

本期要目

壹. ROCLING-2008 Call For Participation	第二頁
貳. 出席 LREC2008 心得分享 (鄭秋豫)	第三頁
參. 詞彙標示框架簡介 (黃居仁)	第四~六頁
肆. 專文-電腦輔助語言學習 (蔡明峰、廖元甫)	第七~十六頁

博碩士論文獎 7/1 開始申請

名額及獎項：

1. 博士論文優等獎：一名，獎金二萬元，並頒給學生及指導教授獎狀各乙紙。
2. 博士論文佳作獎：一名，獎金一萬元，並頒給學生及指導教授獎狀各乙紙。
3. 碩士論文優等獎：一名，獎金一萬元，並頒給學生及指導教授獎狀各乙紙。
4. 碩士論文佳作獎：三名，獎金各伍仟元，並頒給學生及指導教授獎狀各乙紙。

申請資格及限制：

1. 國內大專院校博碩士班應屆畢業生從事計算語言學相關研究方向者，由其指導教授推薦。
2. 參賽限制：每位指導教授以推薦一篇博士論文及兩篇碩士論文為限。(含個人指導與共同指導)。

申請期間：7/1~7/31，申請程序及辦法請參閱本會網站：<http://www.aclclp.org.tw/doc/shipreg.htm>。

獎助學生出席國際會議

- 會議名稱：COLING -2008
- 論文題目：OntoNotes: Corpus Cleanup of Mistaken Agreement Using Word Sense Disambiguation
- 獎助學生：禹良治 (成功大學資工所博士生)
- 獎助金額：US\$1,000 元

ROCLING-2008

專題演講主講人背景：

1. **林偉 Lam, Wai** (香港中文大學系統工程與工程管理學系教授)

Lam, Wai received a Ph.D. in Computer Science from the University of Waterloo. He obtained his BSc. and M.Phil. degrees from the Chinese University of Hong Kong. After completing his Ph.D. degree, he conducted research at Indiana University Purdue University Indianapolis (IUPUI) and University of Iowa. He joined the Chinese University of Hong Kong, where he is currently a professor.

His research interests include intelligent information retrieval, text mining, digital library, machine learning, and knowledge-based systems. He has published articles in IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Knowledge and Data Engineering, ACM Transactions on Information Systems, etc. <http://www.se.cuhk.edu.hk/people/wlam.html>

2. **林欽佑 Lin, Chin-Yew** (微軟亞洲研究院主任研究員)

Lin, Chin-Yew is a lead researcher and research manager at Microsoft Research Asia. Before joining Microsoft at the end of February 2006, he was a senior research scientist at the Information Sciences Institute at University of Southern California (USC/ISI) where he worked in the Natural Language Processing and Machine Translation group since 1997. His research interests are automated summarization, opinion analysis, question answering (QA), computational advertising, community intelligence, machine translation (MT), and machine learning. <http://research.microsoft.com/~cyl/>

Conference on Computational Linguistics and Speech Processing

第二十屆自然語言與語音處理研討會, ROCLING XX

National Taiwan Normal University, Taipei, Taiwan

4-5 September 2008

The 20th ROCLING Conference will be held at National Taiwan Normal University, Taipei, on September 4-5, 2008. Sponsored by Association for Computational Linguistics and Chinese Language Processing (ACLCLP), ROCLING is the most historied and major conference in the broad field of computational linguistics, speech processing, and related areas in Taiwan.

ROCLING XX will be hosted by the Information Technology Center and Department of Computer Science and Information Engineering, National Taiwan Normal University. The two-day conference will feature invited talks, paper, and poster sessions.

ROCLING XX invites submissions of original and unpublished research papers on all areas of computational linguistics, natural language processing, and speech processing, including, but not limited to, the following topic areas.

Topics of Interest:

- cognitive/psychological linguistics
- discourse/dialogue modeling
- information extraction/text mining
- information retrieval
- language understanding/generation
- lexicon/morphology
- machine translation/multilingual processing
- named entity recognition
- NLP applications/tools/resources
- phonetics/phonology
- question answering
- semantic web
- semantics/pragmatics
- speech analysis/synthesis
- speech recognition/understanding
- spoken dialog systems
- spoken language processing
- syntax/parsing
- text/speech summarization
- web knowledge discovery
- word segmentation/POS tagging
- others

Important Dates:

Submission Deadline (Extension) : July 14, 2008

Notification of Acceptance: August 11, 2008

Camera-ready Due: August 20, 2008

Conference Date: September 4-5, 2008

Submission Guidelines:

Prospective authors are invited to submit full papers of no more than 25 A4- sized pages in PDF format (please go to the conference website to download the paper guideline and template). Papers will be accepted only by electronic submission through the conference website. The submitted papers should be written in either Chinese or English, and in single column, single-spaced format. The first page of the submitted paper should bear the items of paper title, author name, affiliation, and email address. All these items should be properly centered on the top, followed by an concise abstract of the paper.

Papers should be made in PDF format and submitted online at:

<http://forum.dmc.ntnu.edu.tw/rocling2008/>

出席第六屆語言資源與評估國際學術會議 LREC2008 心得分享

鄭秋豫

中央研究院語言學研究所

語言資源與評估國際學術會議 LREC(6th Conference on Language Resources & Evaluation) 係全球性國際學術會議，每二年舉行一次，此次在摩洛哥舉行的 LREC 2008 <http://www.lrec-conf.org/lrec2008/> 為第六屆會議，主辦單位是位於法國巴黎的 ELRA (European Language Resource Association) 及義大利 Piza 國家研究中心。此次會議共有 645 篇論文，1111 人註冊參加。主要會議是 5 月 28-30 日，會前二日 (5 月 26、27) 是專題課程 tutorial，會後二日 (5 月 31 日、6 月 1 日) 是工作坊 workshop。

我很榮幸自 2006 年起擔任 Oriental-COCOSDA <http://www.slt.atr.jp/o-cocosda> 召集人。Oriental-COCOSDA 的母組織 COCOSDA (The International Committee for the Co-ordination and Standardization of Speech Databases and Assessment Techniques) 成立於 1994 年，亞洲地區有鑑於此一區域的語言在類型學上大異於歐美非洲語言，且需大力推廣與口語語料的收集及建立平台相關的研究，在 1998 年由日本發起，成立亞洲地區區域性國際學術組織 Oriental-COCOSDA。該組織設召集人一人，並有國家及地區代表二至三人，至今共有十三個國家及地區，按英文字母順序分別為：中國、香港、印度、印尼、日本、韓國、馬來西亞、蒙古、尼泊爾、台灣、泰國、越南、巴基斯坦。

