中華民國計算語言學學會
通訊 ACLCLP
The Association for Computational Linguistics and Chinese Language Processing
發行人：陳信宏
主　編：曾淑娟
執行編輯：黃　琪
第十八卷第一期
E-Mail：aclclp@hp.iis.sinica.edu.tw
Website：http://www.aclclp.org.tw
傳 真：(02)2788-1638
地 址：台北市研究院路二段128號中研院資訊所
電 話：(02)2788-3799 ext:1502
劃 撥：19166251

本 期 要 目

ROCLING-2007

「第十九屆自然語言與語音處理研討會」謹訂於九月六日～七日假國立台灣大學舉行，本屆是由國立台灣大學資工所與本會共同主辦。歡迎踴躍投稿及參加，徵稿啓事請參閱本刊第二頁。

獎助學生出席國際會議

- 會議名稱：ICASSP -2007
- 論文題目：Word Topical Mixture Models for Dynamic Language Model Adaptation
- 獎助學生：邱炫盛(台灣師範大學資工所)
- 獎助金額：NT$ 30,000 元

NTCIR 跨語言中文問答系統比賽

中央研究院資訊科學研究所許聞廉老師所主持的「智慧型代理人實驗室」(IASL)，參加 NTCIR-6 跨語言問答系統比賽(CLQA)，獲得「中中」以及「英中」兩個子項冠軍。

NTCIR 爲日本 NII 所主辦的國際會議，與會者多爲亞洲資訊處理相關學者，會議目的在促進相關領域技術交流。第一屆會議開始於 1999 年，每年都舉辦多項資訊處理比賽。今年 NTCIR-6 的比賽項目包括「跨語言文件檢索」、「跨語言問答系統」、「專利文件檢索」、「日文問答系統」、「意見分析」、以及「多模趨勢資訊摘要」等六項。其中「跨語言問答系統」又細分爲中、英、日三種語言混雜的數個子項目。在「中中」子項中，參賽系統須在指定的新聞語料庫(CIRB2.0)中，找出中文題目的中文答案; 而在「英中」子項中，則必須找出英文題目的中文答案; 困難度大幅增加。智慧型代理人實驗室累積多年問答系統開發經驗，於去年獲得「中中」子項冠軍，今年該實驗室再度於「中中」子項以 55.3%答題正確率奪冠（去年爲 44.5%），同時也以 34%之正確率（去年之第一名爲 12%）獲得了「英中」子項冠軍。

在本次「智慧型代理人實驗室」的比賽系統中，爲了提升回傳文章篇數，拉高候選答案召回率(Recall)，特別改進了「答案過濾」以及「答案排序」模組，以便應付召回率提升所引入的其他不相關答案。在「答案過濾」上，除了依據題目類型來過濾答案外，還另外使用了一百多個自動生成的「答案模版」來保留高度相關答案。藉由將多序列並列分析技術(Multiple Sequence Alignment)應用到問答訓練資料文句上，有效地找出適用性廣且準確性高的答案模版。在「答案排序」方面，則是增加了一個新的特徵(feature)：計算候選答案與題目各關鍵字在回傳文句中共同出現程度(co-occurrence)。另外還使用了基因演算法來調整各特徵的權重，以解決人工調整權重的不便以及不確定性。實驗證明，這些新導入的技術除了對「中中」問答有幫助外，對於「英中」問答也有一定的效果。

「智慧型代理人實驗室」表示，到目前爲止所使用的技術都還是屬於「淺層」的語言分析技術，這些技術的優點在於實作容易，成本低，通常不需額外的知識或工具，但其成效有先天上的限制。目前的正確率已經達到「淺層」技術之上限，若要有進一步的突破，則必須要轉向應用「深層」的語言分析技術，在語意框架與文句結構上下功夫。2006/11/24

Conference on Computational Linguistics and Speech Processing
第十九屆自然語言與語音處理研討會

September 6-7, 2007, National Taiwan University, Taipei, Taiwan, ROC
http://nlg12.csie.ntu.edu.tw/~Rocling/index.html

