voice	np7	article		
		vp8	mood	np8	quantifiers		
		vp9	modal/auxiliary	np9	other determiners		

形容詞短語		副詞		介詞短語		連詞	
碼	類型	碼	類型	碼	類型	碼	類型
aj1	pattern	ad1	order	pp1	pattern	cj1	pattern
aj2	set phrase	ad2	modification	pp2	set phrase	cj2	set phrase
aj3	degree	ad3	degree				
aj4	-ed/-ing confusion						
aj5	predicative/attributive						

詞彙		搭配		句法	
碼	類型	碼	類型	碼	類型
wd1	order	cc1	noun/noun	sn1	run-on sentence
wd2	part of speech	cc2	noun/verb	sn2	sentence fragment
wd3	substitution	cc3	verb/noun	sn3	dangling modifier
wd4	absence	cc4	adj/noun	sn4	illogical comparison
wd5	redundancy	cc5	verb/adv	sn5	topic prominence
wd6	repetition	cc6	adv/adj	sn6	coordination
wd7	ambiguity			sn7	subordination
				sn8	structural deficiency
				sn9	punctuation

2.4.4　CLEC 的統計分析

除此之外，先前提到的 Learner Corpus（LC）與 English Corpus of Native Speakers（ECNS），我們也可以從 CLEC 中的分析資訊，得到一些資料，以便於我們可以針對這些資訊，進行我們的程式撰寫與分析。

CLEC 之中的統計表包括：詞頻排列表、拼寫失誤表、詞目表、詞頻分布表、詞目分布表、語法標記頻數表、言語失誤表…等等，在此我們針對其中的幾個表的分析做一些介紹。

1.　超用詞（overused）與少用詞（underused）

在 CLEC 中提到[19]，如果我們比較 FLOB Corpus（Freiburg Lancaster-Oslo/Bergen）與 CLEC 當中的詞彙，利用 Wordsmith 這個軟體分析語料庫，他可以對某一個語料庫和一個參考語料庫作比較分析，看其中哪些詞是超用的，哪些是少用的。

超用詞和少用詞可以呈現出語料庫的特點，學習者語料庫中可以發現，一些漢語拼音的專有名詞都成為超用詞，但英語中的專有名詞卻是少用詞，抑或是當題材不同時，也會有一些影響，例如：學習者語料庫中的作文都是與個人和學校生活有關的詞彙，如：life, school, college, campus, English, friends, knowledge, teachers, students…等等，都是所謂的超用詞；相反的，British, church, European, community, Labour, minister, religious, tax…等等，這些常出現在英語語料庫中的詞，反而成為 CLEC 的少用詞。

2.　詞形使用錯誤

所謂詞形使用錯誤，在 CLEC 中最多的有三種[19]，拼字、造詞、大小寫。拼字及大小寫的錯誤，可能是在查詢字典的時候不小心發生錯誤的，或是並沒有再寫完之後多檢查幾次；但有趣的是 CLEC 中 ST3 和 ST4 的學習者，他們喜歡使用推理的方法以及造詞的規則，自行創造出不存在的詞。例如：maken, limitful, couragely, normy…等詞。

3.　句法使用分析

由 CLEC 中的統計資料中也可以發現句型也是由超用詞與少用詞所影響

的，例如：被動語態（been、by），從屬句（where、which、who...），是學習者所少用的句型；英語中關於時態的句型，也是在此之中所少見的類型，因為這也是學習者最常見的動詞失誤，可能因為怕用錯，所以出現頻率不高。

另外，句法失誤在 CLEC 中分成九種，其中四種佔的比例最大[19]，分別是：結構缺陷、不斷句、不完全句和標點符號。根據分析，結構缺陷和標點符號有密切關係，而不斷句和不完全句有密切關係。其他的句型失誤則自成一類，和不斷句與不完全句有關係，可見不斷句是學習者語料庫中常見的失誤，這種失誤和中文的表達方式很有關係。例如 It is used not only in studying English but also in many ways, dancing is a good exercise, when I begin to learn, I feel it difficult to do, but I do every day, little by little...，這就是很明顯的例子，屬於不斷句（run-on sentences），我們除了應該注意使用者該寫完整的句子之外，還應該注意標點符號的用法，避免由標點符號所組成的不斷句。

2.4.5 CLEC 的製作工具

語料庫是將語言資料輸入資料庫，且加上分類及相關資訊的資料庫；應用在輸入上面的軟體有[19]：WordCruncher, MicroConcord, Longman's Concordancer, Concordance, Concordancer, Lexa, TACT, Wordsmith...等等，經過實驗和比較功能，也因為他們是免費或共享軟體，於是選用 TACT 和 Wordsmith。除此之外，CLEC 的研究人員也撰寫了相關的專門軟體；這些所建立的資料都是純文字檔，方便讀取與加工，而統計方面的資料，則交由 Microsoft Office Excel 來製作，如果是較為複雜的統計，則使用了 SPSS、Statistica 和 Havard Chart。

2.4.6 CLEC 實例說明

表 2-4 的作文是 CLEC 樣本中的 ST2 類型，我們可以很清楚的看到，所有 CLEC 中的句子都是使用 XML 語法來標記，XML 它是一九八六年國際標準組織(International Standards Organization, ISO)公佈的一個名為「標準通用標示語言」(Standard Generalized Markup Language, SGML)的精簡版/子集合。XML 掌握了 SGML 其延展性、文件自我描述特性、以及其強大的文件結構化功能，但 XML 卻摒除了 SGML 過於龐大複雜以及不易普及化的缺點。字面上

來看 XML 是一種標示語言，但嚴格來說它和 SGML 一樣是一種「元語言」
(meta-language)。換言之，XML 是一種用來定義其它語言的語法系統。這正
是 XML 功能強大的主因。它可促進各專業機構、不同產業界、學術界和特定
應用領域發展各自標準的文件和訊息，以利資訊的交換、處理和相關衍生性資
料加值服務[16]。XML 語法讓我們在研究 CLEC 的時候，能夠了解一些與文章
有關的資訊。

　　以表 2-4 來看，01、02 列分別標示出此篇作文作者的資訊，如：<ST 2>，
前面的 ST 表示是學生，後方的 2 則代表該考生是第二類型；接著是<SEX ?>
性別不詳，<SCH GDWYWMDXFSWYXX >就讀學校使用代號，<TITLE A
Shop>代表本篇的作文題目是 A Shop，其餘部份 03～11 列就是考生主要的作
文內容了。標示問號的部分代表沒有作者的該項資訊；另外由於 ST2 類型的學
生文章並非考試，因此未標示出分數是多少。

<p align="center">表 2-4　　CLEC 內容範例</p>

01　<ST 2> <SEX ?><Y ?> <SCH GDWYWMDXFSWYXX> <AGE ?>
02　<WAY ?><DIC ?> <TYP 2> <TITLE A Shop>
03　There is a fruit shop near my home. its [fm3, 1-] name is Many [fm1,-]
04　fruit shop. It's not very big, but it's clean and bright. There is [vp3,-2] two
05　women working in it. The women are very friendly [wd2, 1-] and busy.
06　Every buyer comes to the shop, they both give them smiles and say.
07　[sn9, s] Hello. Can I help you? So every buyer comes to here are very
08　satisfy. [sn8,s]The shop has many different kinds of fruit. There are
09　apples [sn9, s]oranges bananes [fm1,-]pears bananas many [wd6].
10　[sn8,s] So it aways [fm1,-] give [vp3, 1-] the buyers a good time. I like
11　the shop very much.

　　以下是 CLEC 中錯誤標記的方式，分為錯誤位置、錯誤類型與錯誤範圍標
記。錯誤位置是在句子當中使用方括號標記在錯誤單字的後方，括號中左邊標
示錯誤類型，右邊則標示出錯誤範圍，如 03 列後方：its [fm3, 1-] name is Many
[fm1,-] fruit shop.，即表示此句中 its、Many 兩個字是錯誤位置，而錯誤類型
分別是 fm3（大小寫錯誤）、fm1（拼字錯誤）。

　　至於針對錯誤範圍標示部分，有以下幾種分類，分別說明如下(詳細請參閱

[19])：

1. 錯誤字前方導致錯誤：

   05 列：The women are very friendly [wd2, 1-] and busy.
   括號中錯誤範圍標示的 1 在一號前，表示是 friendly 是錯誤，是因為往前
   數一個字的位置的字導致該字發生錯誤，也就是因為 very 也是副詞，而導
   致錯誤。

2. 錯誤字後方導致錯誤：

   04 列：There is [vp3,-2] two women working in it.
   括號中錯誤範圍標示的 2 在一號後，仍表示是 is 發生錯誤，但原因是後方
   的兩個字 two women 所導致的。

3. 本身錯誤：

   03 列：its [fm3, 1-] name is Many [fm1,-] fruit shop.
   此句所示的第二個括號，在一號前後並沒有任何數字標示，此即表示並沒
   有其他的字導致 Many 發生錯誤，是這個字本身就拼錯，並非原本作者所
   要使用的字。

4. 句中結構或標點錯誤：

   08 列：There are apples [sn9,s] oranges bananes [fm1,-] pears...
   此句的第一個錯誤標示範圍僅用一個 s 表示，此代表是句中的結構或標點
   發生錯誤，因此這裡發生錯誤的原因，在於表示多個名詞的時候，需要用
   逗點分開。

   了解錯誤的標記方式與相關資訊之後，在下一章將會詳細的介紹如何利用
這些分類資訊，以設計實作系統。

三、 系統設計與實做

3.1 系統架構與流程

下圖是系統設計架構圖，學習者首先在系統當中輸入英文作文，接著依序經過三個部份的檢查，分別是拼字檢查、文法檢查、同義字檢查，最後將前述三種類型的錯誤檢查結果標示出來，再交給評分系統，由人工進行最後的判別評分，同時把結果輸出到網頁上面。
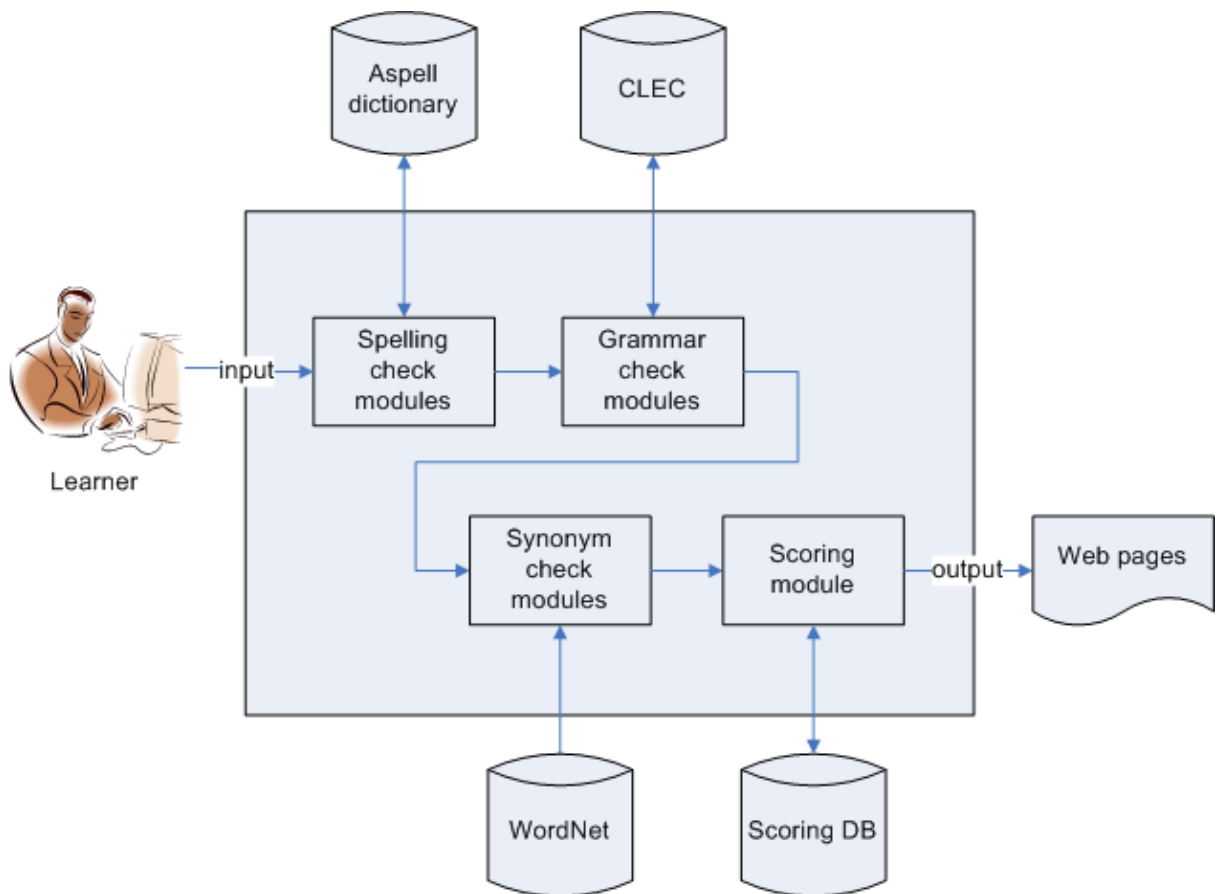
圖 3-1   系統架構圖

基於以上的系統架構，我們繼續詳述系統的程式流程，以便清楚系統是如何處理一篇英語作文，從其中挑出錯誤並且標示出來。
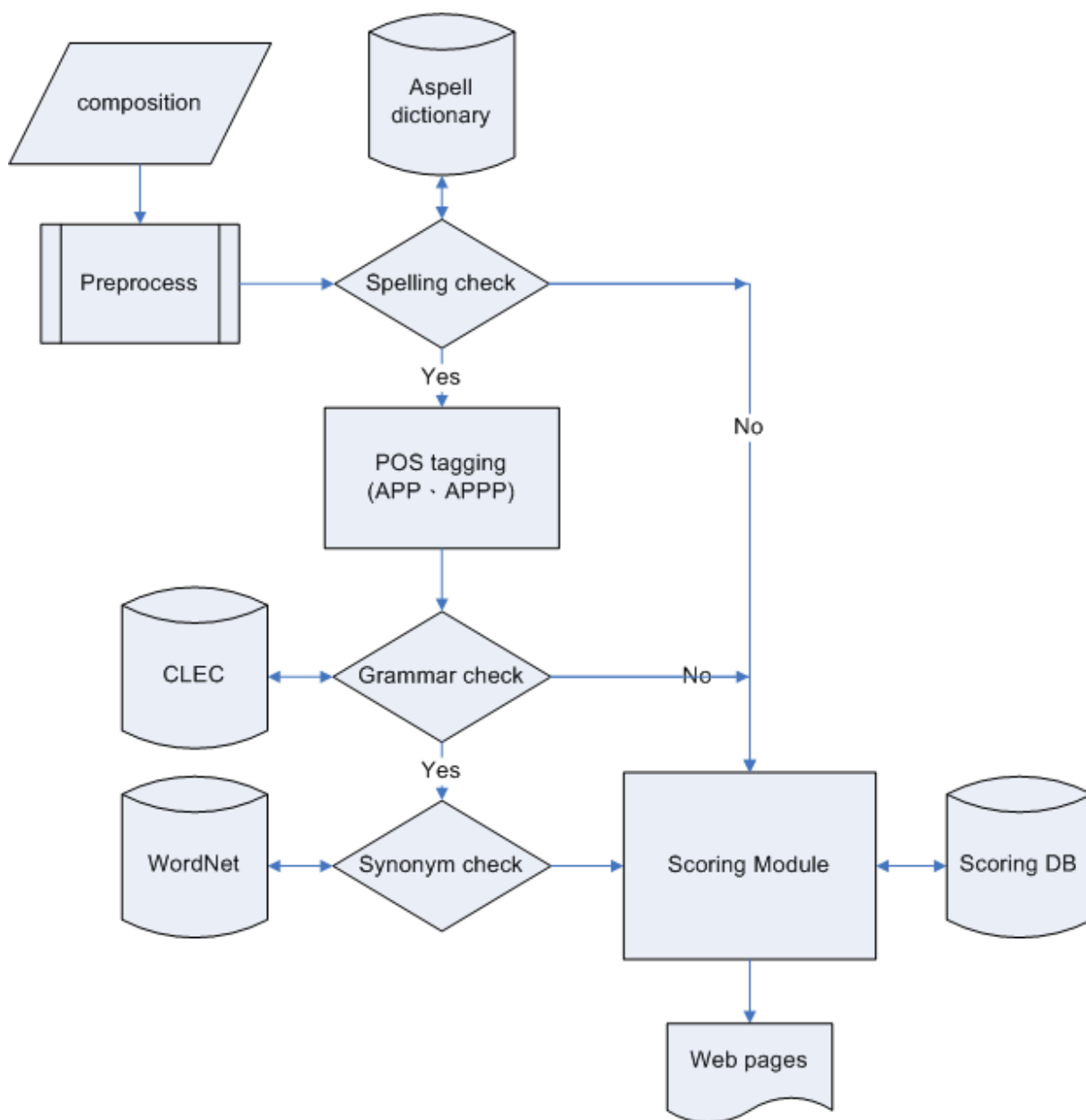
圖 3-2　　系統流程圖

　　從流程圖當中[22]，我們可以看到系統中運作的情形。首先作文會經過前處理，進行斷句、斷詞的程序，接著進入拼字檢查，如果不通過就直接評分，通過的話則繼續以 APP 進行詞性標記，完畢之後便開始以文法模組檢驗，如果上述的拼字、文法的檢查均正確，我們最後送到同義字檢查模組，檢查文章是否文不對題。所有以上的結果，將送至評分系統，可讓閱卷者針對系統檢查的情形加以評分及建議，最後以網頁輸出。

3.1.1　前處理

在文章在進入各模組進行檢查之前，系統進行文章的前處理 (preprocess)，所謂前處理是將學習者的英文作文切分成單詞，以便進行拼字檢查的動作。
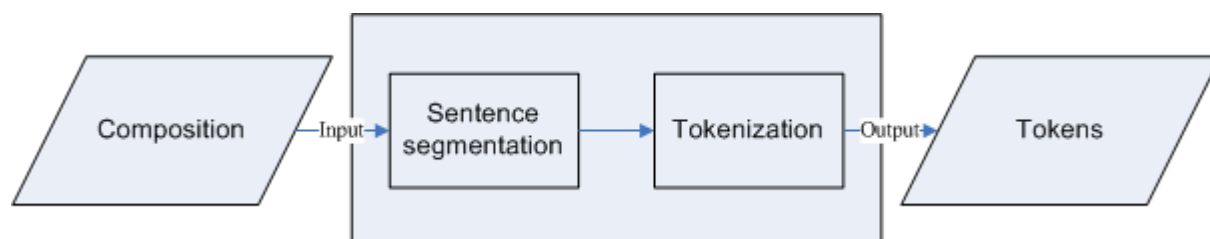
圖 3-3　　前處理流程圖