COCOSDA 每二年均在 LREC 會議集會，Oriental-COCOSDA 召集人援例需代表 Oriental 地區所有國家及代表出席會議。此次 COCOSDA 會議安排在會議最後一天 6 月 1 日的 COCOSDA/Write Workshop，由我做會務報告 15 分鐘，題目是「Asian Activities on Speech Corpora」，並參加有關世界性語言資源開發、共享、評估之研討。雖然集會安排在會議的最後一天，仍有將近五十餘人出席 COCOSDA/Write Workshop，第一場次 COCOSDA session 共有 7 個報告，分別是 1. 台灣的鄭秋豫報告 Oriental COCOSDA，2. 澳洲的 Steve Cassidy 報告大洋洲 Oceania，3. Justus Roux 報告非洲 Africa，4. Chris Cieri 報告北美 North America，5. Khalid Choukri 報告歐洲 Europe，6. Ksenia Shalnova 報告在地資源，7. Nick Campbell 報告語音合成。隨後是 WRITE-FLaReNet session，由 Khalid Choukri 代表 ELRA 和 Chris Cieri 代表 LDC 報告，並全場討論。下一場是 Nancy Ide (Vassar College, USA) 和 Edouard Geoffrois (南美) 報告並討論。最後一場則由美國賓州大學 Alan Lee 及 Aravind Joshi、Johns Hopkins 大學的 Fred Jelinek 及東京大學的 Junichi Tsujii 報告並討論。這個工作坊對於文本和語音的合作，評估平台的建置都有具體討論的結果。對於所有參與討論的學者，都有實質的意義。鄭秋豫將綜合討論的結果，於 2008 年 11 月將舉行的 Oriental COCOSDA2008 國際會議中，向亞洲地區代表作口頭及書面報告。

此次出席 LREC2008，除參加上述 COCOSDA/Write Workshop 外，我並於 5 月 29 日於主要會議中發表論文一篇，題目是「The 2008 Oriental COCOSDA Book Project – In Commemoration of the First Decade of Sustained Activities in Asia (with Shuichi Itahashi)」；應邀於會期第一天 5 月 28 日擔任口頭報告場次 LRs: Infrastructures, Projects, Centers 主席；應邀參加 5 月 30 日的大會專題小組「The Role and Mission of Data Centers, Cost Principles in the Creation and Sharing of Language Resources: Why there is no such thing as a free lunch?」討論，為四位引言人之一，引言報告 15 分鐘，題目為「Asian Activities-The Oriental COCOSDA Community」。

此次會議在北非摩洛哥的 Marrakech 舉行。該城市建設進步，雖無大眾運輸工具，卻有國際水準的會議中心，可同時容納數千人，四周環繞五、四、三星級旅館，距離機場六公里，交通便捷。本次會議僅見到台大資訊系陳信希教授，台灣學者僅二人參加，大陸學者人數也很少。但日本學者很多，韓國學者一人，其他與會學者則來自全世界。我除了代表亞洲地區 13 個國家外，由於曝光率很高，也全程參與了會議及相關活動，因此特別凸顯了來自台灣身份。

詞彙標示框架簡介

黃居仁

中央研究院語言學研究所

詞彙標示框架(Lexical Markup Framework, 簡稱 LMF)是國際標準組織(ISO/TC37)進行中的一項工作,目的在為自然語言處理(NLP)與機讀字典(MRD)的詞彙庫描述建立一個標準化框架。計畫範疇涵蓋牽涉到多語溝通及文化差異的語言資源,對建立與交換這些資源的準則與方法做標準化處理。

詞彙標示框架的目標

詞彙標示框架的目標有三。其一,為詞彙資源的創造與使用提供共用模型。其二,管理詞彙資源間的資料交換。其三,促進個別電子資源的整合以形成大規模的全球性電子資源。

詞彙標示框架的種類包括單語、雙語或多語的詞彙資源。這三種分類亦適用於小型或大型詞彙庫、簡單或複雜詞彙庫乃至於書面或口語詞彙表述。說明的範疇包含構詞學、句法學、計算語意學及電腦輔助翻譯。涵蓋的語言包括所有自然語言,並不侷限於歐洲地區。此計畫在自然語言處理的運用上不受限制。詞彙標示框架能呈現多數辭典,包括 WordNet, EDR 及 PAROLE。

詞彙標示框架的歷史

在過去,已有不少計畫做過詞彙庫標準化的研究,如 GENELEX、EDR、EAGLES、MULTEXT、PAROLE、SIMPLE 與 ISLE。其後,參與 ISO/TC37 的國家代表決定為自然語言處理與詞彙庫呈現訂定標準。2003 年暑假,由美國代表提出的詞彙標示框架計畫誕生。2003 年秋,法國代表提出定義資料模型的技術主張,以俾利做自然語言處理的辭典。到了 2004 初,ISO/TC37 委員會再決定做一項共同 ISO 計畫,由 Nicoletta Calzolari (義)擔任召集人, Gil Francopoulo (法)與 Monte George (美)擔任編輯。迄今此計畫建立出的詞彙標示框架已有 13 個版本,並已分送給各國指定的專家做評論,再經過 ISO 技術會議的討論修正。

現況

這項計畫的 ISO 標準號碼為 24613。到 2007 年 12 月,詞彙標示框架尚處於國際標準草稿(DIS)階段。預定在 2008 年 2 月份進入國際標準最終草稿階段(FDIS),並於當年 9 月定版發行。

隸屬 ISO/TC37 家族的一支

ISO/TC37 標準家族目前被定位為高階指定規格,處理包括斷詞(ISO 24614)、標註(ISO 24611 亦稱為 MAF、ISO 24612 即 LAF、ISO 24615 亦即 SynAF 與 ISO 24617-1 又名 SemAF/Time)、特徵結構(ISO 24610)、多媒體載體(ISO 24616 亦稱為 MLIF)以及辭典(ISO 24613)等內容。這些標準主要建立在提供共同術語的低階指定規格,亦即資料範疇匯總(ISO 12620 的修定版)、語言編碼(ISO 639)、書寫體編碼(ISO 15924)、國家編碼(ISO 3166)與通用編碼(ISO 10646)。