CALL FOR PAPERS

Conference Chair

Hsin-Hsi Chen
National Taiwan University

Program Committee

Kuang-Hua Chen, Co-Chair
National Taiwan University

Berlin Chen, Co-Chair
National Taiwan Normal University

Jing-Shin Chang
National Chi Nan University

Jason S. Chang
National Tsing Hua University

Chia-Ping Chen
National Sun Yat-Sen University

Keh-Jiann Chen
Academia Sinica

Sin-Horng Chen
National Chiao Tung University

Pu-Jen Cheng
National Taiwan University

Jen-Tzung Chien
National Cheng Kung University

Zhao-Ming Gao
National Taiwan University

Wen-Lian Hsu
Academia Sinica

Chih-Chung Kuo
CCL/ITRI

Tyne Liang
National Chiao Tung University

Shou-De Lin
National Taiwan University

Ren-Yuan Lyu
Chang Gung University

Yuen-Hsien Tseng
National Taiwan Normal University

Hsiao-Chuan Wang
National Tsing Hua University

Hsin-Min Wang
Academia Sinica

Chung-Hsien Wu
National Cheng Kung University

Ming-Shing Yu
National Chung Hsing University

Organization Chair

Lun-Wei Ku
National Taiwan University

The 19th ROCLING Conference will be held at National Taiwan University, Taipei, on September 6-7, 2007. 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 XIX will be hosted by the Department of Computer Science and Information Engineering, Department of Library and Information Science, National Taiwan University. The two-day conference will feature invited talks, paper and poster sessions.

ROCLING XIX invites the submission 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.

(a) cognitive/psychological linguistics
(b) discourse/dialogue modeling
(c) information extraction/text mining
(d) information retrieval
(e) language understanding/generation
(f) lexicon/morphology
(g) machine translation/multilingual processing
(h) named entity recognition
(i) NLP applications/tools/resources
(j) phonetics/phonology
(k) question answering
(l) semantic web
(m) semantics/pragmatics
(n) speech analysis/synthesis
(o) speech recognition/understanding
(p) spoken dialog systems
(q) spoken language processing
(r) syntax/parsing
(s) text summarization
(t) web knowledge discovery
(u) word segmentation/POS tagging
(v) others

Paper Submission

Prospective authors are invited to submit full papers of no more than 25 A4-sized pages in PDF or Microsoft Word format (visit our website to download Word Template). Papers will be accepted only by electronic submission through the conference website. Prospective authors without Web access should contact the Program Committee Co-Chair (khchen@ntu.edu.tw) before the submission deadline. 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, with a short abstract of the paper following.

Best Paper Award

The best paper will be selected and announced at ROCLING XIX.

Important Dates

Preliminary paper submission due:	July 2, 2007
Notification of acceptance:	August 10, 2007
Final paper due:	August 20, 2007

Sponsors

Association for Computational Linguistics and Chinese Language Processing (ACLCLP)
Department of Computer Science and Information Engineering, National Taiwan University
Department of Library and Information Science, National Taiwan University

Parallel Session of Designated Topic 9: Computational Linguistics

CIL18 (18th International Congress of Linguists)
http://www.cil18.org
July 21-26, 2008, Korea University, Seoul, Republic of Korea

Important Dates:

- May 31, 2007: Deadline for submitting the abstract.
- August 31, 2007: Notification of acceptance.
- Submission of accepted abstract for publication in the proceedings: February 15, 2008:
- Submission of final paper to be published in CIL18 CD: September 30, 2008

Description:

Computational linguistics is motivated and scientifically rooted in three scientific fields studying language, technology and cognition. It drives its results from novel connections between these fields, such as the use of specialized information technology for qualitative and quantitative linguistic research or the exploitation of linguistic theory and insights for the development of novel software applications. Other important links involve the combination of linguistic theory and formal models of cognition as well as the embedding of language competence into multi-modal computer interfaces. Overarching formal and experimental themes are the computability of language and models of language communication, language acquisition and language change.

In line with the theme of the Congress, the emphasis of the session will be on computational models and applications of linguistic analyses and theories. Such implementations may serve the better understanding and gradual improvement of the underlying linguistic theories. They may also serve some concrete functional needs in advanced software applications.

But papers are also invited that present research on other novel connections between linguistic, cognitive and engineering approaches with the aim to further our understanding of language or to provide useful and user-friendly applications for our daily life.

The range of computational methods for processing language has broadened significantly. Non-discrete mathematical approaches involving statistical methods, optimization techniques and neural networks play an increasingly dominant role in computational linguistics. For this session, we explicitly extend the focus beyond purely linguistic methods to include so-called hybrid approaches to language processing such as combinations of statistical techniques and linguistic models as well as informative comparisons between competing paradigms.

Form and submission of abstracts:

An abstract (.pdf or .doc file) should be up to 3 pages long, including data and references. The abstract should start with the title of the paper, followed by the text of the abstract.
Please do not include the author's name in the abstract. On a separate page, please give the author's name, affiliation, e-mail address, telephone number, mailing address, the paper title and the session number(title).