當學習者將英語作文輸入之後，首先便進行前處理的動作，在這裡分成兩個部份[22]：

1. 切成句子：

首先先將文章依據句點切分成許多句子，這部分系統忽略某些縮寫的部份，如：Mr.、Dr....等。

2. 切成單字：

作文經過以上的切分，已成為單一的句子，此時系統再依據單字之間所出現的空白切成單字，這邊將縮寫 don't 以及句中出現逗點時，是緊連單字的情況考慮進去。

經過前處理之後，我們就會將這些處理好的文字，送到 Aspell 裡面去檢查，以便進行之後步驟。

3.2　拼字檢查模組

系統將學習者的文章切成多個單字(token)之後，接著進行的就是將已經切好的單字送入拼字檢查模組，進行錯誤檢查。我們所使用 Aspell 的版本是

0.60.3，並且在模組中以函式庫的方式呼叫使用。
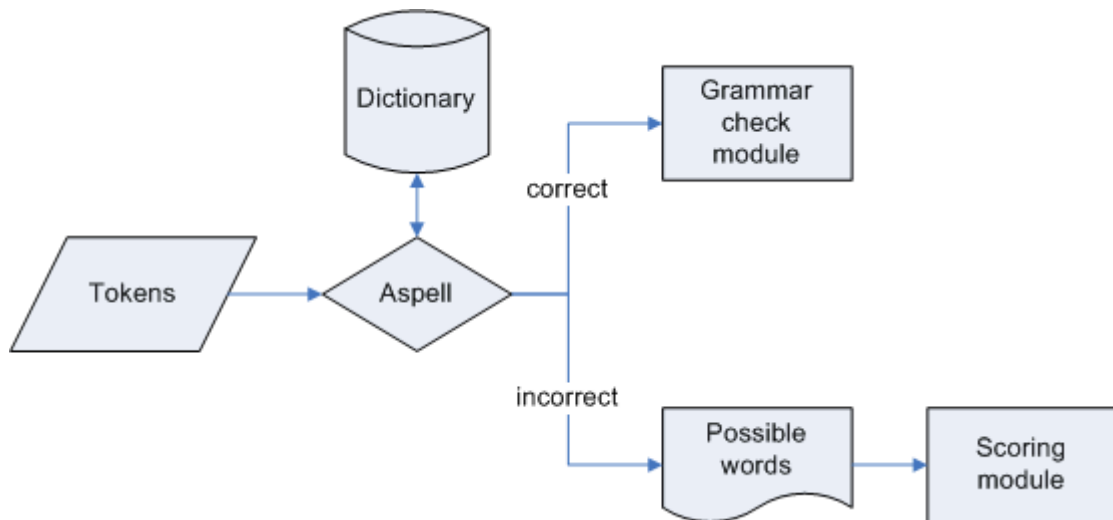
圖 3-4　拼字檢查模組流程圖

　　進行檢查時，程式會將剛才的 token 放入字典中進行比對，如果拼字檢查
正確，就繼續進行下個模組；因為如果拼字錯誤，則可能導致在文法檢查模組
中，標記詞性的時候無法辨識該字而導致文法判別錯誤。因此如果未通過拼字
檢查的句子，我們便不送到文法檢查模組，而直接送往評分模組。

　　而經過統計之後，發現關於拼字的錯誤大致有下列三種：

1. 拼字錯誤(spelling)：

例：said→siad、important→importen、lucky→luckly
此類型的錯誤可直接由 Aspell 找出來，但是比率並非達到 90%的原因，是因為
存在某些特例無法解決。如：every thing。這兩個字都沒有錯誤，文法也沒有
錯誤，但是大家都知道在英文文法中並非以此種方式使用，因而 CLEC 標記為
錯誤。
想要解決此一類型的問題，我們可以用例外(exception)的方式記錄在程式模組
當中，直接將錯誤挑出。

2. 人工造詞(word building)：

例：using→useing、useless→useness、happily→happyly
造成這些錯誤的原因，多半是對於詞性的變化不甚熟悉，而產生這種張冠李戴
的情形。此類型同 1 也可由 Aspell 找出。

24

有些例外的情形，如：I be live that learning English...，很明顯出錯是用法加上拼字錯誤的問題，應該是 I believe that learning English...，這類無法由 Aspell 找出來的問題，我們就會利用文法模組加以解決，因為 be、live 都是動詞，我們知道一個句子裡面如果有兩個動詞，是需要用連接詞將其分開的。

3. 大小寫錯誤(capitalization)：

例：「we [fm3,-] will make it better.」、「I am sure that china[fm3,-] will...」、「For example, In [fm3,1-] 1960 the life」
上述三種情形，分別是句子開頭沒有大寫、專有名詞沒有大寫、逗點後面應該是小寫。開頭或句中需要大小寫的辨別，這可以由程式解決；但專有名詞的部份，則需要在程式當中撰寫所遇到的例外才可以辨別。

以上錯誤除了可以用軟體(如：Aspell)來檢查之外，關於例外的情形是需要文法檢查模組來輔助的，從以上針對詞形的分析範例中，說明了系統分析的過程及相關的內容。除此之外，其餘的錯誤部份，我們也是針對分析的結果，再加以撰寫相關的文法檢查模組來檢查錯誤,這也是在下一節中主要的研究重點。

3.3 文法檢查模組設計

文法檢查模組包含以下幾個部份，依序如後：詞性標記(APP)、結構標記(APPP)、文法檢查(modules)、CLEC 驗證。
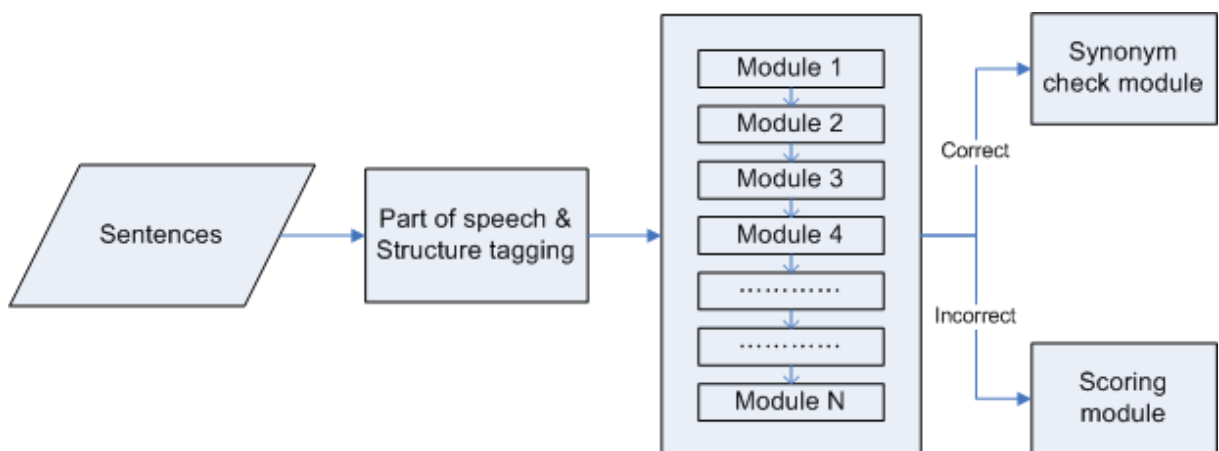
圖 3-5　　文法檢查模組流程圖

從此圖即可以清楚的看出，剛才經過 Aspell 檢查無誤的句子，經過詞性以及句子結構的標記，便依照這些標記進行文法檢查，句子經過每一個文法檢查的模組檢查過後，如果全部都正確無誤，就將資訊送到同義字檢查模組；如果發生錯誤，便一一記錄下來錯誤的模組有幾個，並且送到評分模組。

3.3.1　詞性及句構標記

我們採用的 parser 是 Apple Pie Parser(APP)。我們將通過拼字檢查的句子，送進 APP，APP 會對句中的單字作出詞性標記以及句子的結構；之後，再將句子送進 Apple Pie Parser Parser(APPP)進行樹狀標記，標記完之後，再交由文法模組做錯誤檢查。
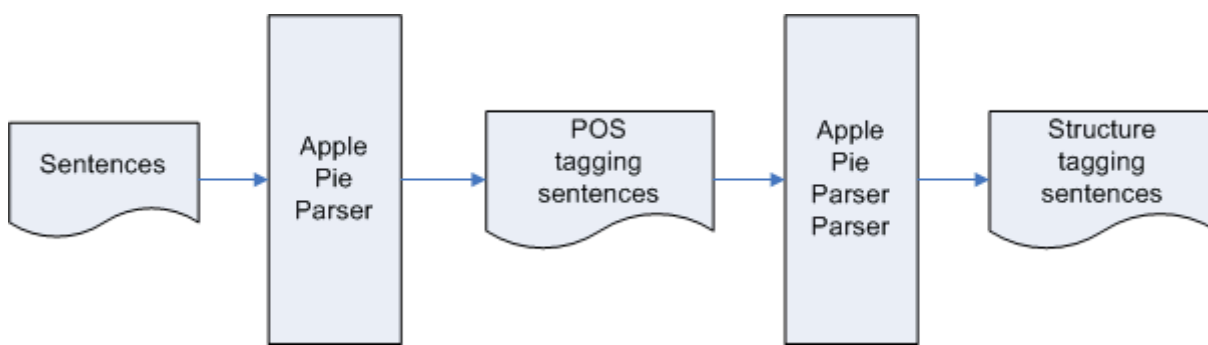
圖 3-6　　APP 及 APPP 解析句子流程圖

3.3.2　文法檢查模組驗證方式

系統將文章做過以上的處理之後，就會送到文法檢查模組中做檢查，但是這些文法模組的設計分類方式，以及相關的驗證，需要大量的樣本來做測試。因此，我們將依照文法模組在 CLEC 各類型句子中的驗證結果，去設計及撰寫文法模組，如下圖所示：
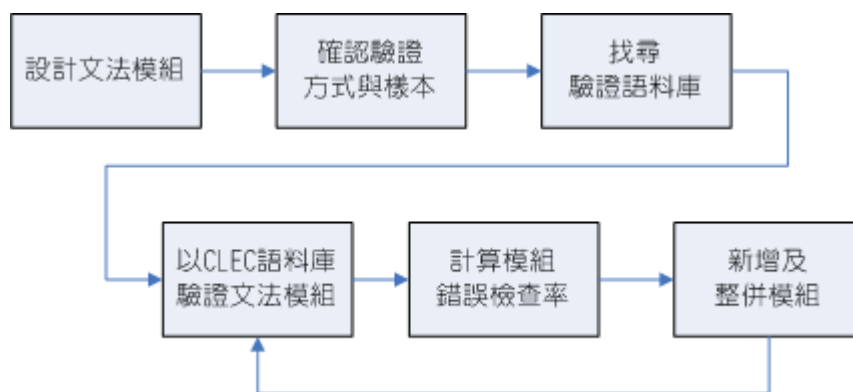
圖 3-7　　文法檢查模組驗證方式設計流程圖

　　我們每設計一個文法模組，便以 CLEC 中的錯誤句子做測試，測試完畢後，便統計尚有多少數量及類型錯誤還未能找出，針對模組不足之處再加以補強，並且重複這樣的步驟，以新增各類型文法檢查模組。

　　實做的方式則如下圖所示：
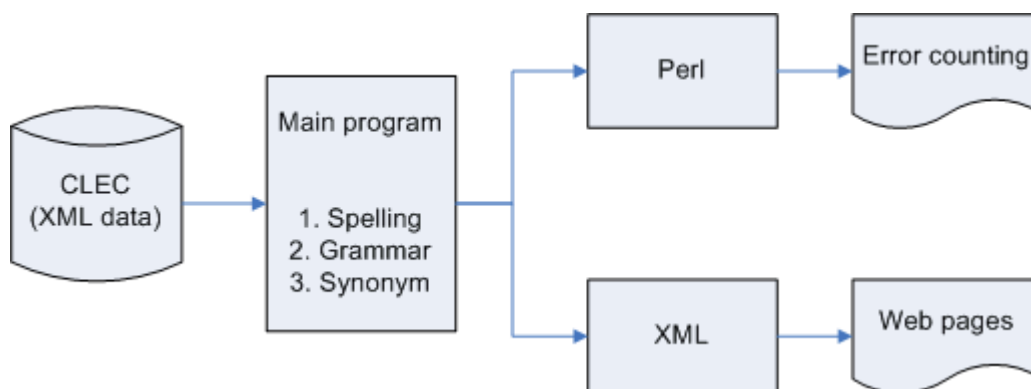
圖 3-8　　文法檢查模組驗證方式實作流程圖

　　圖中 CLEC 的資料送進我們的文法檢查模組中檢驗，檢驗後主要分成兩個部份：

1.　後端錯誤統計：

　　利用 perl 去呼叫檢驗程式，將 CLEC 中有錯但尚未檢查出錯誤的句子經過主程式檢驗之後做成列表，由於 perl 對於字串的處理能力很強，所以在這個部份能很快的計算出句子數目，以利後續在撰寫程式時，確認哪類型的文法錯誤數量較多，需要先行解決。

2. 前端網頁測試：

後端撰寫完畢的文法檢查模組，可以由前端網頁呼叫，測試目前文法檢查模組可以偵測的狀況有哪些。並將檢查出的拼字錯誤、文法錯誤、語義錯誤的訊息呈現在網頁上。

3.3.3 CLEC 樣本統計

系統藉由以上檢驗方式，首先統計出兩項資料：

1. CLEC 原本樣本句數：

我們將 CLEC 中的 ST2～ST6 樣本輸入 perl 中，雖然已知在 CLEC 中的詞數為一百萬詞，但並非所有文章中的句子都有發生錯誤，因此我們僅統計有發生錯誤的句數有多少，同時將它分類。

2. CLEC 句數錯誤類型比率：

表 3-1 是 CLEC 語料庫經過 perl 程式檢驗後，而得到的各類型錯誤數量。

表 3-1　CLEC 錯誤類型及句數統計表

錯誤類型	句數	錯誤類型	句數	錯誤類型	句數
[ad1]	151	[np2]	56	[sn5]	25
[ad2]	72	[np3]	678	[sn6]	131
[ad3]	17	[np4]	109	[sn7]	125
[aj1]	30	[np5]	209	[sn8]	2523
[aj2]	29	[np6]	1251	[sn9]	1888
[aj3]	156	[np7]	436	[vp1]	776
[aj4]	37	[np8]	99	[vp2]	321
[aj5]	10	[np9]	45	[vp3]	1499
[cc1]	154	[pp1]	309	[vp4]	285
[cc2]	167	[pp2]	380	[vp5]	300
[cc3]	673	[pr1]	340	[vp6]	1548
[cc4]	228	[pr2]	94	[vp7]	318

[cc5]	48	[pr3]	203	[vp8]	57
[cc6]	24	[pr4]	116	[vp9]	484
[cj1]	60	[pr5]	69	[wd1]	210
[cj2]	21	[pr6]	12	[wd2]	1375
[fm1]	6586	[sn1]	1116	[wd3]	3405
[fm2]	1119	[sn2]	1232	[wd4]	1613
[fm3]	2402	[sn3]	42	[wd5]	1185
[np1]	74	[sn4]	43	[wd6]	119
				[wd7]	811

統計之後，在 CLEC 裡面的錯誤句數總共有 37895 句，但由於上表並不容易看出來各類型錯誤比率是多少，因此我們再針對 CLEC 中的言語失誤分類表上面的 11 大類做出分類和百分比。

表 3-2 是以 CLEC 中 11 大類的錯誤類型，分別統計數量而得到的，並且計算各大類所佔的比率。

表 3-2   CLEC 錯誤類型比率統計表

詞形[fm]	10107	26.7%
動詞短語[vp]	5588	14.7%
名詞短語[np]	2957	7.8%
代詞[pr]	834	2.2%
形容詞短語[aj]	262	0.7%
副詞[ad]	240	0.6%
介詞短語[pp]	689	1.8%
連詞[cj]	81	0.2%
詞彙[wd]	8718	23.0%
搭配[cc]	1294	3.4%
句法[sn]	7125	18.8%
總計	37895	100.00%

圖 3-9 是將表 3-2 以統計圖的方式呈現，讓我們可以更清楚的看出各類型錯誤在整體錯誤類型中所佔的比率。
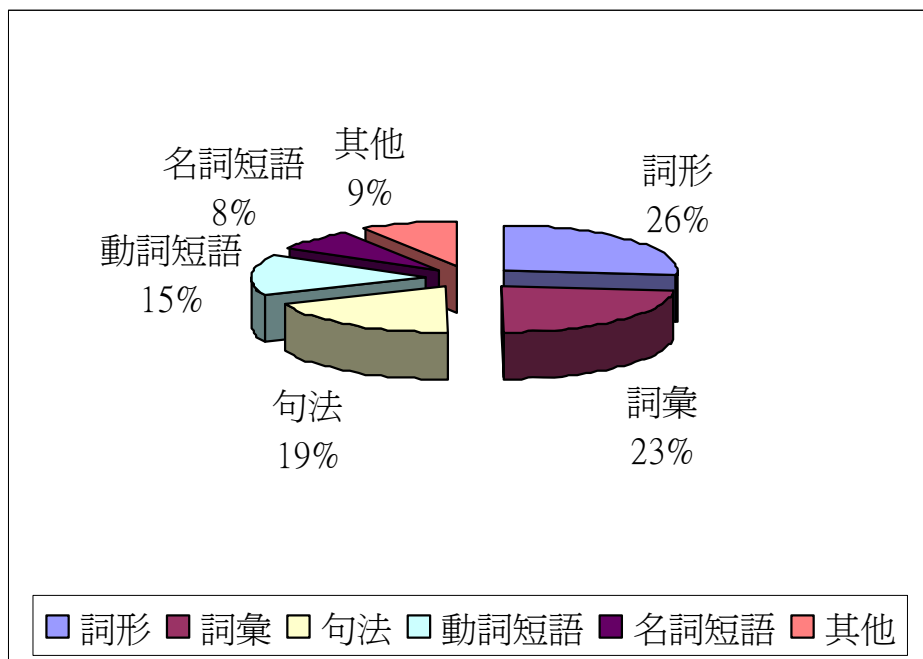
圖 3-9　　CLEC 錯誤類型及句數分類統計圖

經過我們的 perl 程式統計，可以看到上圖是使用不同的圖表來表示目前 CLEC 中的錯誤樣本句數，其中代詞、形容詞短語、副詞、介詞短語、連詞、詞彙、搭配在統計圖中表示「其他」，而前五名數量的錯誤數量多寡依序分別是「詞形」、「詞彙」、「句法」、「動詞短語」、「名詞短語」…等。

3.4　文法檢查模組實作

在文法檢查模組當中，我們是用以下幾種方式來進行分析：

1. 詞性：

語言當中所包含的詞彙範圍很廣，詞數可達成千上萬，因此我們需要詞性來做為這些單字的分類，APP 依照這些單字在句中的結構位置，分析出最恰當的詞性，讓我們可以得到分析的詞性之後，以詞性做為文法檢查的依據。

例如：被動語態➔be 動詞+過去分詞，就可以表示「is broken」、「was opened」、「were selected」…等句型，而不需一一列舉。

2. 單字：

當 APP 不能夠解析出正確的詞性時，或者有時屬於同詞性的單字，但是用法不同時，我們就無法由詞性來做文法檢查。

例如：A 或 AN 的用法。這兩者都是被 APP 解析成[DT]，我們如果使用詞性，就無法判別兩者的不同，因此需要以單字本身作判斷。

3. 句構：

除了一般的詞性、單字的文法判斷之外，另外還有屬於需要句子結構的分析部份，才能幫助我們做上述 1、2 的文法判別。

例如：時態一致性的判斷。由於在這個模組裡面，我們需要找到主要子句和從屬子句中的動詞，進行一致性的判別，首先需要先找出"子句"的位置，而在 APP 的解析中，單字經由各種的組合，而成為某些句子結構，例如：從屬子句就被解析為[SBAR]。我們便可以經由此來做接下來的文法判別。

　　分析完畢後，使用 C++語言撰寫相關程式模組，並針對以上三者加以分析探討，看是否可檢查出文章的錯誤之處，如果通過這部份的檢驗，或是發生不足以影響文法的警告(warning)，那麼將繼續進行同義字檢查模組，進行其餘的錯誤檢查部份；但如文法已是錯誤(error)，程式就直接進入評分模組。

　　以下將詳細說明五個文法模組，以了解模組是如何在系統中運作，以達到文法檢查的目的。

3.4.1　比較級

　　這個模組主要是解決下列幾種類型的文法錯誤，並且附上程式流程圖，以便清楚地知道程式是如何依照 APP 所解析出來的詞性，進行錯誤檢查。

　　我們首先觀察大量 CLEC 中此類型錯誤的句子，如：

　　The footprint was more larger for him.

　　I am taller and more heavy than last year.

　　It is more stronger and lively than before.

　　New computer is more smaller than the first one.