這個兩個階層架構的標準遵循下列兩項簡單共通準則,以構成具內部一致性的標準家族:

- 高階指定規格提供了上層的結構框架,其內容將由標準化常數元素充實之。
- 低階指定規格提供標準化常數元素與後設資料。

基本標準

語言學的常數元素如/女性/或/及物/在詞彙標示框架中並沒有被規範，但被記錄在資料範疇匯總(DCR)中。資料範疇匯總是一項全球性資源，由 ISO/TC37 負責維持，依照 ISO/IEC 11179-3:2003。這些常數元素要用來充實高階結構。

詞彙標示框架遵守由物件管理組(OMG)規範的統一建模語言(UML)。結構由 UML 的階級圖表呈現，例子由 UML 的例子圖表或物件圖表呈現。

需要補充說明是 XML 的檔案類別定義(DTD)是由詞彙標示框架的附件所規範。

模型結構

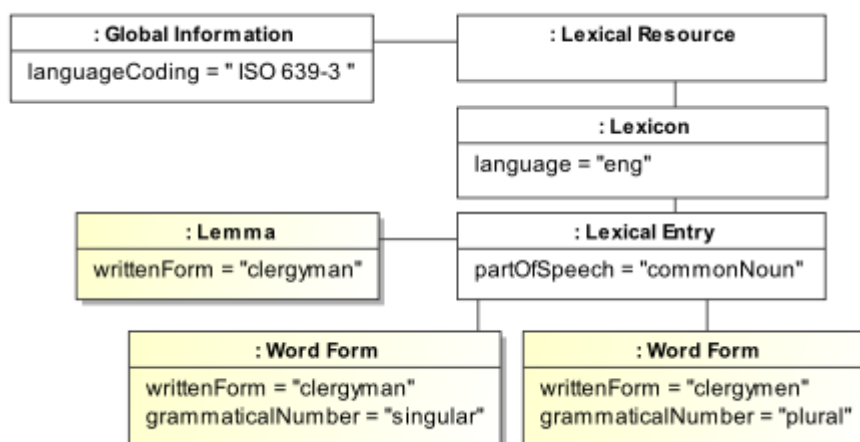
詞彙標示框架包含下列組成成分：

- 核心封包：也就是描述詞條資訊之基本階級的結構。
- 核心封包的擴充：它以框架表達，該框架描述核心組成成分的再利用，此核心組成成分與特定詞彙資源所需的額外組成成分相連接。

這些擴充乃特別為下列領域而設置，包含構詞學、機讀字典、自然語言處理句法學、自然語言處理語意學、多語標記、自然語言處理典型模式、多字詞表達模式與限制表達模式。

簡例

在下面的舉例中，詞目 *clergyman* 以及兩個詞尾變化形 *clergyman* 跟 *clergymen* 都與詞條有關。語言編碼套用於整個詞彙資源，而語言特性的值則適用於整個詞彙庫，正如下方 UML 圖所示。



詞彙資源(Lexical Resource)、全球資訊(Global Information)、詞彙庫(Lexicon)、詞條(Lexical Entry)、詞目(Lemma)以及詞形(Word Form)這幾個元素定義出詞彙庫的結構。在詞彙標示框架的檔案中有對這幾項元素做說明。另一方面，語言編碼(languageCoding)、語言(language)、詞類(partOfSpeech)、一般名詞(commonNoun)、書寫形(writtenForm)、文法號碼(grammaticalNumber)、單數(singular)、多數(plural)等則是從資料類別登記中取出的資料種類。這些標示能充實結構內容。而 ISO 639-3、clergyman、clergymen 這幾項值都只是基本的特性字串。eng 這個值則參照 ISO 639-3 所定義的語言清單。

再加上一些附加資訊如 *dtdVersion* 與 *feat* 後，同樣的資料可以以下列 XML 段落表示：

```
<LexicalResource dtdVersion="14">
  <GlobalInformation>
    <feat att="languageCoding" val="ISO 639-3"/>
  </GlobalInformation>
  <Lexicon>
    <feat att="language" val="eng"/>
    <LexicalEntry>
      <feat att="partOfSpeech" val="commonNoun"/>
      <Lemma>
        <feat att="writtenForm" val="clergyman"/>
      </Lemma>
      <WordForm>
        <feat att="writtenForm" val="clergyman"/>
        <feat att="grammaticalNumber" val="singular"/>
      </WordForm>
      <WordForm>
        <feat att="writtenForm" val="clergymen"/>
        <feat att="grammaticalNumber" val="plural"/>
      </WordForm>
    </LexicalEntry>
  </Lexicon>
</LexicalResource>
```

要注意的是，這個例子相當簡單。但詞彙標示框架能夠呈現更為複雜的語言描述，XML 標記也會更加複雜。

外部連結

- 相關網站
 - LMF web site
 - LIRICS web site
- 最新科學通訊
 - Gesellschaft für linguistische Datenverarbeitung GLDV-2007/Tubingen: Lexical Markup Framework ISO standard for semantic information in NLP lexicons
 - Language Resources and Evaluation LREC-2006/Genoa: Lexical Markup Framework (LMF)
- 相關科學通訊
 - Language Resources and Evaluation LREC-2006/Genoa: The relevance of standards for research infrastructures

參見

- 構詞學
- 機器翻譯，有關於幾種多語標記的資訊(於方法段)。
- 詞網(WordNet)，最著名的英語語意辭典。

電腦輔助語言學習

蔡明峰 廖元甫
台北科技大學

1. 簡介

口語除了是人與人之間最方便的溝通工具，也是某一國家地區風土人情的直接體現，因此要深入了解某一外國文化的最好方法，就是親自學習好某一外國語言，直接了解他們的想法，並與其溝通。然而要在非外語環境中學習好外國語言通常並不容易，早期要自行學習多需靠單向的錄音帶，廣播與電視教學節目，或是簡單的語音學習機，所幸近十年來，電腦的功能逐漸進步，價格也降低很多，雙向互動式電腦輔助語言學習（computer assisted (or aided) language learning, CALL）[1] 或是電腦輔助發音訓練系統（Computer Assisted Pronunciation Training, CAPT）軟體因此逐漸普及。