Workshop on
Linguistic Studies of Ontology:
From Lexical Semantics to Formal Ontologies and Back

CIL18 (18th International Congress of Linguists) http://cil18.org
July 21-26, 2008, Korea University, Seoul, Republic of Korea

Organizer: Chu-Ren Huang
E-mail address: churenhuang@gmail.com Fax: 886-2-27856622

Description:

Recent developments in the study of ontology have important implications for cognitive science, knowledge engineering, and theoretical linguistics. In particular, research on lexical ontology deals with how concepts are lexicalized and organized across languages and cultures. This workshop aims to explore this new departure in linguistic studies by building upon the three important premises assumed in Fellbaum (1998), Schalley and Zaefferer (2007), and Huang et al. (2007): First, that lexicalized concepts have a special status in every language (as opposed to concepts that require complex coding), second that lexically coded concepts can be shared by different languages, and third that lexicalization universals are relevant for the construction of cross-lingually portable formal ontologies.

Following the references cited above, topics of this workshop include foundational issues pertaining to the relation between formal ontology and linguistic ontologies, as well as descriptive issues pertaining to the interface between conceptual ontologies and lexica. In particular, we would like to focus on the following issues during this workshop:

- Cross-lingual portability of upper-ontologies
- Ontology-based approaches to comparative linguistics
- Ontology enrichment: from concept formation via complex coding to lexicalisation
- Possible relevance of formal ontological principles (e. g Roles cannot subsume Types) to psychological/linguistic reality

References:

Fellbaum, Christiane. 1998. WordNet: An electronic lexical database. MIT Press.
Huang, Chu-Ren et al. Eds. To Appear (2007) Ontologies and the Lexicon. Cambridge University Press.
Schalley, Andrea C. and Zaefferer Dietmer. Eds, 2007. Ontolinguistics. Mouton-De Gruyter.

Submission of Abstracts:

A two-page abstract including everything should be sent electronically to both cil18@cil18.org and churenhuang@gmail.com. An **MS Word** and/or **.pdf** file may be accepted.

Important Dates:

- Deadline for Abstract Submission: **May 31, 2007**
- Notification of Acceptance/Rejection: August 31, 2007
- Submission of accepted abstract for publication in the proceedings: February 15, 2008:
- Submission of final paper to be published in CIL18 CD: September 30, 2008

Please send the abstract and the author's information to both cil18@cil18.org and uszkoreit@dfki.de.

Call for Papers
International Journal of
Computational Linguistics and Chinese Language Processing

Special Issue on

Cross-Lingual Information Retrieval and Question Answering

Cross-lingual information retrieval (CLIR) and cross-lingual question answering (CLQA) have recently gained more attentions as different ways of cross-lingual information access for different needs. CLIR and CLQA are much more complicated challenges compared with monolingual IR and monolingual QA. Many practical problems are to be solved to make their techniques more effective. For examples, they need to deal with query/question translation, question classification, document retrieval, and answering. In recent years, a number of techniques have been made to improve the state of the art. The goal of this special issue is to solicit papers showing recent advances and applications in CLIR and CLQA. Prospective authors are invited to submit their works to this special issue. Subjects of interest include, but are not limited to:

- Statistical models or language models for CLIR and CLQA
- Query/question translation
- Unknown word translation or name entity translation
- Question analysis/classification
- Passage/answer extraction
- Cross-lingual retrieval
- Cross-lingual categorization and classification
- Cross-lingual text mining
- Cross-lingual semantic analysis
- Machine translation/transliteration
- Resource for CLIR and CLQA

Important Dates

Submission deadline: June 30, 2007
Notification of acceptance: September 30, 2007
Camera-ready due: December 15, 2007
Tentative publication date: March 2008

Instructions for Authors

All manuscripts are subject to anonymous peer review. The style for manuscript is available at the homepage of International Journal of Computational Linguistics and Chinese Language Processing (http://www.aclclp.org.tw/journal/index.php). The authors should submit their papers in PDF to one of the guest editors.