　　English becomes more important and more important

以上五句都有錯誤，我們把以上的句子都經過 APP 解析，並且觀察解析後的詞性，便可以歸類出來下列三種錯誤類型：

1. 副詞比較級[RBR]+單音節形容詞[JJ]
2. 形容詞比較級[JJR]+形容詞比較級[JJR]
3. 副詞比較級[RBR]+多音節形容詞[JJ]+and+副詞比較級[RBR]+多音節形容詞[JJ]
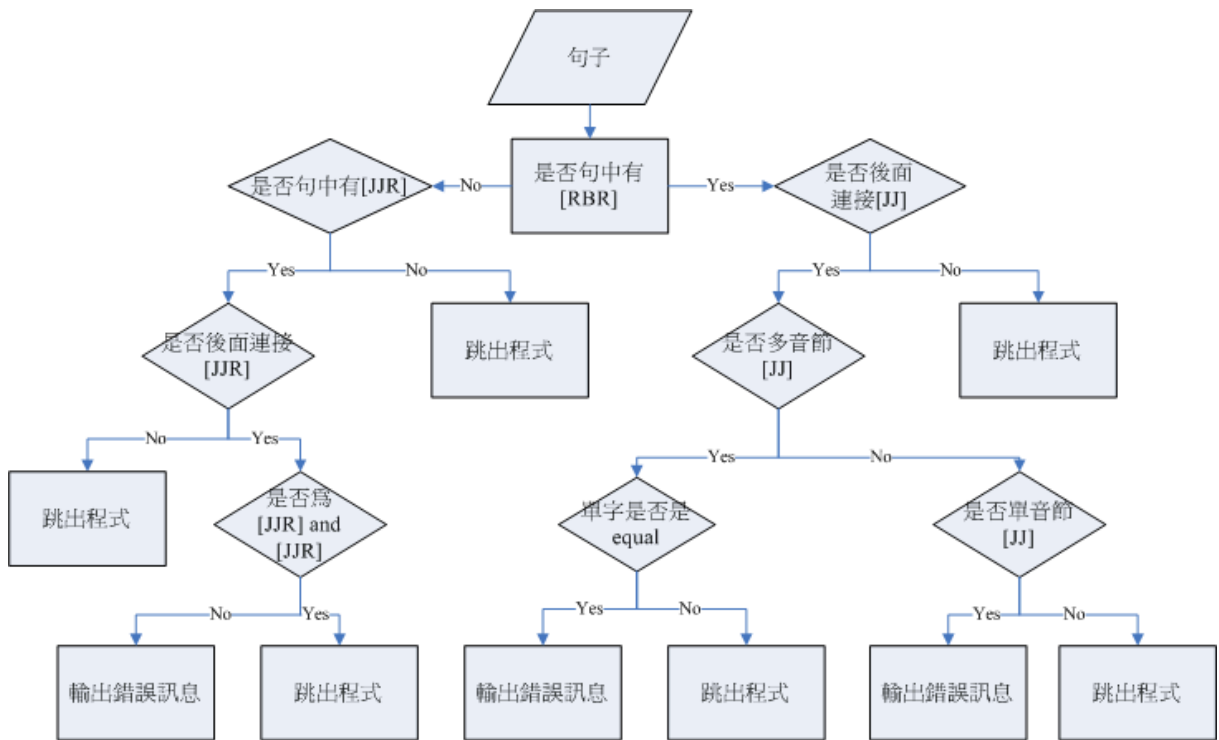
圖 3-10　比較級檢查模組流程圖

1. 副詞比較級[RBR]+單音節形容詞[JJ]

例句：John is more tall than Tom.

APP 的解析結果如下：

```
<Tree>
<CNode tag="S" >
  <CNode tag="NPL" >
    <CLeaf tag="NNP" word="John"   index="0" />
  </CNode>
  <CNode tag="VP" >
    <CLeaf tag="VBZ" word="is"   index="1" />
    <CNode tag="SBAR" >
      <CNode tag="SS" >
        <CNode tag="ADJP" >
          <CLeaf tag="RBR" word="more"   index="2" />
          <CLeaf tag="JJ" word="tall"   index="3" />
        </CNode>
        <CNode tag="PP" >
          <CLeaf tag="IN" word="than"   index="4" />
          <CNode tag="NPL" >
            <CLeaf tag="NNP" word="Tom"   index="5" />
            <CLeaf tag="-PERIOD-" word="."   index="6" />
          </CNode>
        </CNode>
      </CNode>
    </CNode>
  </CNode>
</CNode>
</Tree>
```

圖 3-11　比較級例句在 APP 中解析輸出(1)

由圖 3-11 中 APP 的解析結果可以發現，當句子裡面出現「比較級」加上「形容詞」的時候，APP 會將比較級單字解析成「副詞比較級(RBR)」加上「形容詞(JJ)」。

就此句而言，由於 more 是形容後面的 tall，因此程式將它當成[RBR]，而含有單音節的形容詞此時的詞性變化，應該直接將 tall 寫成形容詞比較級 taller，並且刪除前方的 more 才對；這也呼應了 CLEC 的介紹，由於學習者常常不清楚某些詞性變化的規則，因而製造了錯誤的用法。正確的寫法是，如果此句的 tall 改成多音節的單字時，才可以用 more 來修飾。例如：John is more handsome than Tom.

2. 形容詞比較級[JJR]+形容詞比較級[JJR]

例句：John is more taller than Tom.

形容詞比較級相連時，APP 會把比較級的單字解析成[JJR]加上[JJR]。

此句中 more 後面接著 taller，也就是有「形容詞比較級」相連接的時候，從流程圖的判斷中可以看出，程式會找出這個錯誤；但是在剛才的規則之下其實還有些例外情形，如：The days are getting longer and longer.。此句雖然也出現形容詞比較級相連接的情形，但仍然是正確的用法，因此我們多將句子經過一次是否為[JJR]+and+[JJR]的檢查，以確保檢查結果正確。

```
<Tree>
<CNode tag="S" >
  <CNode tag="NPL" >
    <CLeaf tag="NNP" word="John"   index="0" />
  </CNode>
  <CNode tag="VP" >
    <CLeaf tag="VBZ" word="is"   index="1" />
    <CNode tag="ADVP" >
      <CNode tag="NPL" >
        <CLeaf tag="JJR" word="more"   index="2" />
      </CNode>
      <CLeaf tag="JJR" word="taller"   index="3" />
    </CNode>
    <CNode tag="PP" >
      <CLeaf tag="IN" word="than"   index="4" />
      <CNode tag="NPL" >
        <CLeaf tag="NNP" word="Tom"   index="5" />
        <CLeaf tag="-PERIOD-" word="."   index="6" />
      </CNode>
    </CNode>
  </CNode>
</CNode>
</Tree>
```

圖 3-12   比較級例句在 APP 中解析輸出(2)

以上的例子目的是要找出在句子當中這幾種類型的錯誤用法。而在流程當中需要用到單音節與多音節的判斷，因此我們另外加入此類判斷的副程式。另外，我們也在程式當中加入一些例外情形。

例如：英國著名作家 George Orwell 喬治奧威爾的一部作品《動物農莊》（Animal Farm），裡面提到「All animals are equal, but some animals are

more equal than others.」，這算是反諷用法，在一般的文法當中，是不會有所謂"更"平等(more equal)的詞出現的。

3.4.2 過去分詞

學習者通常在學習動詞的時候，常搞不清楚時態的變化而導致錯誤，尤其是在過去分詞的使用上，因為時態對於華人來說，並不是原本母語中所擁有的東西，我們只會在句子當中加入表示時間的詞，例如：我已經「吃」過蘋果了，並不會因為時間的不同而改變「吃」這個字；但在英語當中的「吃」卻因為時間不同而有所謂的時態變化，這邊就要用「I have eaten the apple.」。

例如 go 在動詞三態中的變化是「go, went, gone」，屬於不規則的動詞三態變化。在 CLEC 的錯誤類型當中，我們也發現過去分詞使用錯誤的比例相當高，常常會有「have(has、had)+過去式動詞」的情形出現；另外，被動語態的用法也是學習者容易忽略的部份，會有「be 動詞+過去式動詞」或「主詞+過去分詞」的情形發生。因此本模組將針對這三種情形進行檢查。

1. have(has、had)+過去式動詞[VBD]
2. be 動詞+過去式動詞[VBD]
3. 主詞+過去分詞[VBN]

圖 3-13 過去分詞檢查模組流程圖

35

1. have(has、had)+過去式動詞[VBD]

例句：The plane has took off.
此句經過 APP 解析輸出如下：

```
<Tree>
<CNode tag="S" >
  <CNode tag="S" >
    <CNode tag="NPL" >
      <CLeaf tag="DT" word="The"   index="0" />
      <CLeaf tag="NN" word="plane"   index="1" />
    </CNode>
    <CNode tag="VP" >
      <CLeaf tag="VBZ" word="has"   index="2" />
      <CNode tag="SBAR" >
        <CNode tag="SS" >
          <CNode tag="VP" >
            <CLeaf tag="VBD" word="took"   index="3" />
            <CNode tag="ADVP" >
              <CLeaf tag="RB" word="off"   index="4" />
            </CNode>
          </CNode>
        </CNode>
      </CNode>
    </CNode>
  </CNode>
  <CLeaf tag="-PERIOD-" word="."   index="5" />
</CNode>
</Tree>
```

圖 3-14　過去分詞例句在 APP 中解析輸出(1)

一般來說，如果有「be 動詞+動作」的句子，其中的動作必須使用「be 動詞+過去分詞」，也就是被動語態來表示，而不是過去式動詞。所以當程式中遇到過去式動詞[VBD]之後，就開始尋找前面的詞是否有 has，當找到的時候程式就輸出錯誤訊息，因為他並不符合過去式動詞的用法，換句話說，其實兩個動詞根本是不能相連接的，所以如果沒有找到[VBN]的情況下，則繼續尋找有無[VBN]，繼續進行其他的判斷。
這句正確的寫法應為「The plane has taken off.」。

2. be 動詞+過去式動詞[VBD]

例句：He was knew to people.
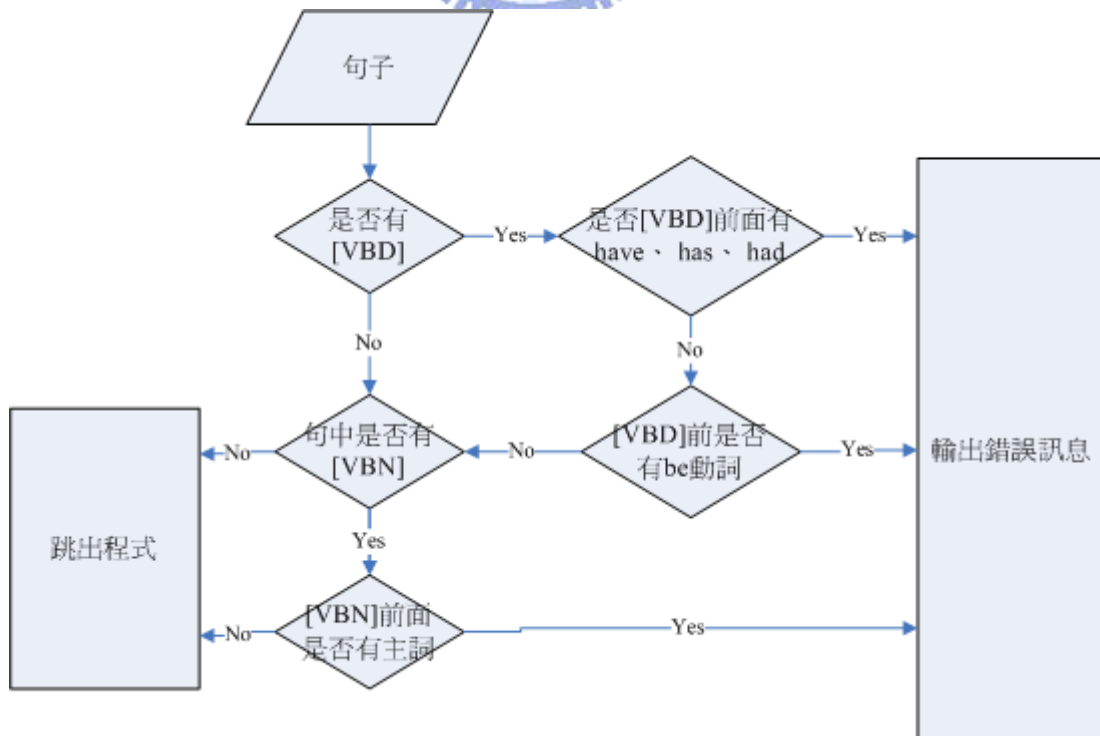
此句經過 APP 解析輸出如下：

```
<Tree>
<CNode tag="S" >
  <CNode tag="NPL" >
    <CLeaf tag="PRP" word="He"   index="0" />
  </CNode>
  <CNode tag="VP" >
    <CLeaf tag="VBD" word="was"   index="1" />
    <CNode tag="SBAR" >
      <CNode tag="SS" >
        <CNode tag="VP" >
          <CLeaf tag="VBD" word="knew"   index="2" />
          <CNode tag="PP" >
            <CLeaf tag="TO" word="to"   index="3" />
            <CNode tag="NPL" >
              <CLeaf tag="NNS" word="people"   index="4" />
              <CLeaf tag="-PERIOD-" word="."   index="5" />
            </CNode>
          </CNode>
        </CNode>
      </CNode>
    </CNode>
  </CNode>
</CNode>
</Tree>
```

圖 3-15　過去分詞例句在 APP 中解析輸出(2)

我們根據解析之後的詞性進行文法判斷，首先找到句中的過去式動詞
knew[VBD]，結果在他的前面發現有 be 動詞 was 存在，根據流程圖所示，程
式可以在這邊找到錯誤，需要更正為「be 動詞+過去分詞」。
同(1)，這句也可由另外的角度來看，在一個句子中，動詞是不可以相連的，除
非有連接詞分開。

3. 主詞+過去分詞[VBN]

例句：The hamburger eaten by my brother.

此句經過 APP 解析輸出如下：

```
<Tree>
<CNode tag="S" >
  <CNode tag="NPL" >
    <CLeaf tag="DT" word="The"   index="0" />
    <CLeaf tag="NN" word="hamburger"   index="1" />
  </CNode>
  <CNode tag="VP" >
    <CLeaf tag="VBN" word="eaten"   index="2" />
    <CNode tag="PP" >
      <CLeaf tag="IN" word="by"   index="3" />
      <CNode tag="NP" >
        <CLeaf tag="PRP$" word="my"   index="4" />
        <CNode tag="NPL" >
          <CLeaf tag="NN" word="brother"   index="5" />
          <CLeaf tag="-PERIOD-" word="."   index="6" />
        </CNode>
      </CNode>
    </CNode>
  </CNode>
</CNode>
</Tree>
```

圖 3-16　過去分詞例句在 APP 中解析輸出(3)

　　同上，我們找到句中的[VBN]，接著在他的前面找到 A hamburger[NPL]，
這邊所謂的主詞，是含有名詞的最小名詞子句[NPL]，因此在缺少 be 動詞的情
況之下，程式找到錯誤，應該將此句更正為「A hamburger was eaten by my
brother.」才對。

3.4.3　A 和 AN 判別

在文法當中，不定冠詞 A 和 AN 的用法，也是常被提到的，在 CLEC 中，也佔了一定的比例。這部份不需要使用 APP 解析出來的詞性，只需要檢驗單字之間的關係即可。

我們歸類下列四種情形，都是錯誤的狀況，提供程式作為檢查之用。

1. A+母音開頭單字
2. A+子音開頭單字(例外)
3. AN+子音開頭單字
4. AN+母音開頭單字(例外)
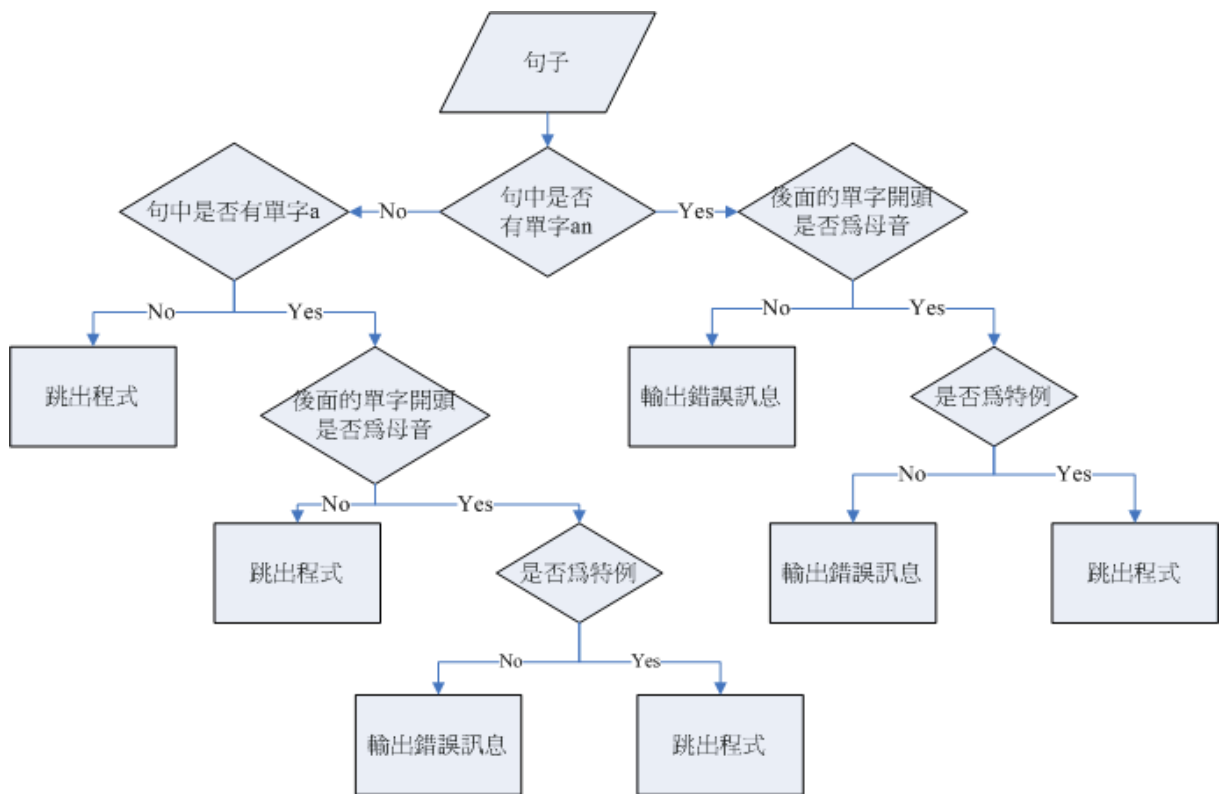
圖 3-17　A 和 AN 判別檢查模組流程圖

一般來說，大部分 A 和 AN 的用法，都可以用單字開頭的字母作為判別的標準，只要開頭是母音的單字就用 AN，是子音的就用 A；但是，嚴格說來，其實是要根據單字的發音為準，發音的開頭是母音就用 AN，是子音才用 A，

但是程式中不容易做到，因此我們使用特例的方式，直接寫入一些例外的字，讓程式的判別能夠正確。

這個模組中並不需要用到詞性的解析，只需要運用到單字，也就是說不需要經過 APP，就可以直接以文法檢查模組判斷正確與否。以下分別說明能夠判斷的四種類型。

1. A+母音開頭單字

例句：This is a apple.
Apple 的開頭不能使用 A 作為不定冠詞。

2. A+子音開頭單字(例外)

例句：I only spent a hour to study.
這是 A or AN 的例外用法，由於 hour 的 h 是不發音的，因此這時候仍然需要使用 an 來當作不定冠詞。

3. AN+子音開頭單字

例句：An train has just arrived the station.
Train 的開頭是子音，因此不能使用 an 作為不定冠詞。

4. AN+母音開頭單字(例外)

例句：He studied in an university.
固然 university 的開頭字母是母音 u，但是在發音來說，開頭的發音是[ju]，並非母音，因此這邊的不定冠詞應該使用 a 才對。