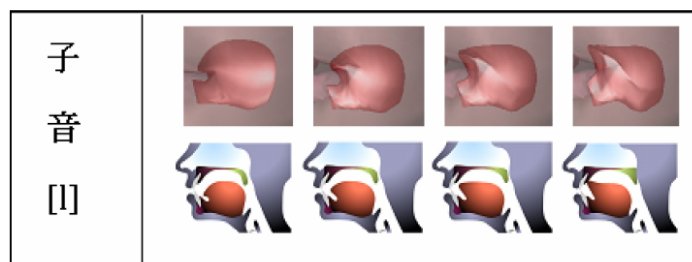
這些 CALL 軟體除了提供使用者一個無壓力的學習環境，與傳統錄音帶，廣播與電視教學節目，或是簡單的語音學習機進步的是可以偵測個別學習者的個人發音、強弱、音調、節奏，給予適當的糾正，例如圖一的 MyET [2] 軟體，可以指出學習者錯誤的發音發生在那一個字，那一個音節，讓學習者可以針對自己有問題的地方，一個字一個字聽到自己的發音與標準發音有何差異，進而加以改善。這類軟體通常也提供圖片或是 3D 動畫等回饋機制 [3]（請見圖二），教導學習者正確發音的技巧，包括口形與舌頭位置。

這些都可以讓使用者針對個人的錯誤反覆的練習，不僅節省了人力、時間，同時也可達到較高的學習效果。例如台灣艾爾科技發行的『MyET』，Power English 的『Tell Me More』軟體等，都已經成熟到足以在正規的語言教學課程中，擔任一定輔助角色（老師通常會要求同學在下次上課前，先自行練習到可以通過軟體的評量標準的程度，以加強學習效果）。

以下先簡介國內外 CALL 發展現況，再進一步介紹 CALL 的核心技術 - 口語評量系統架構，尤其是發音評量，發音錯誤偵測與韻律評量方法，最後是一點簡單討論。



圖一：MyET 電腦輔助英文學習軟體畫面



圖二：成大湯士民、莊則敬、吳宗憲開發之電腦輔助英文學習軟體畫面-舌位動畫與發音口腔圖比較

2. 國內外發展現況

2.1 國外發展現況

許多國內外的學術單位都投入不少心力在 CALL system 的開發上，傳統上以英文學習，語言治療（如構音困難等溝通障礙者或難語症（dyslexia）等治療）為主要研發方向，但目前的國際局勢變成研發中文 CALL 系統最為迫切，另外重要的應用則是自動外國語口語熟練程度測試與用來保存瀕臨危險語言（如台語，客語），也有作成小型的手持裝置的 CALL 系統，以方便攜帶，隨時可以自行學習。

例如在國外方面，包括 Univ. of Cambridge 的 Prof. Steve Young, SRI International (SRI), International Standards for Language Engineering (ISLE), Kyoto Univ (東京大學) 的 Prof. Tatsuya Kawahara (河原 達也), Toyohashi University of Technology (豐橋技術大學) 的 Prof. Seiichi Nakagawa (中川 聖一), University of Erlangen-Nuremberg 的 Andreas Maier, 大陸安徽科大訊飛 (iFLY) 的王仁華教授, The Hong Kong University of Science 的 Prof. Brian Mak 等等，都有相關研究或是實際系統。以下舉例說明：

Univ. of Cambridge 的 S.M. Witt 與 Prof. Young 早在 2000 年即提出以 log-posterior probability 為基礎的 goodness of pronunciation (GOP) [4] 評分，與利用 systematic mispronunciations 觀念之 explicit error modeling 方法，基本上 GOP 其可視為一種 confidence measure, systematic mispronunciations 則是以自己熟習的語言的發音取代不熟習目標語言的發音的現象，這兩種方法實際上一直沿用至今。

ISLE [5] 為一個針對義大利與德國人所設計的英語發音學習系統（見圖三），此系統主要的功能為偵測發音錯誤的位置與發音錯誤的類型。發音錯誤偵測部份主要利用 likelihood 對每個音素來做可信度分析。針對發音錯誤的音素，再利用事先定義好的錯誤規則來偵測錯誤類型。

SRI 的 EduSpeak [6] 主要是供英文學習，其利用 speaker adaptation 技術結合 native 與 non-native 的語音訓練辨認器，使得系統在辨識率方面有較好的效果，並結合 log-posterior score, duration score, 及 rate of speech (ROS) 三個分數來對語音作評分。

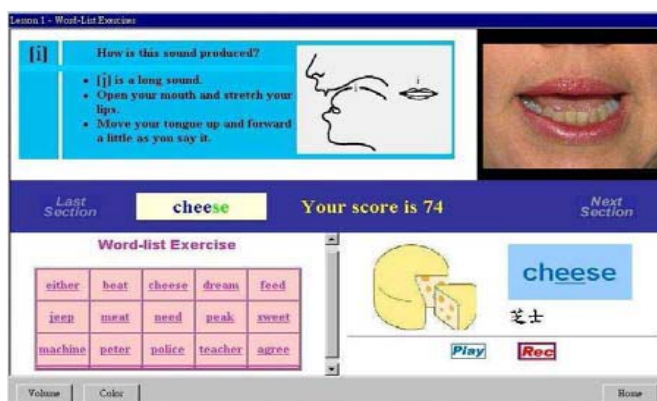
Kyoto Univ. Yasushi 與 Prof. Kawahara 等人則提出以 pronunciation error networks [7-9] 來偵測日本學生常犯發音錯誤，並根據偵測出的錯誤類型，挑選適當的練習語句讓使用者加強練習。

The Hong Kong University of Science 的 PLASER [10]（見圖四）則是針對母語為廣東話的中國人所設計。利用英文與廣東話的語料一起訓練聲學模型。根據評估的結果，75% 的使用者在使用此系統二至三星期後，在英文發音的正確性上均有所提升。