Guest Editors

Wen-Hsiang Lu (whlu@mail.ncku.edu.tw)
National Cheng-Kung University, Taiwan
Jianfeng Gao (jfgao@microsoft.com)
Microsoft Research, USA
Min-Yen Kan (kanmy@comp.nus.edu.sg)
School of Computing, National University of Singapore, Singapore
Cheng-Zen Yang (czyang@saturn.yzu.edu.tw)
Dept. of Computer Science & Engineering, Yuan Ze University, Taiwan

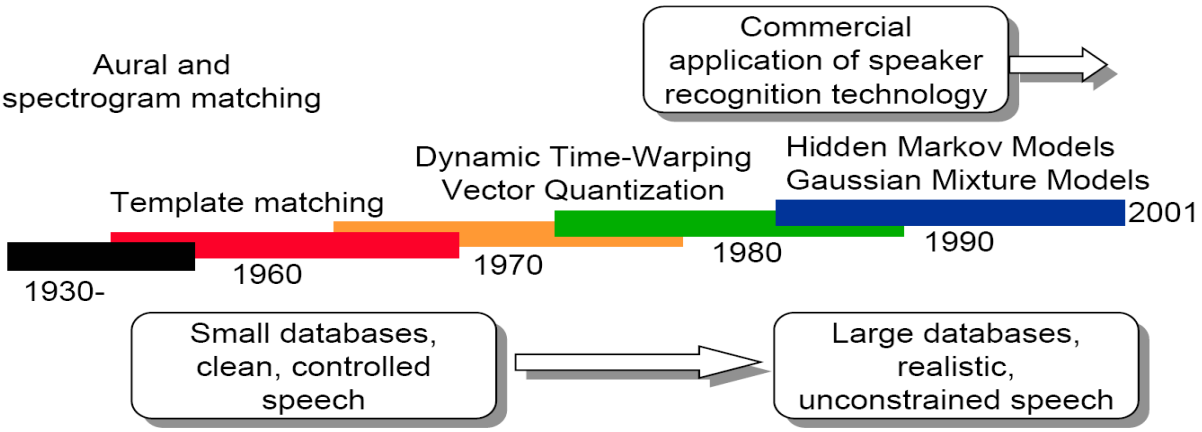
強健性語者驗證技術簡介

張文杰、廖元甫、莊堯棠
國立台北科技大學電子工程研究所
國立中央大學電機工程研究所
wcchang@ee.ncu.edu.tw　yfliao@ntut.edu.tw　ytjuang@ee.ncu.edu.tw

1. 前言

透過言語傳遞是人類最常用來溝通的方式之一，而語音信號因包含不同層面的豐富資訊以致能讓聽話人足以瞭解所表達之意。隨著數位科技的蓬勃發展以及網際網路和無線通訊的創新普及，使得語音辨識(Speech Recognition)的應用愈來愈廣泛，其中在語者辨識(Speaker Recognition)的應用更是目前生物辨識技術上最具研究潛力的發展方向。

回顧過去幾十年來的研究發展我們可知，語音辨識技術確實已有顯著地進步，從圖一中[1] 看到早期的樣版比對問題在中期推演出 Dynamic Time Warping (DTW)技術，讓獨立詞(Isolated Word)或多個連接詞彙(Connected Words)的辨認問題得以容易地解決，到後期更爲了希望能解決在連續語音辨識任務的問題，CMU 及 IBM 這兩個知名的研究單位都提出了隱藏式馬可夫(Hidden Markov Model, HMM) [2] 方法來解決。而有趣的不只是語音辨識技術到目前已有穩定且不錯的應用效果，當語音辨識技術套用在語者辨識的任務時，我們才發現一樣能有效地將胡亂無章的聲音信號轉換成帶有身份資訊的語者特徵，所以語者辨識的應用可說是有著相同速度的發展。另外我們從圖一亦可發現早期在語者辨識的研究都侷限在少量且安靜環境所錄製的語料庫，通常也都是在朗讀(Reading)及單一時間點(Single Session)的條件下完成，漸漸地隨著語者辨識技術的成熟，開始演變到大量且較近乎現實情況的語料，使得辨識目標不僅要能不限制說話內容(Text-Independent)，還必須考量通道(Channel)變化及說話人身體狀態的差異性。因此以這樣的發展趨勢來看，加速發展語者辨識技術不只是語音信號處理中最重要的分支研究，在未來語者辨識更勢必成爲資訊安全中新一代可靠的認證機制。