3.4.4 基本動詞使用

在動詞的使用上，我們提出幾個需要注意的地方，主要是屬於動詞出現的次數；通常句子中只會有一個動詞，如果出現兩個動詞的話，動詞之間就必須以連接詞分開。另外，也檢查多個動詞的狀況。

因此這個模組在檢查下列兩種情形：

1. 句中沒有任何動詞。
2. 句中有兩個以上的動詞，但是動詞之間並沒有出現連接詞、不定詞、關係代名詞、助動詞…等。
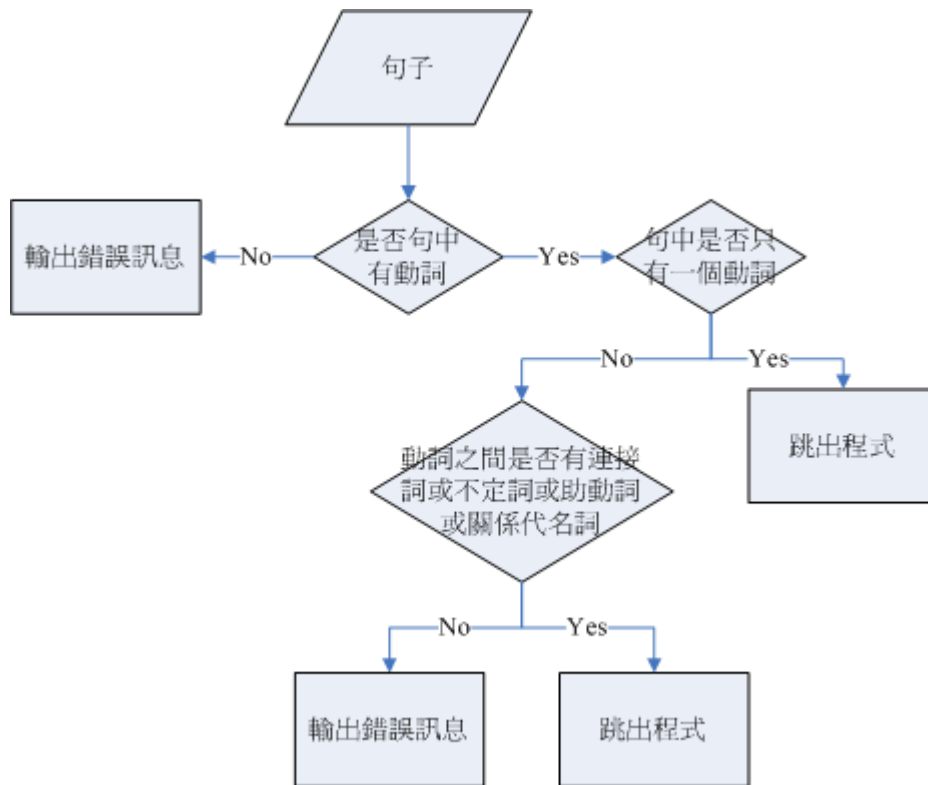
圖 3-18　基本動詞使用檢查模組流程圖

1. 句中沒有任何動詞：

例句：In China, during the last year and just before the armistice.

任何句子中都需要出現動詞，我們列出所有動詞的列表，只要是經過 APP 之後，沒有被解析成原形動詞[VB]、現在式動詞[VBP]、現在式或過去式動詞[VBX]、過去式動詞[VBD]、第三人稱單數動詞[VBZ]其中之一的，就算是在句中沒有找到任何的動詞，我們就在程式當中送出錯誤訊息。

當我們將例句送進 APP 解析，出來的結果是：

```
<Tree>
<CNode tag="S" >
  <CNode tag="PP" >
    <CLeaf tag="IN" word="In"   index="0" />
    <CNode tag="NPL" >
      <CLeaf tag="NNP" word="China"   index="1" />
    </CNode>
  </CNode>
  <CLeaf tag="-COMMA-" word=","   index="2" />
  <CNode tag="SS" >
    <CNode tag="PP" >
      <CLeaf tag="IN" word="during"   index="3" />
      <CNode tag="NPL" >
        <CLeaf tag="DT" word="the"   index="4" />
        <CLeaf tag="JJ" word="last"   index="5" />
        <CLeaf tag="NN" word="year"   index="6" />
      </CNode>
    </CNode>
  </CNode>
  <CLeaf tag="CC" word="and"   index="7" />
  <CNode tag="SS" >
    <CNode tag="NP" >
      <CNode tag="NPL" >
        <CLeaf tag="RB" word="just"   index="8" />
      </CNode>
      <CNode tag="PP" >
        <CLeaf tag="IN" word="before"   index="9" />
        <CNode tag="NPL" >
          <CLeaf tag="DT" word="the"   index="10" />
        </CNode>
      </CNode>
    </CNode>
    <CNode tag="NPL" >
      <CLeaf tag="NN" word="armistice"   index="11" />
      <CLeaf tag="-PERIOD-" word="."   index="12" />
    </CNode>
  </CNode>
</CNode>
</Tree>
```

圖 3-19　基本動詞使用例句在 APP 中解析輸出(1)

我們可以看到，句中時間和地點都描述的很清楚，但是從 APP 解析出來的詞性就可以知道，這個句子並沒有主詞，也找不到動詞，因此這句是錯誤的。

2. 句中有兩個以上的動詞，但是動詞之間並沒有出現連接詞、不定詞、關係代名詞、助動詞…等。

例句：I expect you write to me.

如果句中出現兩個以上的動詞，根據 CLEC 中句子顯示，大致有可能的句型是以上幾種。本句經過 APP 解析之後，得到的結果如下圖：

```
<Tree>
<CNode tag="S" >
  <CNode tag="S" >
    <CNode tag="NPL" >
      <CLeaf tag="PRP" word="I"   index="0" />
    </CNode>
    <CNode tag="VP" >
      <CLeaf tag="VBP" word="expect"   index="1" />
      <CNode tag="SBAR" >
        <CNode tag="SS" >
          <CNode tag="NPL" >
            <CLeaf tag="PRP" word="you"   index="2" />
          </CNode>
          <CNode tag="VP" >
            <CLeaf tag="VBP" word="write"   index="3" />
            <CNode tag="PP" >
              <CLeaf tag="TO" word="to"   index="4" />
              <CNode tag="NPL" >
                <CLeaf tag="PRP" word="me"   index="5" />
              </CNode>
            </CNode>
          </CNode>
        </CNode>
      </CNode>
    </CNode>
  </CNode>
  <CLeaf tag="-PERIOD-" word="."   index="6" />
</CNode>
</Tree>
```

圖 3-20　基本動詞使用例句在 APP 中解析輸出(2)

可以看出動詞是 expect 和 write，有兩個動詞，但是由於文法中規定句子是不能出現兩個動詞的，除非有上述規則中的詞性夾在動詞之間。因此這句發生錯誤，應該更正為 I expect you to write to me.才對。

3.4.5　時態一致

　　這個模組主要是將時態的問題和句子結構結合寫成的，從句子的結構中去判別主要子句的動詞和附屬子句動詞時態之間的關連。

我們在程式模組中主要判別的有以下幾種情形：

1. 主要子句中動詞是現在式、現在完成式、未來式，從屬子句使用任何的時態都可以。

2. 主要子句中動詞是過去式，從屬子句動詞必須用過去式或過去完成式。
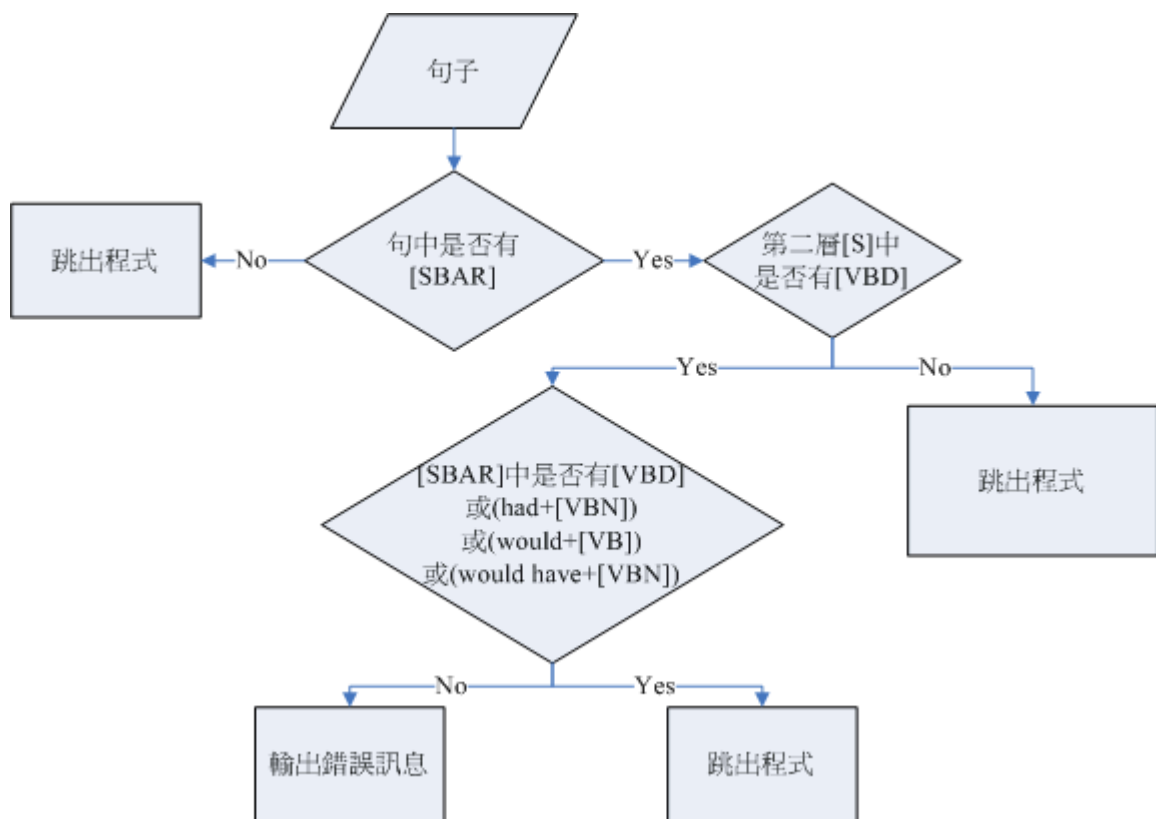
圖 3-21　時態一致檢查模組流程圖

1. 上面已經敘明，無論主要子句的動詞是現在式、現在完成式、未來式，由於從屬子句中的動詞時態不限，因此，無須在模組中做檢查。

2. 主要子句中動詞是過去式，從屬子句動詞必須用過去式或過去完成式。

例句：She knew she will make the right choice.

此句經過 APP 解析得到：

```
<Tree>
<CNode tag="S" >
  <CNode tag="NPL" >
    <CLeaf tag="PRP" word="She"   index="0" />
  </CNode>
  <CNode tag="VP" >
    <CLeaf tag="VBD" word="knew"   index="1" />
    <CNode tag="SBAR" >
      <CNode tag="SS" >
        <CNode tag="NPL" >
          <CLeaf tag="PRP" word="she"   index="2" />
        </CNode>
        <CNode tag="VP" >
          <CLeaf tag="MD" word="will"   index="3" />
          <CNode tag="VP" >
            <CLeaf tag="VB" word="make"   index="4" />
            <CNode tag="NP" >
              <CNode tag="NPL" >
                <CLeaf tag="DT" word="the"   index="5" />
                <CLeaf tag="NN" word="right"   index="6" />
              </CNode>
              <CNode tag="NPL" >
                <CLeaf tag="NN" word="choice"   index="7" />
                <CLeaf tag="-PERIOD-" word="."   index="8" />
              </CNode>
            </CNode>
          </CNode>
        </CNode>
      </CNode>
    </CNode>
  </CNode>
</CNode>
</Tree>
```

圖 3-22　時態一致例句在 APP 中解析輸出

　　首先可以在結果當中找到[SBAR]以進入程式流程，經過追蹤[SBAR]之下的 LEAF，可以得到原形動詞 make，因此在句中找到錯誤，應該將句子更正為

She knew she had made the right choice.才對。

　　除以上五個文法檢查模組之外，尚有許多文法檢查模組，我們盡可能的觀察大量的 CLEC 的錯誤類型句子，以便找出程式模組當中不足的地方，並且將缺少的部份加入，同時我們將某些文法規則整併在一起，在判斷的時候就可以減少一些程式執行的時間，使程式更有效率。

3.5　同義字檢查模組

　　通過以上的拼字檢查模組以及文法檢查模組之後，系統將進行同義字的檢查模組。
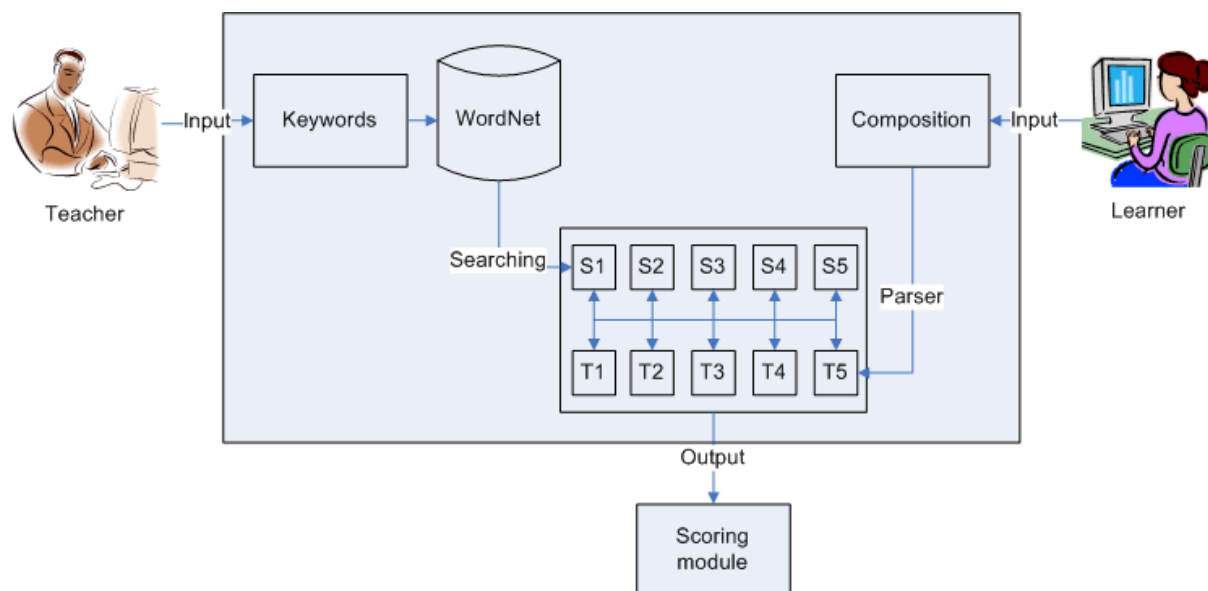
圖 3-23　同義字檢查模組架構圖

　　由圖可知，老師可以針對作文題目指定關鍵字，輸入 WordNet 之後，產生許多同義字，這些同義字與學習者文章的單字(Tokens)逐一做比較，看是否有相符的單字，然後再將結果送到評分系統。
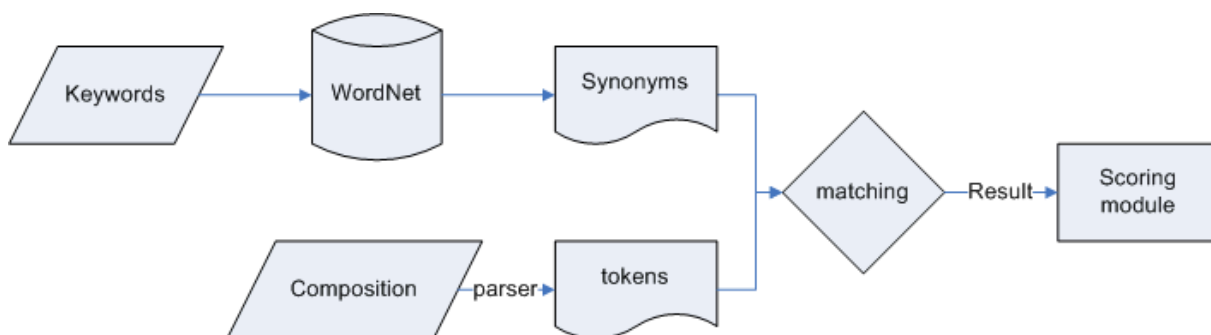
圖 3-24　同義字檢查模組流程圖

　　當學習者的文章通過本系統之拼字檢查模組、文法檢查模組、同義字檢查模組之後，所有的結果都會送交評分模組，提供老師做最後的評分以及建議，我們將在下一節介紹。

3.6　評分模組

　　包括拼字檢查模組、文法檢查模組、同義字檢查模組，在三者之中所發生的錯誤，均會呈現在這個部份，老師可以依據所標示出來的地方加以評分。
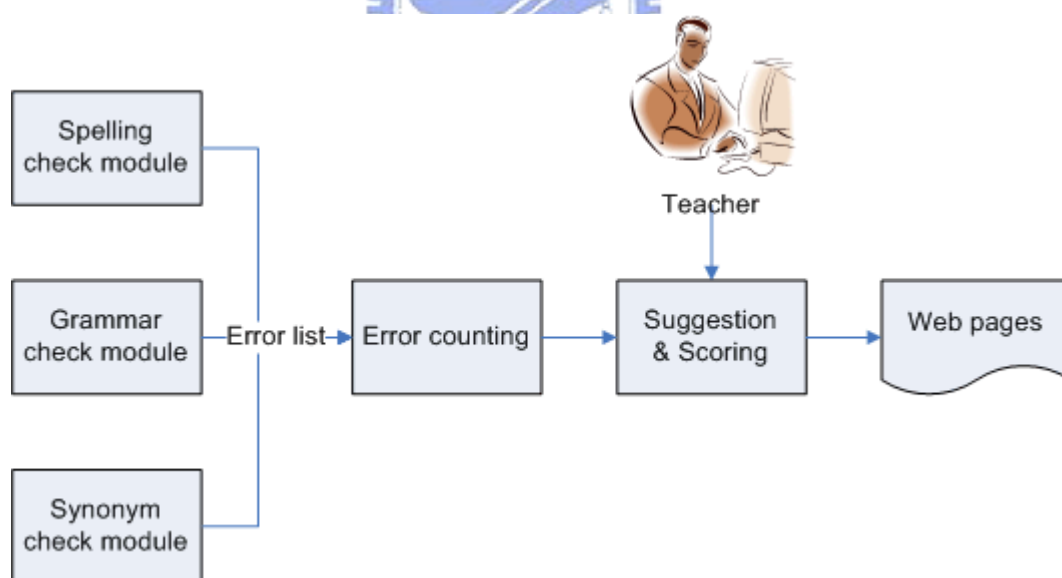
圖 3-25　評分模組流程圖

　　評分模組接收來自於拼字檢查模組、文法檢查模組、同義字檢查模組的錯誤訊息，將他們彙整之後，提供給老師過目，老師便可針對每個拼字、文法、同義字錯誤的地方酌予扣分，也可以給予相關的評語或正確的用法，等到最後確認之後，便可以送出結果，將結果送出並且儲存，學習者便可以在網頁瀏覽