安徽科大 iFLY 主要研究自動化『普通話水平測試』(見圖五), 希望可以取代主觀的普通話水平測試主考官, 其困難點在於如何正確學習人類感知的評斷標準。



圖三：ISLE 電腦輔助英文學習軟體畫面



圖四：Hong Kong University of Science 的 PLASER 電腦輔助英文學習軟體畫面



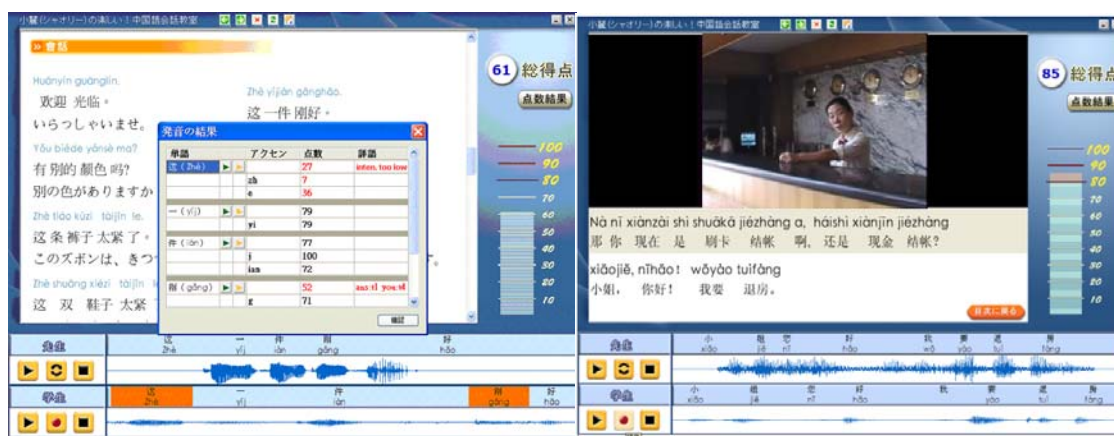
圖五：iFLY 自動化『普通話水平測試』軟體畫面

2.2 國內發展現況

在台灣較知名的 CALL application 有 My English Tutor (My ET) 及總合科技的 Train Speech [11]。MyET 主要是針對發音、強弱、音調、節奏四個部份分別給一個分數。並利用一個側面的舌位動畫與一些發音建議來提供使用者正確發音的回饋介面。Train Speech 主要的核心是利用 IBM Via Voice 的語音辨識器針對發音的部份來評分, 且根據辨識的結果給使用者一些改正發音的建議。

此外工研院資通所也已開發認知仿真語言數位學習技術 [12]，其為 Text Independent 之語音發音評量技術，具有即時動態語音文字同步的特性，不受限於特定的文字內容，可即時回應目前語者語音與文字的對應情形，並針對音素、單字及整篇文章之發音給予評量分數。已技轉給昱泉國際公司，並協助其開發完成兒童美語學習軟體 A+Talk。

在研究方面則有台大李琳山，清華王小川，張智星，成大吳宗憲，長庚呂仁園，臺灣師範大學國語教學中心，高師大黃文禎與周福強（原先在銘傳大學任教）等等老師進行相關研究。其中李琳山老師（與台大語文中心國際華語研習所）與臺灣師範大學國語教學中心皆與艾爾科技合作，發展 MyCT 華語學習課程，王小川與吳宗憲老師亦針對語言治療應用做研究發展，張智星老師在中英文皆有發展（見圖六），呂仁園老師非常有興趣開發台語 CALL 系統，周福強則開發了一套手持英語學習裝置-牙牙學語機（Ya-Ya Language Box） [13-14]（見圖七）。



圖六：清華張智星老師開發之電腦輔助中文學習軟體畫面



圖七：周福強開發之手持式牙牙學語機（Ya-Ya Language Box）

3. 口語評量系統

3.1 口語評量系統架構

通常發音評量都是利用語音辨認技術，自動評量使用者發音，韻律，節奏語音量強弱等語音特性的正確性與否，並給予適當建議。主要的系統架構有兩種：

一是根據老師的語音，直接比對學生的測試語音與老師標準語音間的誤差程度，基本上是要學生發音越接近老師越好，但這種方式會受限於老師錄音語料的多寡。一是利用

較大量語音訓練出的語音辨認器作語音驗證，以機率方式計算學生的語音是否接近語音模型學習到的統計分佈。後者的好處是可以任意產生大量測試文句，不受老師錄音數量限制，在評量方法的研究上也較有發展性，因此現在的系統大多採取第二種作法。

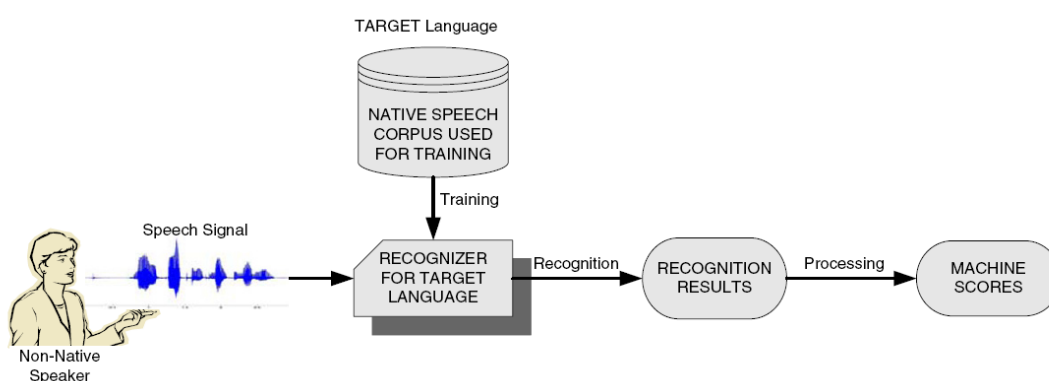
而在第二種系統實作上，通常也有兩種作法，一是建立一套標準學習目標語言的語音辨認器，只使用這個標準語音辨認器作驗證，如圖八所示之架構。另一種則是建立數套語音辨認器，尤其是包含標準學習目標語言與使用者之母語語言之辨認器，如圖九所示之架構。後者的好處是當學習者發音與目標語言差異太大時，仍可以得到較準確之辨認結果作發音評分。

此外因為目的是希望機器的評分結果可以客觀且與人的感知接近，因此此類研究通常需要先有一套由多位語言教學專家（或是 **native speaker**）評分與標記發音錯誤地方的 **database**，當作是標準答案（**ground truth**）。機器評分結果的好壞，則由相關係數公式（如式一）計算機器評分（ x ）與專家評分（ y ）結果相關性是否夠高，一般期待相關係數至少可以達到 0.8 之間。錯誤偵測的好壞則多由 **recall** 與 **precise** 高低，**ROC** 曲線或是機器偵測結果向量（ x_{d1} ）與專家錯誤標記向量（ x_{d2} ）的交相關係數（式二）判定。