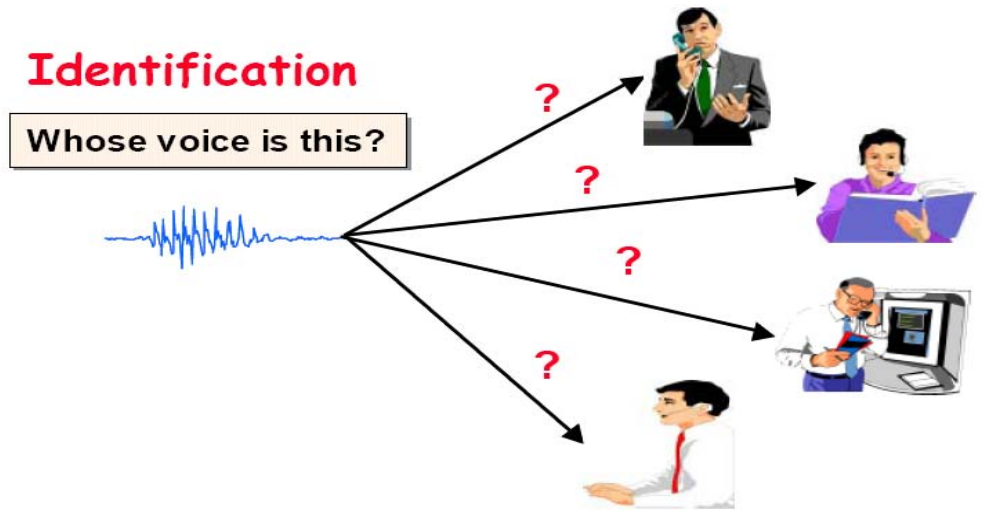


圖一、語者辨識演進圖

在本文章中，我們會先就語者辨識作一簡單介紹，並說明目前語者辨識領域中最主要的架構以及國際間的研究概況。接者我們將描述語者辨識技術於實際應用上所面臨的難題，並介紹當前強健性語者辨識的相關技術。

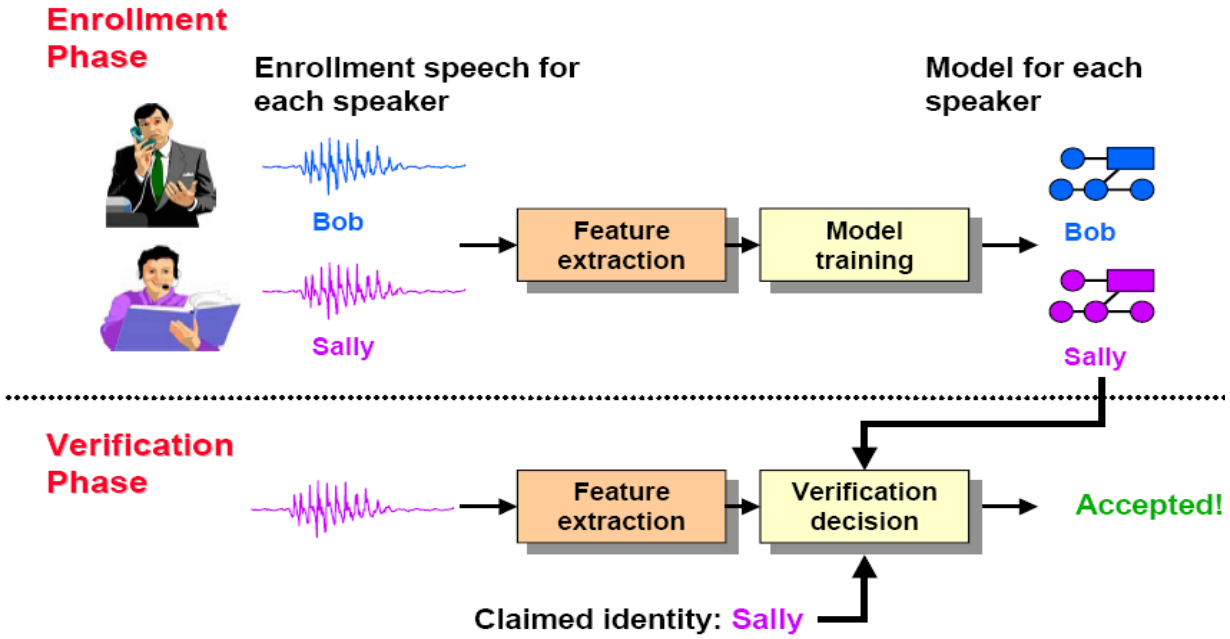
2. 語者辨識

一般來說，語者辨識依據使用目的不同可分爲兩大類 [3]：語者識別(Speaker Identification)和語者驗證(Speaker Verification)。語者識別必須事先針對一組已知身份的使用者收集各自的語音資料，並擷取出資料中的語者特徵參數或進一步建立語者模型。當未知身份的測試語音輸入系統時，系統會經由某種聲音相似度計算或是模型比對來決定出說話者的身份，而由於說話者並未宣稱其身份，所以語者識別可謂之爲一對多(one to many mapping)的分類處理，另外更廣泛地來說，我們都會假設這個未知身份的測試語音必須來自於已知身份的成員之一，因此語者識別通常也可稱之爲封閉集合識別(Closed-set Identification)，如圖二 [1] 所示。