自己的學習成果。

3.7 軟體程式建構技術

我們在開發以上各模組時，均是採用所謂 MVC(Model-View-Controller)的概念，此架構概念源自 Smalltalk 語言[13]，我們先簡單用以下的圖示說明[7]：
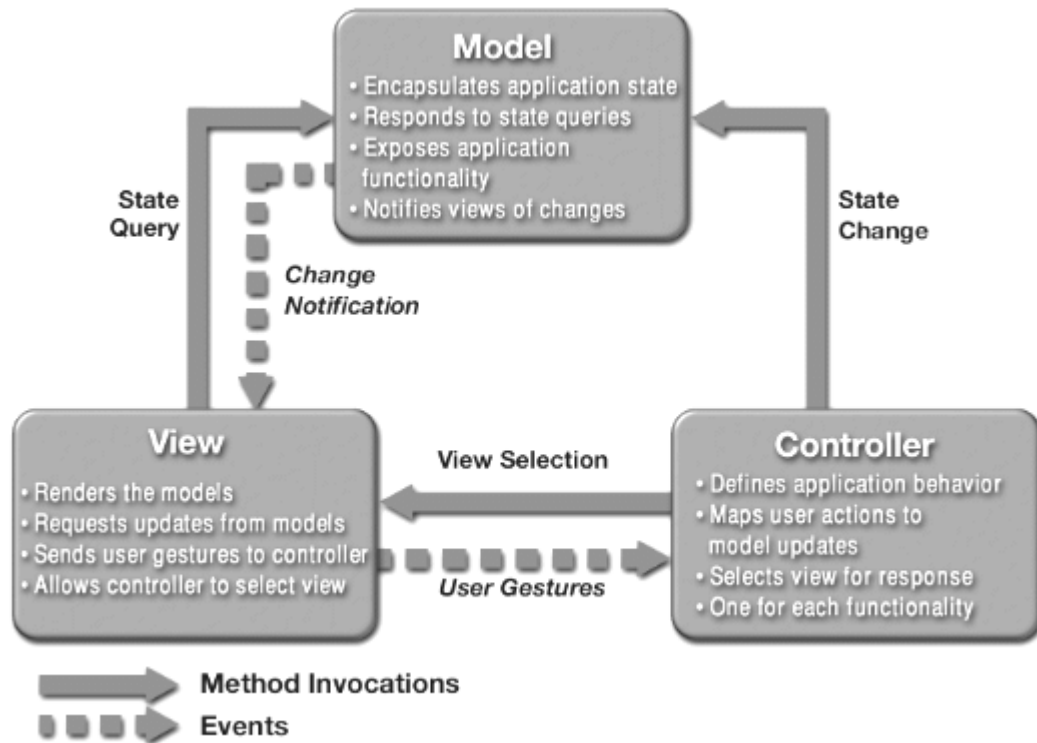
圖 3-26　MVC 概念圖

1. Model：

如同系統中的各個模組，擁有各項的功能，如：拼字檢查、各種文法檢查、同義字檢查…等，當使用者在 View 端呼叫時，Model 便開始執行，並將結果送交給 View；另外，如果在未來需要開發新的模組，也不需要再次更改或編譯主程式，只需要對有更動的模組做編譯即可。

2. View：

簡單來說，View 是用來呈現 Model 的執行結果的，但是在輸出結果之前，結果必須經過 Controller 篩選，再將結果呈現在使用者的面前。

48

3. Controller：

Controller 是一個操控程式，他接收來自使用者的要求，並且支配 Model 的執行和 View 的輸出，以便完成使用者的需求。他沒有做任何輸出或處理，只是決定調用哪個 Model 來處理這個要求，以及使用哪個 View 作輸出的動作而已。

我們決定使用這樣的方式，是因為對於開發系統有以下的優點：

1. 程式碼可重複使用：

當系統更換版面設計，或者被不同的使用者介面（如：WAP、FLASH…等）所要求的時候，並不需要再次撰寫程式碼，只需要使用原本的 Model 來支應這樣的要求即可。

2. 系統開發容易：

由於我們已經將 Controller 和 View 分離，因此當我們更動系統的資料庫，甚至程式結構的時候，我們只需要更改 Model 為相對應的設定即可，此時並不會影響到其他兩者的結構。對於系統開發來說，我們已經將不同層面的開發獨立，而不會相互影響。

四、 系統效能與分析

本研究的系統是基於本實驗室之前所開發之答題驗證系統，但此系統之文法檢查模組在設計的時候，並沒有經過相關的驗證，以及如 CLEC 般清晰而有系統的分類。因此，我們將比較有無 CLEC 驗證之系統的錯誤檢查率。

4.1 系統效能

目前系統是以 CLEC 為基準來設計實作的英語作文檢查系統，前三章已介紹系統的架構與相關模組，本節將呈現系統 CLEC 分類以及錯誤句子的檢查結果比率。

我們計算系統錯誤檢查率的方式，如(1)式所示：

$$錯誤檢查率 = \frac{(CLEC錯誤總句數 - 剩餘錯誤句數)}{CLEC錯誤總句數} \times 100\% \tag{1}$$

表 4-1　系統錯誤檢查率統計表

錯誤類型	剩餘錯誤句數	CLEC 錯誤總句數	錯誤檢查率
詞形	370	10107	96.34%
詞彙	1643	8718	81.15%
句法	1403	7125	80.03%
動詞短語	988	5588	82.32%
名詞短語	483	2957	83.67%
其他	672	3400	80.24%
總計	4941	37895	83.96%
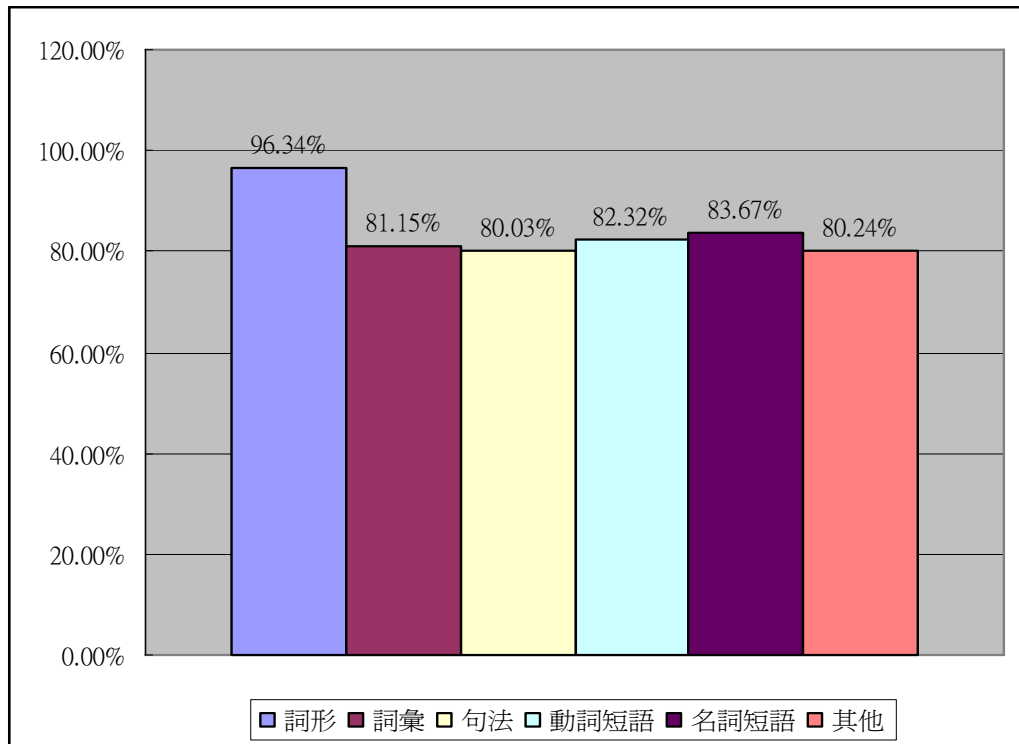
圖 4-1　　系統錯誤檢查率統計圖

　　由圖所示，依照 CLEC 中的錯誤類型分類來撰寫文法檢查模組，並且加上拼字檢查模組檢查的結果，超過百分之八十的錯誤均可以被找出來，同時，詞形的部份更達到百分之九十六。

　　下表列出的文法檢查模組中，包括實驗室之前開發之系統文法模組更新，以及新文法模組的撰寫。

模組名稱	模組功能
COMPARATIVE	比較級的用法
SUPERLATIVE	最高級的用法
PRESENT PARTICIPLE	現在分詞的用法
PAST PARTICIPLE	過去分詞的用法
USE A OR AN	A 和 AN 的用法
MISSING VERBPHRASE	基本動詞使用
SEQUENCE OF TENSES	時態一致
PUNCTUATION	需要正確的標點
REPEAT WORDS	重覆詞性或單字
MIXED CAPITALIZATION	大小寫混雜
AUXILIARY VERBS	助動詞的用法
DOUBLE NEGATIVE	雙重否定句
WORDNET SYNNONYM	同義字檢查
SUBJECT VERB DISAGREEMENT	動詞不一致
ADJECTIVE DISAGREES WITH NOUN	形容詞與名詞不一致
NOT A QUESTION	非疑問句
USE OR INSTEAD OF NOR	or,nor 的使用
USE NOR INSTEAD OF OR	or,nor 的使用
NEED OBJECTIVE PRONOUN	須要正確的代名詞
NEED SUBJECTIVE PRONOUN	須要主格代名詞
NEEDS POSSESSIVE	須要所有格
USE FEWER INSTEAD OF LESS	使用 fewer 代替 less
USE WELL INSTEAD OF GOOD	使用 well 代替 good
PRONOUN NOT REFLEXIVE	代名詞不是反身代名詞
STARTING WITH AND OR BUT	and、or、but 開頭的句子
SENTENCE CAPITALIZATION	句首大寫
PUNCTUATION INSIDE OF QUOTATIONS	引號內少了標點
MISSING QUOTATIONS	遺漏引號

4.2 系統比較

　　本系統的開發，是基於實驗室中之前的答題驗證系統，當時在設計系統時，文法檢查模組的撰寫，並沒有依照一個清楚而完整的錯誤分類去做，因此效能也與基於 CLEC 開發的系統有一些差距，以下的表 4-3 將呈現未經 CLEC 分類之系統檢查率。

表 4-3　　未經 CLEC 分類之系統錯誤檢查率統計表

錯誤類型	剩餘錯誤句數	CLEC 錯誤總句數	錯誤檢查率
詞形	1021	10107	89.90%
詞彙	4136	8718	52.56%
句法	2228	7125	68.73%
動詞短語	2509	5588	55.10%
名詞短語	1357	2957	54.11%
其他	1643	3400	51.68%
總計	12894	37895	65.97%
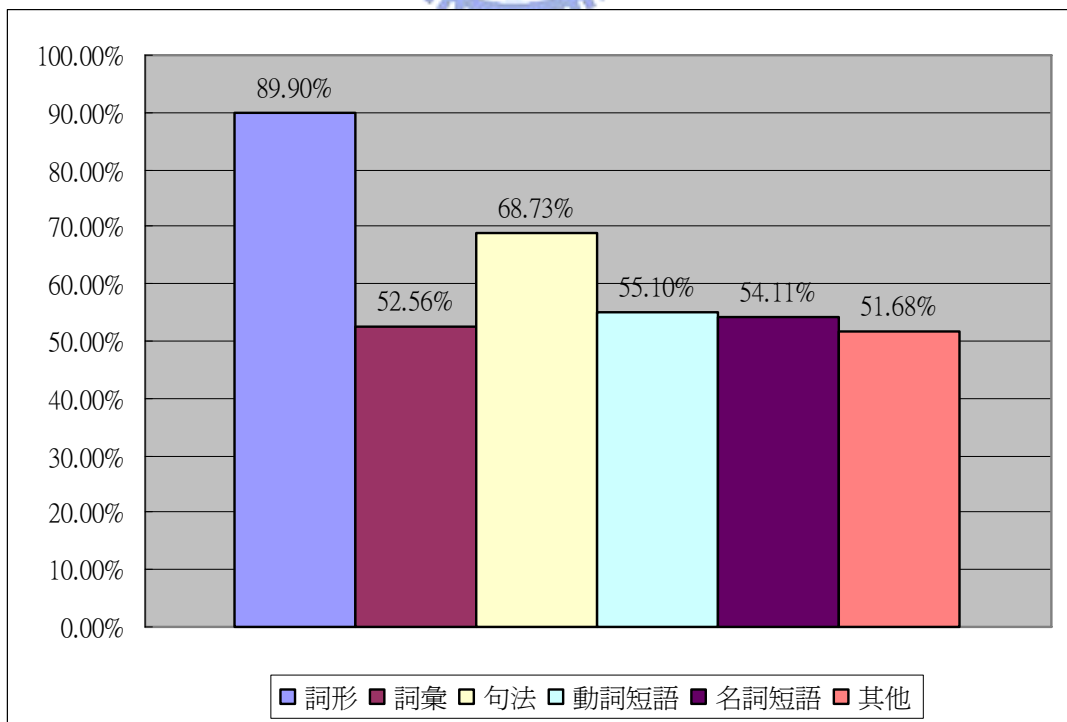
圖 4-2　　未經 CLEC 分類之系統錯誤檢查率統計圖

將上圖的數據，與 4.1 中經 CLEC 分類之系統檢查率比較，我們做成下面的圖表(表 4-4 及圖 4-3)，以便可以更清楚的呈現結果。

表 4-4 中的(1)表示未經分類之系統，(2)則表示本系統。

表 4-4　系統錯誤檢查率比較表

錯誤類型	剩餘句數(1)	剩餘句數(2)	錯誤檢查率(1)	錯誤檢查率(2)	提升比率
詞形	1021	370	89.90%	96.34%	6.44%
詞彙	4136	1643	52.56%	81.15%	28.59%
句法	2228	1403	68.73%	80.31%	20.25%
動詞短語	2509	988	55.10%	82.32%	30.22%
名詞短語	1357	483	54.11%	83.67%	29.56%
其他	1643	672	51.68%	80.24%	28.56%
總計	12894	4941	65.97%	83.96%	19.48%
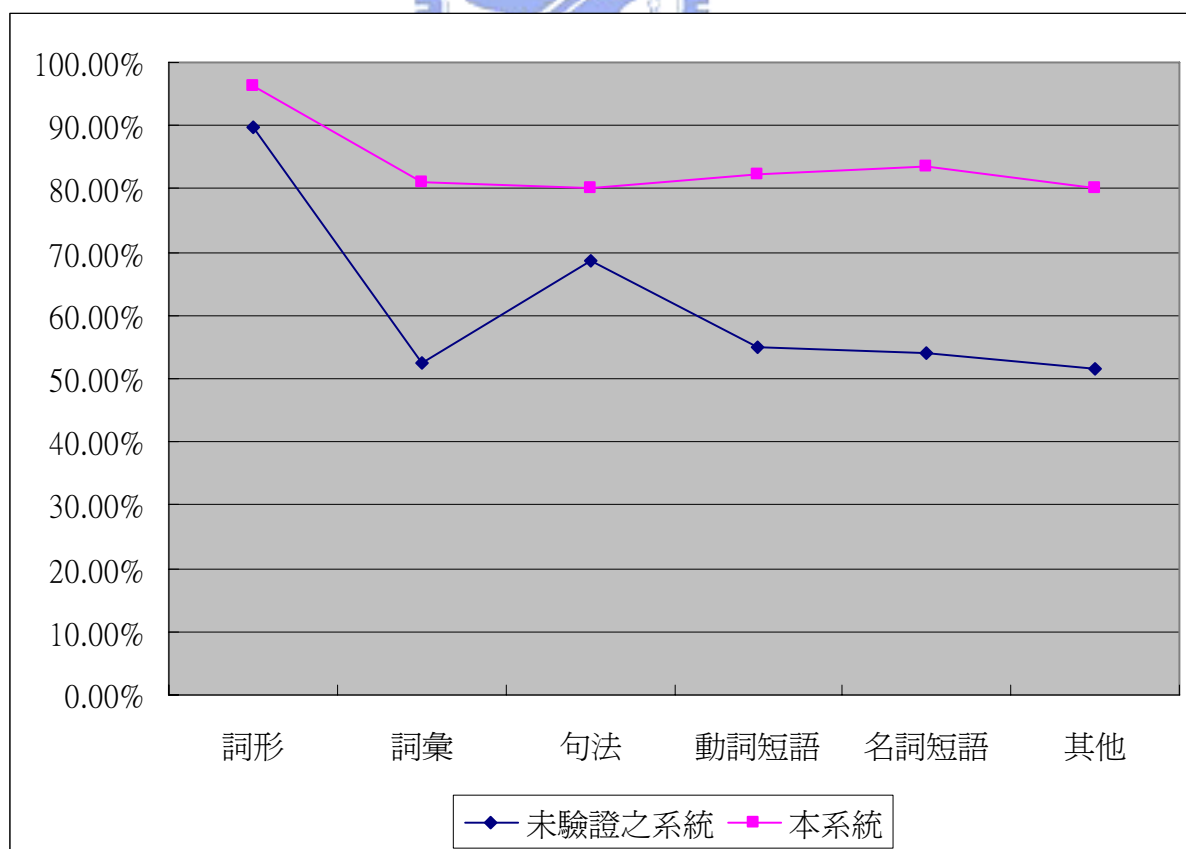
圖 4-3　系統錯誤檢查率比較圖

我們將兩個系統比較之後，做成系統錯誤檢查率比較圖，從圖中可以清楚的看到，關於詞形的錯誤，因為文法檢查模組的新增及加強而有提升，其餘詞彙、句法、動詞短語、名詞短語以及其他類型錯誤，錯誤檢查率皆有相當幅度的提升。

五、 系統實例展示

5.1 系統範例

系統中包含拼字檢查模組、文法檢查模組、評分模組，這一節中將依照順序，以不同的範例，呈現系統運作的狀況。

1. 學習者輸入樣本：
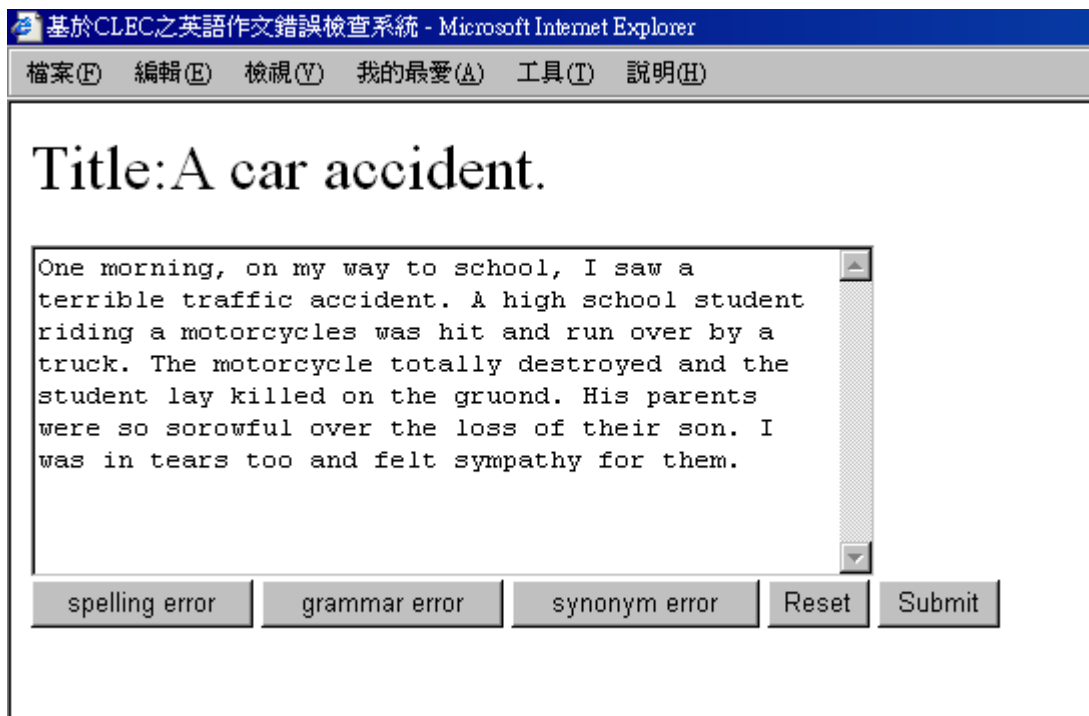
圖 5-1 作文錯誤檢查系統輸入畫面

由此圖可以看到，上方為此次作文的題目「A car accident.」，學習者可以直接將答案填在文字輸入方塊裡面。填答完畢後，便可以按下右下角的 Submit 按鈕送出答案，系統便可以即時予以回饋。