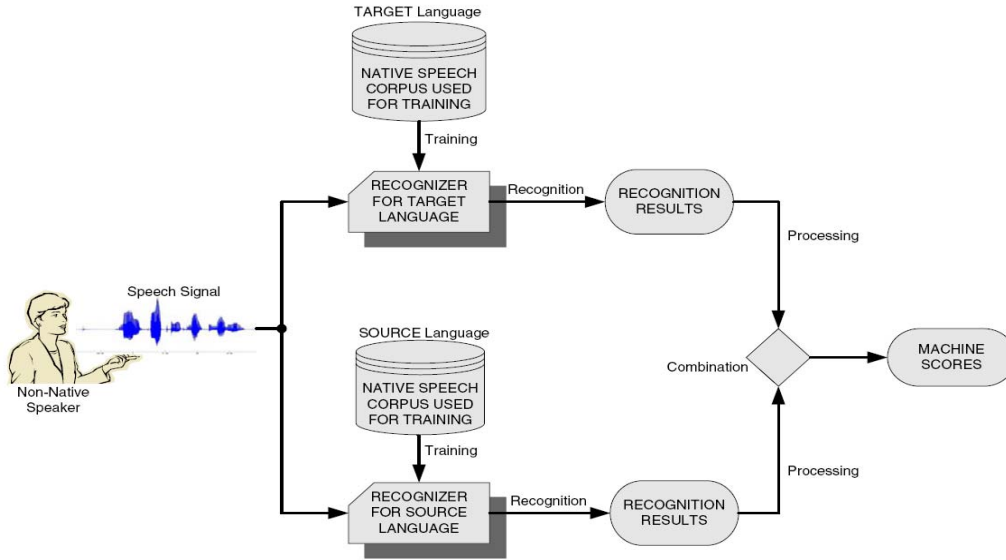
以下先說明在發音評量方面的標準作法，再簡介韻律評分的標準方法（強弱、音調、節奏）。

$$r = \frac{S_{xy}}{S_x S_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

$$CC_{d1,d2} = \frac{x_{d1}^T x_{d2}}{\|x_{d1}\|_E \|x_{d2}\|_E} \quad (2)$$



圖八：使用目標語言之口語評量系統架構



圖九：同時使用目標與母語語言之口語評量系統架構

3.2 發音評量與錯誤類型方法

近幾年來相關的研究主要可分為發音評分與發音錯誤類型偵測兩部份，在發音評分的部份，最通用的評分方法是 S.M. Witt 等人在 2000 年提出的 Goodness of Pronunciation (GOP)，針對句子中的每個音素 (T) 做評分，再將個別 phone 的分數平均而成整句的分數。這種作法是以後置機率 (posterior probability) $P(O|T)$ 為基礎，其作法如圖十所示，基本上是 phone-level 的 duration normalized log-posterior probability (式三)：

$$\begin{aligned}
 P(T|O) &= \sum_{i=1}^N (|\log(P(T_i|O^{(T_i)}))| / NF(T_i)) / N \\
 &= \sum_{i=1}^N (|\log(\frac{P(O^{(T_i)}|T_i)P(T_i)}{\sum_{q \in Q} P(O^{(T_i)}|q)P(q)})| / NF(T_i)) / N \\
 &\approx \sum_{i=1}^N (|\log(\frac{P(O^{(T_i)}|T_i)}{\max_{q \in Q} P(O^{(T_i)}|q)})| / NF(T_i)) / N
 \end{aligned} \tag{3}$$

其中 $O^{(T_i)}$ 是觀察到的語音參數， Q 是所有的 HMM models，即 q 是所有 T_i 可能被讀錯的 phoneme， $NF(T_i)$ 是句子音框 (frame) 的長度， $P(O^{(T_i)}|T_i)$ 則是 HMM 算出來的 likelihood。

此方法也可以再引進語言學上已經歸納出來可能發音錯誤規則，尤其是系統性的發音錯誤 (systemic mispronunciation)，當作先驗知識，例如表一中列出的台灣學生常見的英文發音錯誤，將這些先驗知識用 explicit error modeling 方法 (見圖十一) 表示出來，並取代圖十中的 phoneme loop，以加強其效果，公式如下：

$$\begin{aligned}
 P(T|O) &= \sum_{i=1}^N (|\log(P(T_i|O^{(T_i)}))| / NF(T_i)) / N \\
 &= \sum_{i=1}^N (|\log(\frac{P(O^{(T_i)}|T_i)P(T_i)}{\sum_{q \in Q_{error}^{T_i}} P(O^{(T_i)}|q)P(q)})| / NF(T_i)) / N \\
 &\approx \sum_{i=1}^N (|\log(\frac{P(O^{(T_i)}|T_i)}{\max_{q \in Q_{error}^{T_i}} P(O^{(T_i)}|q)})| / NF(T_i)) / N
 \end{aligned} \tag{4}$$

與上式不同的地方在於現在 $q \in Q_{error}^{T_i}$ ，即 q 不再是所有 T_i 可能被讀錯的 phoneme，而是限制在某些共同錯誤 (common errors) 上，如此可以加強較困難分辨的音素的解析度。

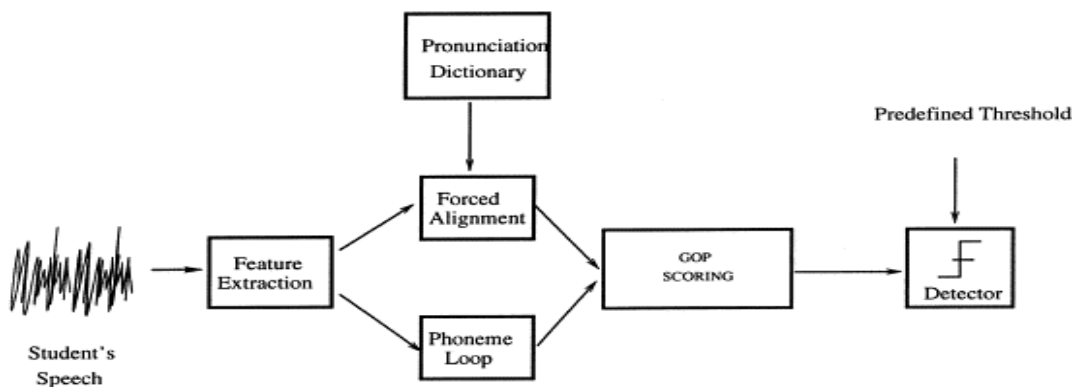
除此方法外，還可以使用 log-likelihood, likelihood ratio, best log-likelihood, ranking-based confidence measure，甚至 phoneme recognition rate, rate of speech (ROS), duration score 與各式 prosodic feature score 等各種評分方式，或是其組合。