圖二、語者識別示意圖

在語者驗證方面，如圖三 [4] 所示，我們亦可稱爲語者認證(Speaker Authentication)或是語者偵測(Speaker Detection)，其主要是用來鑑定說話者宣稱之身份的真實性，且不同於語者識別之處在於系統會先獲知使用者表明的身份，所以這會成爲一對一(one to one mapping)的決策問題，也因此判斷錯誤的情況會包含兩種，其一是錯誤接受(False Accept)，即將冒充者(Imposter)誤認爲其所宣稱的身份，另外是錯誤拒絕(False Reject)，即誤將真實語者視爲冒充者。此外我們通常會假設冒充者是系統中未知的身份，其語音資料幾乎都無法事先取得，因而我們又可稱之爲開放集合驗證(Open-set Verification)。



圖三、語者驗證示意圖

3. 相關研究概況

語者辨識在美國 NIST (National Institute of Standards and Technology)是自從 1996 年起就已經開始舉辦語者辨識評比(Speaker Recognition Evaluation ,SRE) [5]，每年所有參賽單位均使用 NIST 所提供的標準對話電話語料(Conversational Telephone Speech)，其評比項目共包含單一語者偵測(1-Speaker Detection)、雙語者偵測(2-Speaker Detection)、語者分段(Speaker Segmentation)以及語者追蹤(Speaker Tracking)等。其中單一語者偵測是在判斷這段語音是否爲假設語者(Hypothesized Speaker)的說話的聲音，其實就相當於語者驗證；雙語者偵測則是必須在一段兩人對話的語音中決定出假設語者是否包含在其中；而語者分段是須先這段語音中各語者的聲音區段找出，進而將這些區段依據語者身份做分群；另外在語者追蹤項目中則是將一段語音中屬於某一個假設語者的區段一一的標示出來。目前在相關期刊包括 IEEE Transaction on SAP、Digital Signal Processing、Speech Communication、JASA、Computer Speech and Language 等與研討會 ICASSP、Interspeech、ICSLP 等論文中大多已採用 NIST 語者辨識評比的語料來進行實驗，因此 NIST SRE 可說是國際公認的語者辨識評比基準(Benchmark)。

除此之外，由 Linguistic Data Consortium (LDC) [6] 所發行的 YOHO、KING、TIMIT 等亦是常被使用的語料庫，而歐洲語系中由 European Language Resource Association (ELRA) [7]所發行的 SIVA、POLYCOST、PolyVar 等語料庫也是有人作爲語者辨識的研究資源。另外值得一提的是關於漢語的語者辨識評比也在 2006 年首度於中文口語語言處理國際會議舉辦 [8]，評比項目主要是語者識別和語者驗證兩大主題，針對通道不匹配的問題亦有做相當仔細之條件設計。雖然說漢語的評比在歷史發展上遠不及 NIST SRE，但也算是爲漢

語研究開啓一道技術交流的大門，畢竟漢語是聲調(Tonal)語言中最具代表的語言，未來發展絕不容忽視。

4. 主流方法概述

語者辨識可分爲訓練與測試兩階段的操作模式。在訓練階段主要是擷取使用者之註冊(Enrollment)資料的語者特性，一般會再將此語者特性表示爲語者模型，而測試階段則是估測使用者所輸入的語音和系統中語者模型之間的相似性。此外，根據運用層面的不同又可分爲特定文本(Text-Dependent)與非特定文本(Text-Independent)兩大類。特定文本中，使用者說話的內容會受限在某些關鍵詞語或文句，且測試時使用者說出的語音內容必須和訓練語料相同，由於系統已知測試語音內容，所以通常會事先建立各語者所屬的音素(Phoneme)模型，以便控制和語者特性無關的聲音變異，目前這類語者辨識運用最常使用的方法是隱藏式馬可夫模型。至於非特定文本則不會語音內容有任何限制，也因此在任務上會比特定文本較困難些，這類的運用普遍採用高斯混合模型(Gaussian Mixture Model, GMM) [9] 來描述語者在頻譜上特性分佈。