2. 拼字及文法錯誤：

本範例輸入五個句子，在上方「Original Sentences」處將五個句子列出，

圖 5-2　英語作文錯誤檢查系統錯誤輸出畫面

此圖主要分為五個部份，分別說明如下：

(1) 原句列表：

評分模組將原來學習者的文章，分別列成五個句子，讓評分老師能夠清楚的看到原本的句子。

(2) 錯誤統計：

此部份統計整篇文章的字數共有多少，拼字錯誤的字數有多少；也提供文章的總句數，以及發生文法錯誤的句數，讓評分老師直接就可以看出來拼字以及文法的錯誤比例，進而在下一個步驟進行扣分的動作。

(3) 拼字錯誤：

只要句中發生拼字錯誤，Aspell 就會挑出錯誤的字，並且在評分系統中顯示錯誤的單字是在哪個句子當中，同時顯示數個建議字於後方，提供參考。

(4) 文法錯誤：

系統會把文法發生錯誤的句子列出，同時輸出錯誤訊息提供評分老師參考，如果一個句子中發生兩個錯誤，也會將兩個錯誤分別列出。

(5) 評分選項：

當按下 Score 鍵，即代表準備開始評分，評分的畫面將在後面介紹。

3. 評分模組：
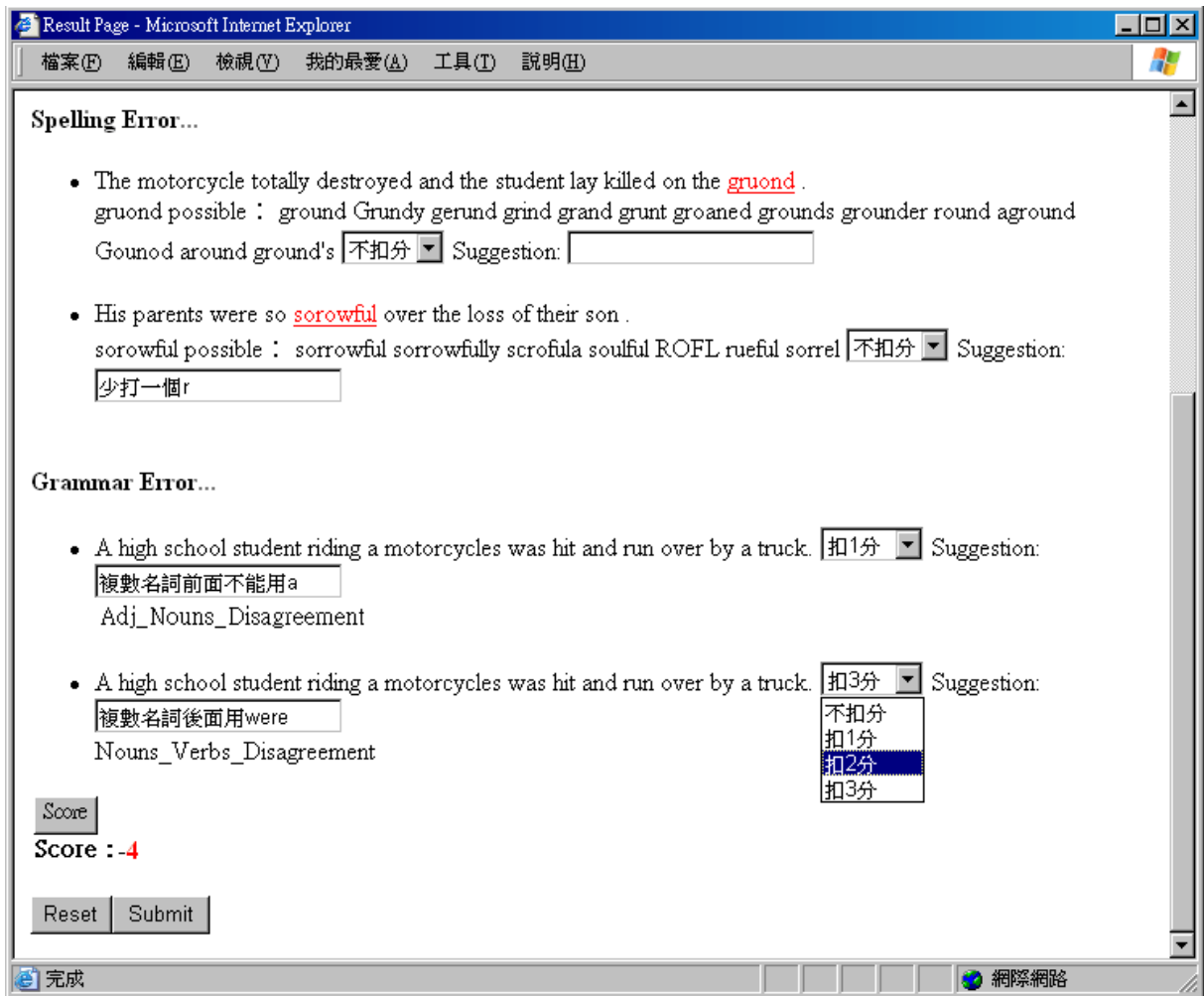
圖 5-3   評分畫面

　　當在上一個步驟按下評分按鈕之後，就會出現可以選擇扣分的下拉式選
單，評分老師可以選擇扣分或者不扣分，最後方則會出現一個文字方塊，提供
老師可以輸入相關的錯誤更正建議，提供學習者正確的改正資訊。
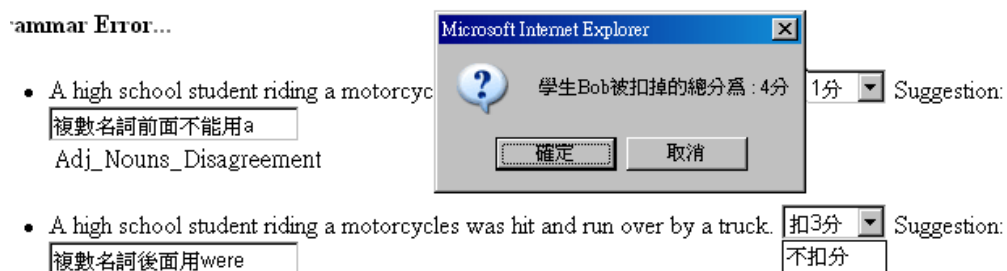
圖 5-4   評分送出畫面

5.2 系統限制

本研究是基於 CLEC 中的錯誤類型，作為文法模組的撰寫依據，因此，驗證範圍主要也以此為主。而本系統未能檢查出來的錯誤，大致可以分為以下三種類型。

1. 人工輸入錯誤：

CLEC 語料庫在開發的時候，純粹是用人工所輸入的，因此錯誤在所難免，經過觀察，有下列幾類錯誤：

(1) 標點錯誤：

這邊所謂的標點並非 CLEC 句中的標點符號，而是用來標記錯誤位置及範圍的括號，由於人工疏失，有些句子中的括號在輸入電腦時發生錯誤，而導致檢查模組產生錯誤的判斷。其餘還有標點疏失，或是標記範圍及位置錯誤，導致系統模組判斷錯誤。

(2) 單字錯誤：

在輸入的時候，本來的樣本是錯誤的，但是由於人工輸入的緣故，在不自覺的情況下直接將之更正，因而發現有些句子並沒有錯，但是卻有錯誤的括號標記在句子當中。

2. 部份文法問題：

例如時態，有些句子在句中以括號標記錯誤位置及範圍，但是實際上整句並沒有錯誤，直到看了前後文句之後才得知，是前後的時態不一致，本系統僅做句中的時態一致性判斷，並沒有檢查前後句時態不一致的情形，因此系統無法判斷出錯誤。

除此之外，由於 APP 在句型正確的時候，解析出來的詞性多半是正確的，但是對於"錯誤"的句型，因為 APP 是使用 bottom-up probabilistic chart 的方式產生詞性的，因此有時候會產生詞性解析錯誤的情形，而導致文法檢查的時候，產生錯誤的判斷。

3. 語意問題：

本研究中在語意的探討，僅侷限在同義字的比對，但真正語言的使用上，語意的問題尚無法偵測，例如：I could study and play in the same time.，此句的錯誤在於介係詞應該是用 at，但是就程式的角度而言，卻因為 in 和 at 都是介係詞，而無法確切的找出錯誤所在。

4. 特定名詞：

例如地名、國名、專有名詞…等。

5.3 論文摘要之測試結果討論

除 CLEC 語料庫外，本系統尚測試實驗室十餘位同學及交通大學數十篇碩博士論文，共計一百篇論文之英文摘要[26]，測試結果如下：

表 5-1 英文摘要測試數據

	拼字檢查		文法檢查
拼字錯誤字數	588	文法錯誤句數	391
總字數	20092	總句數	982

接著我們將以上之拼字及文法錯誤內容，分析如下：

1. 拼字錯誤：

我們觀察系統所找出的錯誤字，主要有兩種：

(1) 學習者對於單字不熟悉而發生錯誤，例如：softwares。

(2) 論文中使用較多專有名詞，但我們知道一般學習者的作文之中，專有名詞的使用頻率並不高。因此，這部份如果需要加強，則可以增加相關專有名詞的字庫予以加強，我們未來展望中也有加以說明。

2. 文法錯誤：

系統測試一百篇碩博士英文摘要後，當中所找出的錯誤數量，如下所示：

表 5-2　　英文摘要錯誤統計表

錯誤類型	數量	百分比
Nouns_Verbs_Disagreement	83	21.23%
mixed capitalization	74	18.93%
More than two verbs in the sentence	63	16.11%
Present participle	41	10.49%
The tense of verbs agreement	34	8.70%
Can't start with XXX	30	7.67%
Past participle	16	4.09%
Doublewords	10	2.56%
Wrong Tense	8	2.05%
Adj_Nouns_Disagreement	8	2.05%
Missing Verb Phrase	7	1.79%
Superlatives or Comparative error	6	1.53%
punctuation error	6	1.53%
Need comma	3	0.77%
Verbs after Infinitive must be present	2	0.51%
	391	
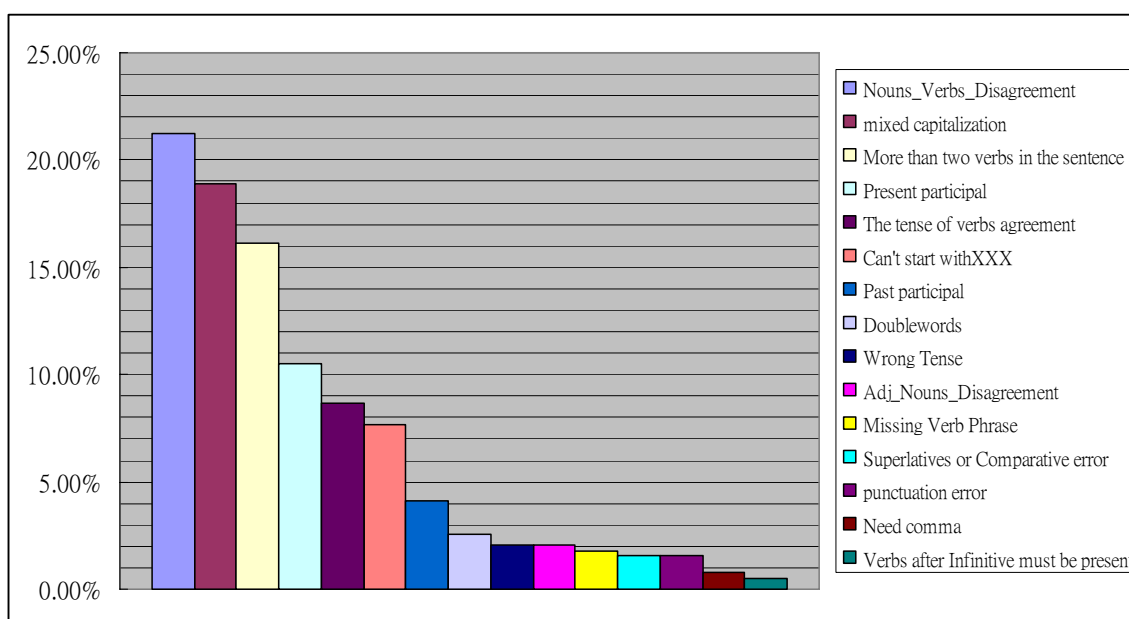
圖 5-5　　英文摘要錯誤統計圖

上面的統計方式，是想要得知各種錯誤類型的數量佔總錯誤數量的比例。我們分析所得到的結果如下。

1. 需加強模組：

依照數量來看，有六種錯誤類型比例較高，可能是需要再加強的，分別是「名詞與動詞不一致」、「句中大小寫混合」、「句中有兩個以上的動詞」、「現在分詞」、「過去分詞」、「動詞時態之一致性」。

2. 錯誤類型：

分析以上的錯誤類型，其中「句中大小寫混合」比例較高的原因，主要也是因為博碩士論文的英文摘要中專有名詞較多而導致的。其餘的錯誤類型，可能要再予以加強。

經過測試一百篇博碩士論文中的英文摘要，我們希望可以藉由這些分析出來得資訊及圖表，得知本系統目前可能還要加強之處，提供作為未來發展的依據及目標。

六、 結論與未來展望

6.1 結論

針對如何實作及設計一個作文錯誤檢查系統，才能有較高的錯誤檢查率，且符合華人的英語使用習慣，我們有著這樣的解決方式。

首先針對較高的錯誤檢查率方面，我們希望能尋找大量的樣本，讓系統可以在新增或更新模組時，運用這些樣本檢測系統的錯誤檢查率。因此，我們試著以語料庫做為我們的檢驗方式。同時，由於英語對我們來說是外來語，語料庫的類型便縮小至學習者語料庫；再者，眾多學習語料庫當中，屬於華人且樣本數量足夠的便是中國學習者英語語料庫(CLEC)了，他不但有著大量的錯誤句型樣本，且類型是學習者的作文，正好符合了我們的需求所在。

因此，我們運用了其錯誤分類，重新設計規劃實驗室系統的文法檢查模組的分類，讓我們能夠新增以及更新已有模組，使得分類更為完善；另外，語料庫的大量樣本是以學習背景分布在高中至大學英語科系的學生，包含範圍也廣，對於系統模組在提升效能的過程當中，測試的錯誤類型多樣化，也較具代表性。

整體而言，不僅是文法檢查模組，對於拼字檢查模組的缺失，部份無法直接由拼字模組檢查出來的錯誤，也在 CLEC 當中有著詳細的分類及錯誤類型，讓我們可以針對那些錯誤樣本，來撰寫相關的文法檢查模組，以便在檢查拼字的時候，補強其不足之處，使得文法模組在 CLEC 當中的錯誤檢查率平均達到百分之八十以上。

最後評分模組的部份，我們讓老師能夠針對每一個拼字錯誤、文法錯誤、以及同義字錯誤做出個別的扣分，同時，也能給予個別錯誤的建議，給予學習者相對應的學習回饋，使學習者在自我練習的時候，能夠更有效。

綜合以上，本系統以 CLEC 為驗證藍本，同時加上實驗室系統為基礎，建立了一個英語作文錯誤檢查系統，不但可讓學習者自我學習，對於系統效能，也做了大幅度的提升。

6.2 未來展望

　　對於本研究中所實作的作文錯誤檢查系統，提出以下兩點作為未來可以繼續研究的方向：

1. 語意研究：

本研究中語意的部份僅做同義字的比對，也就是針對文不對題的情況加以解決，未來如果可以加入相關的語意判斷，例如前後句文意的連貫檢查，或是加入其他的語料庫(如：語意網)做為分析樣本，以便增強目前系統尚未發展完全的功能。

2. 增加學習者適性化學習機制：

學習者在學習英文的時候，必定會遇到許多問題，也會不斷的想要尋求解答，但是對於自己的學習狀況，可能不甚了解；因此，我們可以在系統中加入學習者學習歷程的記錄，記錄下來學習者常犯的拼字或是文法錯誤，並且加以歸類，進而可以做出統計及分析。在下一次學習者進行學習的時候，系統便可以提出建言，使得學習者不至於重蹈覆轍，不但可以做到適性化學習，同時也真正能加強學習者的英文寫作能力。

3. 增加其他華人語料庫：

未來可以增加更多的華人語料庫到目前系統中，使目前系統的錯誤檢查率可再提升。

參考文獻

[1] Centre for English Corpus Linguistics，[On-line]. Available:
http://www.fltr.ucl.ac.be/FLTR/GERM/ETAN/CECL/cecl.html

[2] GNU Aspell，[On-line]. Available: http://aspell.net/

[3] GNU Emacs，[On-line]. Available: http://www.gnu.org/software/emacs/

[4] GNU 計畫，自由軟體基金會，劉昭宏譯，[On-line]. Available:
http://www.gnu.org/brave-gnu-world/issue-44.zh.html

[5] International Corpus of Learner English，[On-line]. Available:
http://cecl.fltr.ucl.ac.be/Cecl-Projects/Icle/icle.htm

[6] International Ispell，[On-line]. Available:
http://fmg-www.cs.ucla.edu/geoff/ispell.html

[7] Java BluePrints，[On-line]. Available:
http://java.sun.com/blueprints/patterns/MVC-detailed.html

[8] Learner Corpora and SLA Research，[On-line]. Available:
http://leo.meikai.ac.jp/~tono/index.html

[9] Leech, G. (1998b). Preface, in Granger, S. (1998). Learner English on
Computer, XVII, Essex: Addison Wesley Longman Limited.

[10] Longman Corpus Network，[On-line]. Available:
http://www.longman.com/dictionaries/corpus/learners.html

[11] Mitchell P.Marcus, Mary Ann Marcinkiewicz, and Beatrice Santorini
(1993). "Building a Large Annotated Corpus of English: The Penn
Treebank", Computational Linguistics, 19(2)

[12] Proteus Project-Apple Pie Parser，[On-line]. Available:
http://nlp.cs.nyu.edu/app/

[13] Smalltalk.org，[On-line]. Available: http://www.smalltalk.org

[14] Spell-checking With Opera，[On-line]. Available:
http://www.opera.com/support/tutorials/opera/spellcheck/index.dml

[15] WordNet，[On-line]. Available: http://wordnet.princeton.edu/

[16] XML 台灣資訊網，[On-line]. Available:
http://www.xml.org.tw/Function/Fglossary1.asp?key=XML

[17] 九年一貫課程與教學網，[On-line]. Available:

http://teach.eje.edu.tw/9CC/fields/2003/language05-source.php

[18] 李佩瑛，數位島與電子報第七期，[On-line]. Available:

http://content.ndap.org.tw/main/epaper_detail.php?doc_id=1046&epaper _id=21

[19] 桂詩春、楊惠中，中國學習者英語語料庫，上海外語教育出版社，上海，2002。

[20] 財團法人語言訓練測驗中心，[On-line]. Available: http://www.gept.org.tw/

[21] 國內語料庫建設一覽表，[On-line]. Available:

http://blog.csdn.net/yujun00/archive/2005/12/01/541633.aspx

[22] 張鈞凱，「英語文試題檢測與答題驗證系統設計與實作」，國立交通大學，碩士論文，民 94 年。

[23] 黃希敏，敦煌英語教學電子雜誌，[On-line]. Available:

http://cet.cavesbooks.com.tw/htm/m0131100.htm

[24] 黃希敏，敦煌英語教學電子雜誌，[On-line]. Available:

http://cet.cavesbooks.com.tw/htm/m050307.htm

[25] 楊惠中、衛乃興，中國學習者英語口語語料庫建設與研究，上海外語教育出版社，上海，2005。

[26] 樣版國度，個人部落格，[On-line]. Available:

http://web.caidiy.com/plate/web/mediamsg.jsp?UI=falldog&MI=4232

附錄

姓名：柯政邦

論文名稱：多媒體內容在手持行動裝置之情境及內容協調與呈現

Original Sentences：

The growing computing power of mobile devices supports multimedia contents presentation on these devices.