例如，SRI EduSpeak System 則結合了 Phoneme posterior score, rate of speech (ROS) 與 duration score 等三種分數來對句子做評分 [6]。Seiichi Nakagawa 等人在 2003 年，則使用 log-likelihood, likelihood ratio, best log-likelihood, a posteriori probability, phoneme recognition rate, rate of speech 等各種評分方式來對句子評分，最後實驗發現結合 log-likelihood, best log-likelihood, phoneme recognition rate, rate of speech 與專家所評斷的結果有較高的相關性 [15]。

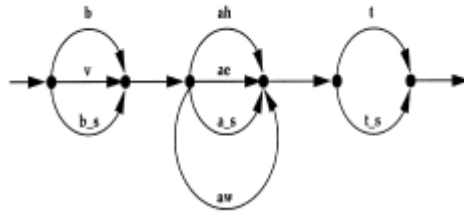
而在發音錯誤偵測方面，通常則將評分結果再經過 threshold 判斷，進一步將分數太低的部分判定為發音錯誤，作法亦如圖十所示，但為了提高偵測正確率，目前作法通常會使用 phone-dependent threshold，也一定會引進語言學上已經歸納出來可能發音錯誤規則，當作先驗知識。例如 Yasushi Tsubota 等人在 2002 年利用 pronunciation error network 來偵測日本學生發音錯誤的類型，並利用 LDA 針對發音錯誤的部份作驗證 [7-9]。2004 年 Jong-mi Kim 等人則是根據韓國人的發音習慣建出一些可能的發音錯誤規則，辨識的時候利用這些發音規則來找出使用者發音與正確發音上的差異 [16]。

Vowel Error Type	Example	Consonant Error Type	Example
/A/ → /C/	not → nought	/v/ → /f/	van → fan
/u/ → /U/	fool → full	/z/ → /s/	rise → rice
/i/ → /O/	seat → sit	/b/ → /p/	bill → pill
/ʃ/ → /A/	cat → cart	/d/ → /t/	made → mate

表一：台灣學生常見英文發音錯誤表（部分，取自清華莊雅琪論文[17]）



圖十： Goodness of pronunciation 評分系統架構

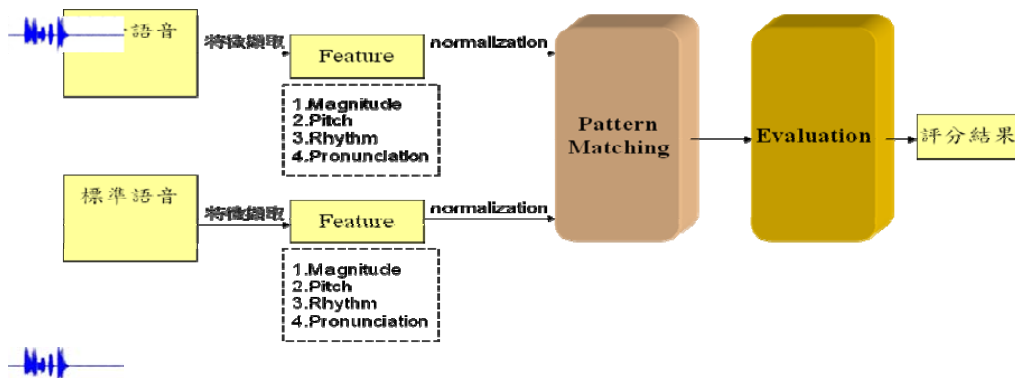


圖十一： explicit pronunciation error modeling network

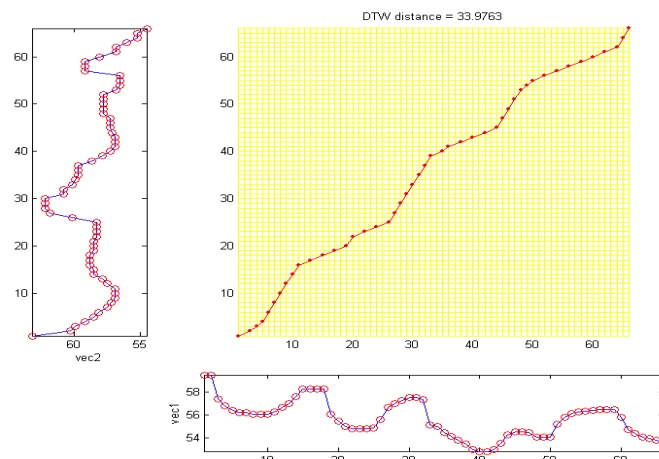
3.3 韻律評量

在這裡我們將強弱、音調、節奏三部份統稱為韻律，相對於發音評量，韻律評量的研究較少，且大部分假設系統中存有老師的標準語音，因此作法上大都根據老師的標準語音，直接比對學生的測試語音與老師標準語音間的 pitch 與 energy contours 與 phone/word duration 長短的誤差程度，系統架構如圖十二所示。這種作法基本上使用 dynamic time warping (DTW)，圖十三為比對 pitch contour 時的示意圖，因此可視為一種 template pattern matching 的作法。

但是假設系統中存有老師的標準語音，有時並不可行，因此在韻律評量方法方面，應該要能如同發音評量使用 HMM 模型般，發展出一些韻律模型來作韻律評量，如此一方面可以避免需要原始老師語音的問題，一方面可能可以如同發音評量使用 explicit error modeling 的方法，以加強評量的效果。



圖十二： 韻律評量系統架構圖（取自清華李俊毅論文[19]）



圖十三：使用 dynamic time warping (DTW) 比對 pitch contour 之示意圖（取自清華李俊毅論文[18]）

4. 討論

在這篇簡介中，我們只 review 了口語評量的技術部分，因此缺少了幾個重要的部分，一是學習理論與教材設計的探討，一是忽略了如何收集語料的設計，尤其是評分標準的設定。原因是教材設計，評分方法與收集語料都需要語言學家的幫忙，尤其是收集實際從事語言教學的老師與專家的多年教學經驗，這部分需要進一步與語言教學專家討論。