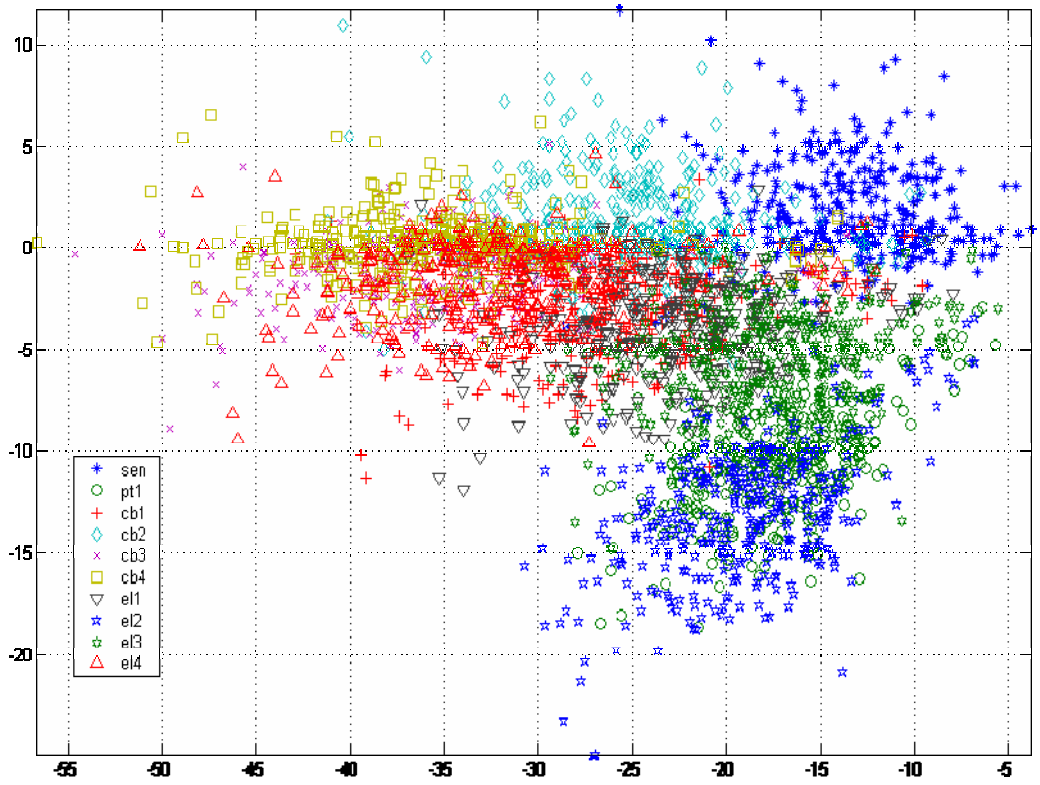
再者考量到一般基於統計模型的語者辨識技術所使用的聲學特徵參數同樣來自於語音辨識，可能面臨因爲訓練語料不夠充足而造成辨識效能嚴重衰退的問題，所以目前最主流的解決方式乃採用 Reynolds 等人 [10] 所提出的 GMM-UBM 方法，這是一種基於模型調適的架構，利用具有大量語者資訊的資料庫來訓練一通用背景模型(Universal Background Model, UBM)，此模型可代表一般非特定語者的聲音特性。接者使用者根據各自少量的訓練語料，藉由最大事後機率估算法(Maximum A Posteriori Estimation)將通用背景模型調適成個別語者的特定模型(Speaker Dependent Model)。

5. 不匹配問題

上一章節所提及之主流方法，在語者辨識的應用確實能達到相當不錯效果，尤其是在安靜的測試環境下 [11] ，然而這樣的表現是無法滿足現實環境條件的要求，因爲語者辨識常在環境不匹配的情況進行，譬如透過電話連接到銀行作身份驗證就是一個現實例子，經由不同電話來源的對話就人類聽覺來說並非問題，但對於自動化的語者辨識系統卻是影響甚大，而此問題將嚴重導致辨識結果大幅滑落，所以這也是無法將語者辨識應用商業化的最大絆腳石。

在不匹配的問題中，主要包含了電話話筒(Handset)、通道(Channel)及說話狀態(Session)等多種影響條件，使得訓練時經由錄音取得的語音特性和測試時有著極大不一致性，我們透過知名 HTIMIT [12] 語料庫的分析會更加明白，圖四是將同一句話經過不同電話話筒所得到的頻譜特徵參數顯示在前兩個維度的分佈情況，在圖中能明顯看出不同的電話話筒造成特徵分佈有相當大的偏移，且在分佈範圍上也有相當大的差異，這樣的結果不僅說明話