If we can present our multimedia contents in various mobile devices, it will be another business opportunity and urges programming designers to development more software for mobile devices.

But in general, most multimedia contents are designed for PC, so they can only be supported on PC.

For content providers, they often have to redesign presentation style and implementation those multimedia contents which are only supported on PC.

This task will take lots of manpower and times and causing a great burden for the content providers.

Reducing the cost of manpower and times of translation between PC and mobile devices is the main purpose of this research.

We use the context awareness algorithm to analyze each element that will affect the presentation on mobile devices.

One of the most important things in this research is how to adapt non-template based multimedia materials of mobile devices without changing their scenario.

Base on the research in《The Context Aware and Content Adaptation of Template Based Multimedia Presentation on Handset Device - Using Template Based English Test Questions with Multimedia Contents as Examples》, we use a mixture of static and dynamic method to adapt materials.

Improving the XMG player with some new functions; we expect the result makes user's learning easier, comfortable and more effective.

	Spelling Check		Grammar Check
Number of Spelling Errors	1	Number of Grammar Errors	8
Number of Words	217	Number of Sentences	10

Spelling Error...

- Improving the <u>XMG</u> player with some new functions ; we expect the result makes user

XMG possible： MG XML MSG MAG MEG MIG MUG SMOG SMUG XMAS
MC MK SM XX SIG SAG SM'S

Grammar Error...

- The growing computing power of mobile <u>devices</u> supports multimedia contents presentation on these devices .
  Nouns_Verbs_Disagreement

- <u>But</u> in general , most multimedia contents are designed for PC , so they can only be supported on PC .
  [fm1] Can't start with "But" .

- For content providers , they often have to redesign presentation style and implementation those multimedia contents which <u>are</u> only supported on PC .
  The tense of verbs on both sides of "and" should be identical.

- For content providers , they often have to redesign presentation style and implementation those multimedia contents which are only supported on <u>PC</u> .
  [fm3] --> PC mixed capitalization

- This task will take lots of manpower and times and <u>causing</u> a great burden for the content providers .
  [VP6] [Waring] -->Wrong Tense......About Participals......at VBG

- Reducing the cost of manpower and times of translation between PC and mobile devices is the main purpose of this research .
  "PC" This is not proper noun or mixed capital in words.

- One of the most important things in this research is how to adapt non-template based <u>multimedia</u> materials of mobile devices without changing their scenario .
  Nouns_Verbs_Disagreement

- Base on the research <u>in《The</u> Context Aware and Content Adaptation of Template Based Multimedia Presentation on Handset Device : Using Template Based English Test Questions with Multimedia Contents as Examples》 , we use a mixture of static and dynamic method to adapt materials .
  [fm3] --> in《The mixed capitalization

姓名：翁浚恩

論文名稱：設計及製作一個介面產生器用以產生手機內的 JAVA 應用軟體之介面設定使其遙控 PC 上相同之應用軟體

Original Sentences：

　　With the quick advance of technology, screen display of digital TV and mobile system becomes more and more elegant and is able to present fine and vivid multimedia contents.

　　Most of the multimedia contents, such as advertisement, motion pictures, messages, etc., can be displayed on different kinds of platforms.

　　If user can use some simple instruments (such as cell phone, PDA, etc.) to remotely communicate with the multimedia application module in the display device (such as PC monitor, digital TV, etc.), then the control becomes live and interesting.

　　But there are various control instruments and display devices, and different kinds of control methods.

　　If one wants to write the control program or partially modify the control features for the multimedia application module in the display device, then he needs to know the software source code in the multimedia application module that will be remotely controlled, so that he can custom-design a set of remote control programs for each multimedia application.

　　But since there is a large number of multimedia applications, such a custom design becomes time consuming and less efficient.

　　This research proposes an interface generator, similar to the function of the parser generator in the programming code, to automatically generate remote control programs in the cell phone for a specific multimedia application.

　　With this generator, designer does not need to know and write the remote control programs in the cell phone.

　　This will simplify the development procedures and make the control system development and modification more flexible.

　　This research utilizes this interface generator and its algorithm to produce three different kinds of application softwares in the cell phone.

　　These three softwares can remotely communicate with similar application softwares in PC.

　　With this practical example, we demonstrate the feasibility and application of this interface generator and its algorithm.

　　We also quantitatively assess the difference between this interface generator plus its algorithm and the conventional method.

	Spelling Check		Grammar Check
Number of Spelling Errors	3	Number of Grammar Errors	8
Number of Words	319	Number of Sentences	13

Spelling Error...

- This research utilizes this interface generator and its algorithm to produce three different kinds of application <u>softwares</u> in the cell phone .
  softwares possible： soft wares soft-wares software's software softwoods softeners stares softies softwood's wetwares softener's bloatwares leftwards softballs softie's stoneware's flatware's stemware's softball's

- These three <u>softwares</u> can remotely communicate with similar application <u>softwares</u> in PC .
  softwares possible： soft wares soft-wares software's software softwoods softeners stares softies softwood's wetwares softener's bloatwares leftwards softballs softie's stoneware's flatware's stemware's softball's

- These three <u>softwares</u> can remotely communicate with similar application <u>softwares</u> in PC .
  softwares possible： soft wares soft-wares software's software softwoods softeners stares softies softwood's wetwares softener's bloatwares leftwards softballs softie's stoneware's flatware's stemware's softball's

Grammar Error...

- <u>With</u> the quick advance of technology , screen display of digital TV and mobile system becomes more and more elegant and is able to present fine and vivid multimedia contents .
  [fm1] Can't start with "With" .

- Most of the multimedia contents , such as advertisement , motion pictures , messages , etc. , can be <u>displayed</u> on different kinds of platforms .
  [VP6] [Waring] -->Wrong Tense......About Participals......at VBN

- <u>But</u> there are various control instruments and display devices , and different kinds of control methods .
  [fm1] Can't start with "But" .

- If one wants to write the control program or partially modify the control features for the <u>multimedia</u> application module in the display device , then he needs to know the software source code in the multimedia application module that will be remotely controlled , so that he can custom-design a set of remote control programs for each multimedia application .
  Nouns_Verbs_Disagreement

- <u>But</u> since there is a large number of multimedia applications , such a custom design becomes time consuming and less efficient .
  [fm1] Can't start with "But" .

- This research proposes an interface generator , similar to the function of the parser generator in the programming code , to automatically generate remote control <u>programs</u> in the cell phone for a specific multimedia application .
  Nouns_Verbs_Disagreement

- <u>With</u> this generator , designer does not need to know and write the remote control programs in the cell phone .
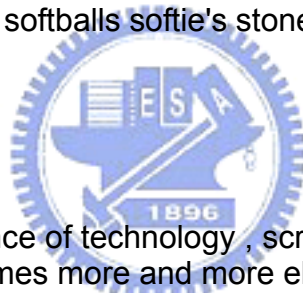 [fm1] Can't start with "With" .

- <u>With</u> this practical example , we demonstrate the feasibility and application of this interface generator and its algorithm .
  [fm1] Can't start with "With" .

姓名：黃承一
論文名稱：以視覺化使用者介面建構方法論應用於儲存系統之人機介面設計及實作

Original Sentences：

　　When the requirements of Man Machine Interface (MMI) are changed, programmers must change the related programs of MMI if they use the conventional MMI development.

　　Therefore, they have to take long time to develop and spend much effort to maintain programs of MMI.

　　The Software Engineering Laboratory of NCTU had developed Visual-Based User Interface Construction Methodology.

　　This methodology can improve the productivity, quality, and maintainability of MMI software.

　　In this thesis, we use this methodology for the MMI design and implementation of storage systems, and try to conquer UI problems using conventional MMI development.

　　For the typical and various Man Machine Interfaces of storage systems, we design a Visual MMI Development for Storage Systems.

　　It includes a Visual Authoring Tool and Generic MMI Engine for storage systems.

　　In order to demonstrate the feasibility of the Visual MMI Development for Storage Systems, we implement a software framework with Generic MMI Engine that can manipulate the output data of Visual Authoring Tool, interact with the Management API of storage systems, and use this system to control a functional storage adapter.

　　A real application example is applied using the proposed approach to demonstrate the applicability of the methodology.

	Spelling Check		Grammar Check
Number of Spelling Errors	11	Number of Grammar Errors	2
Number of Words	197	Number of Sentences	9

Spelling Error...

- Therefore , they have to take long time to develop and spend much effort to maintain programs of <u>MMI</u> .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- The Software Engineering Laboratory of <u>NCTU</u> had developed Visual-Based User Interface Construction Methodology .
  NCTU possible： NC TU NC-TU CUT NUT UNCUT CT NC NT NEUT ACT NCO NWT NAT OCT JCT NET NIT NOT PCT NATO NCAA NATA NATE NITA NETT NOTE

- This methodology can improve the productivity , quality , and maintainability of <u>MMI</u> software .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- In this thesis , we use this methodology for the <u>MMI</u> design and implementation of storage systems , and try to conquer <u>UI</u> problems using conventional <u>MMI</u> development .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- In this thesis , we use this methodology for the <u>MMI</u> design and implementation of storage systems , and try to conquer <u>UI</u> problems using conventional <u>MMI</u> development .
  UI possible： II I U UPI URI UZI UNI AI UR OI UH DUI GUI HUI SUI UK UL UN US UT UV UM UP BI CI DI GI HI KI LI MI NI RI SI TI VI WI PI XI

- In this thesis , we use this methodology for the <u>MMI</u> design and implementation of storage systems , and try to conquer <u>UI</u> problems using conventional <u>MMI</u> development .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- For the typical and various Man Machine Interfaces of storage systems , we design a Visual <u>MMI</u> Development for Storage Systems .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- It includes a Visual Authoring Tool and Generic <u>MMI</u> Engine for storage systems .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- In order to demonstrate the feasibility of the Visual <u>MMI</u> Development for Storage Systems , we implement a software framework with Generic <u>MMI</u> Engine that can manipulate the output data of Visual Authoring Tool , interact with the Management <u>API</u> of storage systems , and use this system to control a functional storage adapter .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- In order to demonstrate the feasibility of the Visual <u>MMI</u> Development for Storage Systems , we implement a software framework with Generic <u>MMI</u> Engine that can manipulate the output data of Visual Authoring Tool , interact with the Management <u>API</u> of storage systems , and use this system to control a functional storage adapter .
  MMI possible： MI MM MIMI MAI MEI MME MOI EMMI AMI MCI MRI

- In order to demonstrate the feasibility of the Visual <u>MMI</u> Development for Storage Systems , we implement a software framework with Generic <u>MMI</u> Engine that can manipulate the output data of Visual Authoring Tool , interact with the Management <u>API</u> of storage systems , and use this system to control a functional storage adapter .
  API possible：APO APIA AP AI PI UPI APE APP APB APR APT ADI ALI AMI ARI CPI

Grammar Error...

- When the requirements of Man Machine Interface CD : MMI CD : are changed , programmers must change the related programs of MMI if they use the conventional MMI development .
punctuation error…

- A real application example is applied using the proposed approach to demonstrate the applicability of the methodology .
[VBG] and [VBN] are combined.

姓名：楊博鈞

題目：互動式多媒體簡訊播放編輯器與傳遞機制在手持行動裝置之呈現-以樣板式多媒體簡訊應用為例子

Original Sentences：

Interactive multimedia applications are getting ever more popular nowadays, there are lots of applications available on the PC, but mobile handsets applications are limited and also not as popular as those for PC.

Currently the most popular application market for mobile handsets is games and entertainment.

Although the MMS message facility on mobile handsets has a multimedia effect, the only facility beyond simple communication is to allow the user to insert or delete as basic editing.

It does not allow the user to change the look and feel of the design, nor the interactive functionality.

It is certainly not easy to write a program to achieve such behaviour on mobile handsets.

But such an application; one with which user could create and edit content on a handset via interactive multimedia templates and which can then be sent on to another mobile recipient, would allow service providers to best utilise and promote the potential of multimedia messaging services.

This research proposes to design the template and transmission facility, which will allow, with small effort, straightforward modification on the handset application in order to create a user personalized multimedia message.

The focus on a simplified editing process, achieved by a well planned and stable GUI (Graphical User Interface), should attract many more users to the specific mobile handset and/or service operator who offers this user friendly interactive messaging facility.

The objective of this research is to create and set-up an actual working editing tool and demonstrate functionality on a mobile handset.

Also to instruct on use of the editing tool on the handset in order to create a demonstrative range of content, including, personalized greeting cards, multimedia messages, teaching materials and test questions, etc. The template system will allow users to change any of these designs directly on the handsets, the user can then preview the messages on the handset before transmission.

Finally it will deliver the multimedia message to another user who also has same application operating on their mobile handset.

	Spelling Check		Grammar Check
Number of Spelling Errors	1	Number of Grammar Errors	5
Number of Words	333	Number of Sentences	11

Spelling Error...

- Although the <u>MMS</u> message facility on mobile handsets has a multimedia effect , the only facility beyond simple communication is to allow the user to insert or delete as basic editing .
MMS possible： MMES MS MAMS MOMS MUMS MM M'S MAS MES MOS MUS MYS FMS HMS MME MRS PMS EMS MKS OMS MN'S MAM'S MOM'S MUM'S MME'S ME'S MO'S MY'S WM'S MA'S MI'S MU'S AM'S CM'S EM'S FM'S MD'S MG'S MR'S MT'S PM'S SM'S TM'S HM'S

Grammar Error...

- Interactive multimedia applications are getting ever more popular nowadays , there are lots of applications available on the PC , but mobile handsets applications are limited and also not as popular as those for PC .
"PC" This is not proper noun or mixed capital in words.

- Currently the most popular application market for mobile <u>handsets</u> is games and entertainment .
Nouns_Verbs_Disagreement

- It is certainly not easy to write a program to achieve such behaviour on mobile <u>handsets</u> .
Nouns_Verbs_Disagreement

- <u>But</u> such an application CD one with which user could create and edit content on a handset via interactive multimedia templates and which can then be sent on to another mobile recipient , would allow service providers to best utilise and promote the potential of multimedia messaging services .
[fm1] Can't start with "But" .

- The focus on a simplified editing process , achieved by a well planned and stable GUI JJ : Graphical User Interface JJ : , should attract many more users to the specific mobile handset and/or service operator who offers this user friendly interactive messaging facility .
"GUI" This is not proper noun or mixed capital in words.

作者：鄭學隆

論文題目：面對面英語口試和視訊英語口試的差異研究-以國中三年級學生為研究對象

Original Sentences：

The main purpose of this research is to investigate if there is a significant difference between the ways of using oral test instruments; namely, face-to-face approach and video-conferencing approach.

The 9th grade high school students are targeted as the subject from this study.

In general, using face-to-face oral test requires support from many test people, a lot of test rooms, and a lot of time for a medium or large scale English oral tests.

Therefore, we study the related theories about English oral tests and think that internet conference might solve the problem with real time multi online function.

And also we think about adopting the oral tests in English in junior high level so as to promote the ability to master a language.

We have 63 sample students which are in the 9th grade in a certain junior high school in Taipei county.

All the students were tested first in face-to-face English oral test, and in video-conferencing English oral test at a later time.

We collect the data acquired in the process of the tests and the questionnaires filled out by the sample students.

And we also analyze qualitatively the attitude toward and operations in taking the face-to-face and video-conferencing English oral tests.

In the research we find that, the students' scores in (1) accuracy, (2) fluency, and (3) communicativity are quite different after taking the video-conferencing English oral tests, and the average scores are also up.

The possible reasons may be as follows: the ability to acticulate is somewhat born by nature and it is hard to change, so are the resulting scores.

And most students feel less nervous in taking video-conferencing English oral test than in taking face-to-face one and thus it increases the total scores.

Before taking the video-conferencing English oral tests, we have an interesting phenomenon that the girl students' scores in both fluency and communicativity are better than the boy students' scores.

But, on the contrary, after taking the video-conferencing English oral tests, the boys' scores in both fluency and communicativity are better than the girl students' scores.

The possible reasons may be as follows: the acceptability to the boy students to the English oral tests is better than that to the girl students.

So we can think about putting English oral tests in the school syllabus to elicit the genuine ability of mastering English

In according to the result of the questionnaire, the majority of students have a positive attitude toward taking the video-conferencing English oral tests.

The way multimedia shows is helpful to the students to answer the questions.

The function of allowing the testees to listen to the questions repeatedly is especially helpful to them to answer the questions after a thoughtful thinking

and reduce their exam nerves so as to elicit the students' genuine ability in English.

Due to the limitation of the time, room, and people, we just use the sample students in the 9th grade in junior high school.

But the techniques to elicit the ability in English isvarious, whether we could get other quite different results or not by means of arranging various test items in the English oral tests is an open issue and is a topic reserved to be discussed.

	Spelling Check		Grammar Check
Number of Spelling Errors	10	Number of Grammar Errors	9
Number of Words	534	Number of Sentences	21

Spelling Error...

- The 9 th grade high school students are targeted as the subject from this study .
  th possible： Th Thu the tho thy THC Rh H T h t nth Ch OH TA Ta Te Ti Tu Ty ah ch eh oh pH sh ta ti to uh Th's

- Therefore , we study the related theories about English oral tests and think that internet conference might solve the problem with real time multi online function .
  internet possible： Internet inter net inter-net interned Internets intent internee interment intranet entente intern internist indent intend interne's

- Therefore , we study the related theories about English oral tests and think that internet conference might solve the problem with real time multi online function .
  multi possible： Miltie Milt milt moult Milty malt melt molt mufti Malta malty muled mullet mulatto mulct mild mould Melita ult Milli Multan milts moults maltier Mali Mulder mauled melter molter mouldy mule mull mulled mute mutt cult must meld mold Marti Matti Melli Molli malts melts molts Milt's milt's malt's

- Therefore , we study the related theories about English oral tests and think that internet conference might solve the problem with real time multi online function .
  online possible： on line on-line aniline incline Olin unlined Ernaline Opaline offline outline Aline Arline oiling unlink engine ogling unaligned unlike Olen enliven onion inkling Nolan alien newline Alaine Elaine Ilene Orlon alone inane owning snowline Onegin mainline nonlinear Angeline Nolana airline only Arliene angling ankling enplane uncling Alane Alene Alina Allin Arlin Elane Elene Orlan Earline aniline's O'Neil

- We have 63 sample students which are in the 9 <u>th</u> grade in a certain junior high school in Taipei county .
  th possible： Th Thu the tho thy THC Rh H T h t nth Ch OH TA Ta Te Ti Tu Ty ah ch eh oh pH sh ta ti to uh Th's

- The possible reasons may be as follows : the ability to <u>acticulate</u> is <u>somewhow</u> born by nature and it is hard to change , so are the resulting scores .
  acticulate possible： articulate inarticulate ejaculate articled actuality

- The possible reasons may be as follows : the ability to <u>acticulate</u> is <u>somewhow</u> born by nature and it is hard to change , so are the resulting scores .
  somewhow possible： somehow somewhat somewhere someway someone some Simeon

- The function of allowing the <u>testees</u> to listen to the questions repeatedly is especially helpful to them to answer the questions after a thoughtful thinking and reduce their exam nerves so as to elicit the students
  testees possible： testers testes tester's tests tasters testes's tastes testis taste's taster's testis's

- Due to the limitation of the time , room , and people , we just use the sample students in the 9 <u>th</u> grade in junior high school .
  th possible： Th Thu the tho thy THC Rh H T h t nth Ch OH TA Ta Te Ti Tu Ty ah ch eh oh pH sh ta ti to uh Th's

- But the techniques to elicit the ability in English <u>isvarious</u> , whether we could get other quite different results or not by means of arranging various test items in the English oral tests is an open issue and is a topic reserved to be discussed .
  isvarious possible： is various is-various usurious ovaries ivories estrous Alvaro's Isidro's aviaries oestrous Osiris Severus safaris savories Esmaria's inferiors avarice ovary's safari's Isidoro's estrus Osiris's Ivory's ivory's Astaire's Isadora's Isadore's

Grammar Error...