此外，我們也忽略了在國際上實際在做第二外國語能力測試時，其中的口語(speaking)測試(oral proficiency interview, OPI) [19]部分，一般是採取與主考官面談的方式(interview)，因此著重在溝通的能力，事先並不可能知道被測試者要如何回答，也就是說我們事先並沒有文字資訊。因此如何模擬此實際情形，真正幫助使用者發展溝通能力，而不是如目前的作法要求使用者照本宣科，是 CALL 系統的一大挑戰。

5. References

- [1] Computer-assisted language learning, http://en.wikipedia.org/wiki/Computer-assisted_language_learning, 2008/7/3
- [2] MyET, <http://www.myet.com/MyETWeb/PersonalizedPage.aspx>, 2008/7/3
- [3] 湯士民、莊則敬、吳宗憲，應用錯誤型態分析於英語發音輔助學習，ROCLING'2005
- [4] Witt and S. Young (2000). "Phone-level Pronunciation Scoring and Assessment for Interactive Language Learning." *Speech Communication* 30(2/3): 95-108.
- [5] ISLE, Interactive Spoken Language Education, University of Hamburg, <http://nats-www.informatik.uni-hamburg.de/~isle/>
- [6] H. Franco et al., "The SRI EduSpeak System: Recognition and Pronunciation Scoring for Language Learning", *Proc. InSTIL, Scotland*, pp. 123-128, 2000
- [7] Y.Tsubota, T.Kawahara, and M.Dantsuji., Recognition and verification of English by Japanese students for computer-assisted language learning system., In *Proc. ICSLP*, pp.1205--1208, 2002. (PDF file)
- [8] K.Imoto, Y.Tsubota, A.Raux, T.Kawahara, and M.Dantsuji., Modeling and automatic detection of English sentence stress for computer-assisted English prosody learning system., In *Proc. ICSLP*, pp.749--752, 2002. (PDF file)
- [9] Raux and T.Kawahara., Automatic intelligibility assessment and diagnosis of critical pronunciation errors for computer-assisted pronunciation learning., In *Proc. ICSLP*, pp.737--740, 2002. (PDF file)
- [10] Brian Mak, M. H. Siu, M. Ng, Y. C. Tam, Y. C. Chan, K. W. Chan, K. Y. Leung, S. Ho, F. H. Chong, J. Wong, J. Lo, PLASER: Pronunciation Learning via Automatic Speech Recognition, *Proceedings of HLT-NAACL, 2003*, (draft) pdf
- [11] 總合科技, Train Speech, <http://www.allzone.com.tw/products.htm>, 2008/7/3
- [12] 前瞻技術中心 - 認知仿真語言數位學習技術, <http://www.atc.itri.org.tw/content/menu-sql.asp?pid=59>, 2008/7/3
- [13] Fu-chiang Chou, "On-line English Pronunciation Training System with Automatic Speech Recognition Technologies", *Proceeding of ED-MEDIA*, pp. 3819-3824, 2005
- [14] Fu-chiang Chou, "Ya-Ya Language Box - A Portable Device for English Pronunciation Training with Speech Recognition Technologies", *Interspeech*, pp. 169-172, 2005
- [15] S. Nakagawa, K. Mori, N. Nakamura, "A statistical method of evaluating pronunciation proficiency for English words spoken by Japanese", *Proc. EuroSpeech-2003*, pp.3193-3196 (2003.09)
- [16] Jong-mi Kim, Suzanne Flynn, What Makes a Non-native Accent?: A Study of Korean English, *ICSLP'2004*
- [17] 莊雅琪，以母語輔助建立個人化辨識網路，2007年清華碩士論文
- [18] 李俊毅-語音評分，2002年清華碩士論文
- [19] The ACTFL Oral Proficiency Interview (OPI), http://www.languageTesting.com/acad_opi.htm

6. Appendix

6.1 Professional associations

- APACALL: Asia-Pacific Association for CALL.
- CALICO: US-based professional association devoted to CALL. Manages a regular annual conference.
- JALTCALL: Japan-based professional association devoted to CALL. Coordinates an annual conference and the JALTCALL Journal.
- PacCALL: Professional CALL association in the Pacific: from East to Southeast Asia, Oceania, across to the Americas.
- PacCALL Australia: Australian Chapter of the Pacific CALL Association. President Renata Chylinski.
- EUROCALL: Europe-based professional association devoted to CALL. Manages a regular annual conference.
- IALLT: US-based International Association for Language Learning Technology. IALLT is a professional organisation dedicated to promoting effective uses of media centres for language teaching, learning, and research. Manages regular conferences.
- Learning Technologies Special Interest Group The Learning Technologies Special Interest Group of the International Association of Teachers of English as a Foreign Language. This UK-based group runs a variety of events and produces a regular newsletter.
- TESOL Teachers of English to Speakers of Other Languages, CALL Interest Section.
- WorldCALL: A worldwide association devoted to CALL and embracing other leading professional associations. The WorldCALL 2008 conference will take place in Japan.
- SLaTE Speech and Language Technologies for Education

6.2 Professional journals dedicated to CALL

- *CALICO Journal* (Computer Assisted Language Instruction Consortium)
- *Teaching English with Technology* (IATEFL Poland)
- *CALL-EJ On-line* (Online Journal)
- *Computer Assisted Language Learning: An International Journal* (Taylor and Francis)
- *CALL Review: the SIG Journal* (The IATEFL Special Interest Group's Newsletter)
- *IALLT Journal* (International Association for Language Learning Technology)
- *JALTCALL Journal* (Japan Association of Language Teaching - Computer-Assisted Language Learning Special Interest Group)
- ON-CALL (Australia) Archives only - now incorporated into CALL-EJ:
<http://www.cltr.uq.edu.au/oncall/home.html>
- *Language Learning and Technology* (Online Journal)
- *ReCALL* (European Association for Computer Assisted Language Learning)
- *Journal of Computer Assisted Learning* (Blackwell - Computer Assisted Learning in general rather than CALL)
- SLaTE
- *AsiaCALL Online Journal* (AsiaCALL OnlinJournal)