- In general , using face-to-face oral test requires support from <u>many</u> test people , a lot of test rooms , and a lot of time for a medium or large scale English oral tests .

  [np3] Adj is not agrees with Noun

- <u>And</u> also we think about adopting the oral tests in English in junior high level so as to promote the ability to master a language .

  [fm1] Can't start with "And" .

- <u>And</u> we also analyze qualitatively the attitude toward and operations in taking the face-to-face and video-conferencing English oral tests .

  [fm1] Can't start with "And" .

- <u>And</u> most students feel less nervous in taking video-conferencing English oral test than in taking face-to-face one and thus it increases the total scores .

  [fm1] Can't start with "And" .

- Before <u>taking</u> the video-conferencing English oral tests , we have an interesting phenomenon that the girl students' scores in both fluency and communicativity are better than the boy students' scores .

  [VP6] [Waring] -->Wrong Tense......About Participals......at VBG

- <u>But</u> , on the contrary , after taking the video-conferencing English oral tests , the boys' scores in both fluency and communicativity are better than the girl students' scores .

  [fm1] Can't start with "But" .

- The possible reasons may be as follows : the acceptability to the boy <u>students</u> to the English oral tests is better than that to the girl students .

  Nouns_Verbs_Disagreement

- So we can think about <u>putting</u> English oral tests in the school syllabus to elicit the genuine ability of mastering English

  [VP6] [Waring] -->Wrong Tense......About Participals......at VBG

- The way multimedia <u>shows</u> is helpful to the students to answer the questions .

  Nouns_Verbs_Disagreement

姓名：謝佳成

論文名稱：多媒體講解呈現播放器與 **PDA** 行動裝置之設計及製作

Original Sentences：

The popularity of network infrastructure, including wired、wireless network resulting ubiquitous networking and ubiquitous learning environment.

Electronic Learning has been developed prosperously toward Mobile Learning.

By the convenience of network infrastructure, Internet resource can be easily accessible and available.

Such that learners can use PDA via Internet infrastructure to acquire the multimedia lecturing learning materials for reading and learning in order to achieve the effect of mobile learning.

Multimedia lecturing applied in learning material is attractive and easily produced.

In general, these materials are usually produced by more computing power PC.

As for PDA readers, content providers have to produce multimedia lecturing learning material for PDA version and that increases the content provider extra overhead.

Learners can use the existing PC multimedia lecturing learning material player to read and learn.

In this research, we discuss the mechanism of multimedia lecturing learning content adaptation from PC to PDA such that multimedia lecturing learning materials from PC can be read and learned from PDA by smallest effort.

In this research, we implement a player of multimedia lecturing learning material for PDA version.

We port the PC player program into PDA player program.

During the porting, we discuss the difference of system resource, program grammar, and supportive libraries between PC and PDA, then compare those items and implement the PDA player.

Make the multimedia lecturing learning material from PC can also be read from PDA, expectedly, achieve the goal of ubiquitous learning material and ubiquitous learning.

	Spelling Check		Grammar Check
Number of Spelling Errors	11	Number of Grammar Errors	4
Number of Words	242	Number of Sentences	13

Spelling Error...

- Such that learners can use PDA via Internet infrastructure to acquire the multimedia lecturing learning materials for reading and learning in order to achieve the effect of mobile learning .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- As for PDA readers , content providers have to produce multimedia lecturing learning material for PDA version and that increases the content provider extra overhead .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- As for PDA readers , content providers have to produce multimedia lecturing learning material for PDA version and that increases the content provider extra overhead .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- In this research , we discuss the mechanism of multimedia lecturing learning content adaptation from PC to PDA such that multimedia lecturing learning materials from PC can be read and learned from PDA by smallest effort .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- In this research , we discuss the mechanism of multimedia lecturing learning content adaptation from PC to PDA such that multimedia lecturing learning materials from PC can be read and learned from PDA by smallest effort .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- In this research , we implement a player of multimedia lecturing learning material for PDA version .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- We port the PC player program into PDA player program .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- During the porting , we discuss the difference of system resource , program grammar , and supportive libraries between PC and PDA , then compare those items and implement the PDA player .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- During the porting , we discuss the difference of system resource , program grammar , and supportive libraries between PC and PDA , then compare those items and implement the PDA player .

PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- Make the multimedia lecturing learning material from PC can also be read from PDA , expectedly , achieve the goal of ubiquitous learning material and ubiquitous learning .
  PDA possible： PAD PD DA PA PTA PDQ PDT PIA PEA ADA EDA FDA IDA RDA PD'S

- Make the multimedia lecturing learning material from PC can also be read from PDA , expectedly , achieve the goal of ubiquitous learning material and ubiquitous learning .
  expectedly possible： expected unexpectedly expectantly excitedly expertly

Grammar Error...

- The popularity of network infrastructure , including wired、wireless network resulting ubiquitous networking and ubiquitous learning environment .
  [wd4] Missing Verb Phrase error.

- By the convenience of network infrastructure , Internet resource can be easily accessible and available .
  "Internet" This is not proper noun or mixed capital in words.

- In general , these materials are usually produced by more computing power PC .
  [fm3] --> PC mixed capitalization

- Learners can use the existing PC multimedia lecturing learning material player to read and learn .
  "PC" This is not proper noun or mixed capital in words.

姓名：簡國安

論文題目：符合 **SCORM RTE1.3.3** 課程架構樣板與教學策略樣板之設計與製作

Original Sentences：

With the progress of multimedia digital technology, the production of digital teaching contents has been improved in terms of amount and quality.

However, each variety of digital teaching contents has been applied with only one single teaching strategy and offers learners only one way of learning.To change the teaching strategies of digital teaching contents, we find it difficult to solve the problems of platforms.

Luckily, ADL has given the definitions as to which kind of teaching contents meet the demands of SCORM, separating teaching contents from teaching contents control, and uses the format of XML ,which is of cross-platform ,to describe the navigation method we see in the teaching materials that matches the Sequencing Rules, so that the revision of the teaching strategies(of teaching contents) has become much easier.

Digital curriculums are almost belong to programmed instruction， In order to accomplish the purpose of programmed instruction,we need aggregate structure of Contents Aggregation from teaching contents first of all.Our aims are to present the shared teaching contents according to the needs of teaching methods and to reuse the teaching contents by bringing in different teaching strategies according to the needs of the teachers so as to meet the demands of individualized teaching and learning.

However, on the internet as learning platform, ADL has not offered the necessary computer programs that can be used to develop teaching strategies---To make our teaching contents more elastic and flexible, the producers of teaching contents are supposed to revise easily their teaching strategies on the same piece of teaching material.

This thesis aims to design and give the performance of a model of curriculum frameworkand a model of teaching strategy to solve the problems of on-line combined curriculum.The characteristics of the system are as followed:

1.The system here will be able to aide the curriculum designer to produce fast a curriculum framework of nine different teaching strategies.

2.The system offers a function for the latter curriculum editor to reproduce the built framework by the former curriculum editor so that the framework can be revised and reused.

3.The system offers the latter curriculum editor to design the curriculum framework from the beginning to the end, completely without the use of models or the formerly-done curriculum to do the work of revision and editing.

4.The system offers the teachers to use fast and easily the shared teaching materials and to settle on certain teaching strategies to build the curriculum of their own so as to meet the demands of teaching.

5.The system offers to combine the teaching materials for the curriculum and, based on the learning conditions and the test results of the students, give guidance so that the proper teaching materials are used and the proper learning is completed.

6.The teachers will be able to send , revise ,and recombine the teaching

materials and the test questions of different units through the system here so that the material will comply with the standards of SCORM1.3.3 .

	Spelling Check		Grammar Check
Number of Spelling Errors	7	Number of Grammar Errors	2
Number of Words	411	Number of Sentences	12

Spelling Error...

- Luckily , ADL has given the definitions as to which kind of teaching contents meet the demands of SCORM , separating teaching contents from teaching contents control , and uses the format of XML , which is of cross-platform , to describe the navigation method we see in the teaching materials that matches the Sequencing Rules , so that the revision of the teaching strategies
ADL possible： ADEL DAL ASL AD AL ADD AOL ADA ADE ADI ADO AIL ALL AWL ADC ADP ADM ADJ ADS ADV ADZ AD'S

- Luckily , ADL has given the definitions as to which kind of teaching contents meet the demands of SCORM , separating teaching contents from teaching contents control , and uses the format of XML , which is of cross-platform , to describe the navigation method we see in the teaching materials that matches the Sequencing Rules , so that the revision of the teaching strategies
SCORM possible： SCORN CORM SCAM SCUM STORM SCORE SCRIM SCHEMA SCHEME SCOUR SCRAM SCRUM SQUIRM SCOT COM SKIM STORMI STORMY SCORCH SCORER SCOURS COMM SCAR SCOW SCUMMY

- Digital curriculums are almost belong to programmed instruction curriculums possible： curriculum's curriculum curricle's curricula curricular

- However , on the internet as learning platform , ADL has not offered the necessary computer programs that can be used to develop teaching strategies---To make our teaching contents more elastic and flexible , the producers of teaching contents are supposed to revise easily their teaching strategies on the same piece of teaching material .
internet possible： Internet inter net inter-net interned Internets intent internee interment intranet entente intern internist indent intend interne's

- However , on the internet as learning platform , ADL has not offered the necessary computer programs that can be used to develop teaching strategies---To make our teaching contents more elastic and flexible ,

the producers of teaching contents are supposed to revise easily their teaching strategies on the same piece of teaching material .
ADL possible： ADEL DAL ASL AD AL ADD AOL ADA ADE ADI ADO AIL ALL AWL ADC ADP ADM ADJ ADS ADV ADZ AD'S

- This thesis aims to design and give the performance of a model of curriculum frameworkand a model of teaching strategy to solve the problems of on-line combined curriculum . The characteristics of the system are as followed :
frameworkand possible： framework and framework-and frameworks framework framework's

- 6 . The teachers will be able to send , revise , and recombine the teaching materials and the test questions of different units through the system here so that the material will comply with the standards of SCORM 1 . 3 . 3 .
SCORM possible： SCORN CORM SCAM SCUM STORM SCORE SCRIM SCHEMA SCHEME SCOUR SCRAM SCRUM SQUIRM SCOT COM SKIM STORMI STORMY SCORCH SCORER SCOURS COMM SCAR SCOW SCUMMY

Grammar Error...

- With the progress of multimedia digital technology , the production of digital teaching contents has been improved in terms of amount and quality .
[fm1] Can't start with "With" .

- However , each variety of digital teaching contents has been applied with only one single teaching strategy and offers learners only one way of learning.To change the teaching strategies of digital teaching contents , we find it difficult to solve the problems of platforms .
[fm3] --> learning.To mixed capitalization

姓名：李清峰

論文題目：操作型試題樣版系統設計與實作

Original Sentences：

   Most of the existing Web based testing systems can not fit the Operation Style tests very well.
   In this paper, we propose the framework of Operation Style Item Templates. With applying the Templates, test makers can edite test questions fast.
   In this paper, we emphasise how to develop operation style item templates system which makes the designers easier to produce the suitable operation style item templates.
   Finally, we design a web based testing system which makes it possible to go on these tests smoothly on this system.

	Spelling Check		Grammar Check
Number of Spelling Errors	1	Number of Grammar Errors	3
Number of Words	87	Number of Sentences	5

Spelling Error...

- With applying the Templates , test makers can <u>edite</u> test questions fast .
  edite possible： edit Edita editor edited EDT Edithe Edie edits Edith elite eddied erudite diet Eadie Eddie Eydie dote edict Deity Ede Edi Ettie deity idiot Idette Odette mediate Dedie ETD Tedie Dita Edee Odie date audit indite it'd Edie's Edita's

Grammar Error...

- Most of the existing Web based testing systems can not fit the Operation Style tests very well .
  [NPL] + VBN

- In this paper , we emphasise how to develop operation style item templates system which makes the designers easier to produce the suitable operation style item templates .
  [wd4] Missing Verb Phrase error.

- Finally , we design a web based testing system which makes it possible to go on these tests smoothly on this system .
  [wd4] Missing Verb Phrase error.

姓名：鄧明芳

論文名稱：**The Design and Implementation of Operation-Style Item Template for Test Items Creation**

Original Sentences：

Assessment testing is a method to know what does student learn in learning process.

We often request student doing something to know what he learns, especially in measurement, geometry and statistics.

Existing online assessment system only provides basic item types, such as true false, choice and essay.

For the operation style item, those online item types can\'t meet the demands.

Furthermore, the authoring of operation style item is difficult for teachers in general.

This paper is to develop skill item template helping teacher to author operation style item.

Through operation style item template, teachers can author operation style item more quickly.

It is possible for each teacher to author operation style item with operation style item template.

	Spelling Check		Grammar Check
Number of Spelling Errors	2	Number of Grammar Errors	2
Number of Words	117	Number of Sentences	8

Spelling Error...

- Existing <u>online</u> assessment system only provides basic item types , such as true false , choice and essay .
  online possible： on line on-line aniline incline Olin unlined Ernaline Opaline offline outline Aline Arline oiling unlink engine ogling unaligned unlike Olen enliven onion inkling Nolan alien newline Alaine Elaine Ilene Orlon alone inane owning snowline Onegin mainline nonlinear Angeline Nolana airline only Arliene angling ankling enplane uncling Alane Alene Alina Allin Arlin Elane Elene Orlan Earline aniline's O'Neil

- For the operation style item , those <u>online</u> item types can't meet the demands .
  online possible： on line on-line aniline incline Olin unlined Ernaline Opaline offline outline Aline Arline oiling unlink engine ogling unaligned unlike Olen enliven onion inkling Nolan alien newline Alaine Elaine Ilene Orlon alone inane owning snowline Onegin mainline nonlinear Angeline

Nolana airline only Arliene angling ankling enplane uncling Alane Alene Alina Allin Arlin Elane Elene Orlan Earline aniline's O'Neil

Grammar Error...

- Furthermore , the authoring of operation style item is difficult for teachers in general .
  Nouns_Verbs_Disagreement

- Through operation style item template , teachers can author operation style item more quickly .
  Verbs after Infinitive must be present.

姓名：連瑞斌

Original Sentences：

The Arrival of third industrial revolution in the twenty first century, led by the internet and information telecommunication technology, has increased high demand for multimedia learning contents .

For both the academic community and industry, it is foremost important to develop high efficiency and low cost methodologies to design high quality MLCs.

To raise the quality level of MLCs, we need to find out the related factors in order to design an efficient quality control mechanism.

This thesis propose two related factors for quality issues: whether the script design of the MLCs fulfills the learning objective, and whether it is attractive enough; in the MLCs developing process, various changes of original requirements or designs may be continuously requested from either inner or exterior customers; if there is a lack of effective control and management mechanisms, there may be causing quality problems.

At present, the design of MLCs usually follows international standards such as the SCORM, but this does not guarantee high-quality MLCs.

Therefore, this research uses the CMMI for Software Engineering developed by the software engineering institute of Carnegie Mellon University, in order to develop a MLCs Quality Control Model .

This MQC model includes an inconsistency detection mechanism in the MLCs developing process, so as to avoid quality problems due to continuous changes in design under a collaboration environment.

The feasibility and compatibility of the proposed methodology will be proved through a client-server system implementation that integrated into MLCs quality control system.

	Spelling Check		Grammar Check
Number of Spelling Errors	13	Number of Grammar Errors	0
Number of Words	241	Number of Sentences	8

Spelling Error...

- The Arrival of third industrial revolution in the twenty first century , led by the <u>internet</u> and information telecommunication technology , has increased high demand for multimedia learning contents .
  internet possible： Internet inter net inter-net interned Internets intent internee interment intranet entente intern internist indent intend interne's

- For both the academic community and industry , it is foremost important to develop high efficiency and low cost methodologies to design high

quality <u>MLCs</u> .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- To raise the quality level of <u>MLCs</u> , we need to find out the related factors in order to design an efficient quality control mechanism .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- This thesis propose two related factors for quality issues : whether the script design of the <u>MLCs</u> fulfills the learning objective , and whether it is attractive enough ; in the <u>MLCs</u> developing process , various changes of original requirements or designs may be continuously requested from either inner or exterior customers ; if there is a lack of effective control and management mechanisms , there may be causing quality problems .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- This thesis propose two related factors for quality issues : whether the script design of the <u>MLCs</u> fulfills the learning objective , and whether it is attractive enough ; in the <u>MLCs</u> developing process , various changes of original requirements or designs may be continuously requested from either inner or exterior customers ; if there is a lack of effective control and management mechanisms , there may be causing quality problems .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- At present , the design of <u>MLCs</u> usually follows international standards such as the <u>SCORM</u> , but this does not guarantee high-quality <u>MLCs</u> .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- At present , the design of <u>MLCs</u> usually follows international standards such as the <u>SCORM</u> , but this does not guarantee high-quality <u>MLCs</u> .
SCORM possible： SCORN CORM SCAM SCUM STORM SCORE SCRIM SCHEMA SCHEME SCOUR SCRAM SCRUM SQUIRM SCOT COM SKIM STORMI STORMY SCORCH SCORER SCOURS COMM SCAR SCOW SCUMMY

- At present , the design of <u>MLCs</u> usually follows international standards such as the <u>SCORM</u> , but this does not guarantee high-quality <u>MLCs</u> .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- Therefore , this research uses the <u>CMMI</u> for Software Engineering developed by the software engineering institute of Carnegie Mellon University , in order to develop a <u>MLCs</u> Quality Control Model .
CMMI possible： CM MI CM-MI CAMMI CAMI EMMI

- Therefore , this research uses the <u>CMMI</u> for Software Engineering developed by the software engineering institute of Carnegie Mellon University , in order to develop a <u>MLCs</u> Quality Control Model .
MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- This MQC model includes an inconsistency detection mechanism in the MLCs developing process , so as to avoid quality problems due to continuous changes in design under a collaboration environment .
  MQC possible： MC QC MAC MIC

- This MQC model includes an inconsistency detection mechanism in the MLCs developing process , so as to avoid quality problems due to continuous changes in design under a collaboration environment .
  MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

- The feasibility and compatibility of the proposed methodology will be proved through a client-server system implementation that integrated into MLCs quality control system .
  MLCs possible： ML Cs ML-Cs Macs Mics Mac's Mic's

姓名：林桂盟

Original Sentences：

Digital Signage is a new developing medium, spreading multimedia information through large flat-panel displays in public.

And now, the most used application of "Digital Signage"is DSA .

My research implemented a playback module of a commercial DSA system.

And we want to extend the system by expanding the format of the DSA system.

There were two ways to extend the DSA system in my research: Look for the suitable format, and support the edit solution for it to reduce the manufacturing cost.

Program the new module by dispatching players of the expanded format.

Therefore we don't need to waste time to program all the player of these formats.

	Spelling Check		Grammar Check
Number of Spelling Errors	3	Number of Grammar Errors	1
Number of Words	106	Number of Sentences	7

Spelling Error...

- My research implemented a playback module of a commercial <u>DSA</u> system .
  DSA possible： SSA DA SA DEA DOA DST ASA BSA DNA GSA ISA USA D'S

- And we want to extend the system by expanding the format of the <u>DSA</u> system .
  DSA possible： SSA DA SA DEA DOA DST ASA BSA DNA GSA ISA USA D'S

- There were two ways to extend the <u>DSA</u> system in my research : Look for the suitable format , and support the edit solution for it to reduce the manufacturing cost .
  DSA possible： SSA DA SA DEA DOA DST ASA BSA DNA GSA ISA USA D'S

Grammar Error...

- Digital Signage is a new developing medium , spreading <u>multimedia</u> information through large flat-panel displays in public .
  Nouns_Verbs_Disagreement