筒特性的變異足以改變原有語音的聲學特性，更可知道未知話筒(Unseen Handset)在辨識時絕對是個傷害。因此在接下來幾個章節當中，我們將介紹一些目前常用來減緩不匹配影響的方法，而在此僅針對語者驗證技術方面加以說明，因爲這些方法亦可補償在語者識別技術上。



圖四、頻譜特徵參數於不同通道之分佈情形

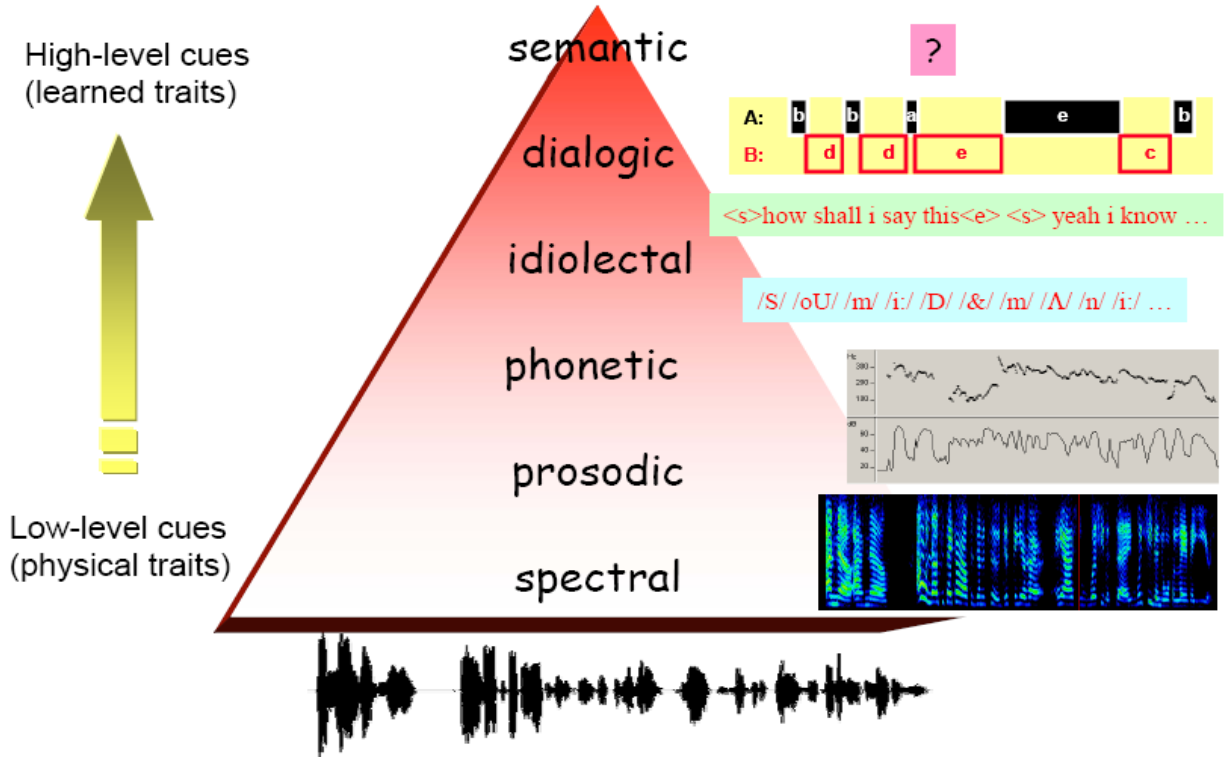
6. 補償方法簡介

以不同層級區分，補償方法大致上可分爲三大類，分別爲強健性特徵參數(Robust Feature)、特徵參數與模型補償(Model Compensation)及特徵參數與分數正規化(Score Normalization)。底下我們就針對這些補償的方法做介紹。

6.1 強健性特徵參數

至今在語者辨識最有效且常用的特徵參數仍以頻譜(Spectrum)爲基礎，因爲頻譜足以反映出大多數的語者訊息，但是當頻譜特徵遇上不匹配問題時就無法繼續維持原有的描述能力[13-15]，所幸頻譜特徵只有表達出最低階層的線索(Cues)，還有更多較高階層的線索尚未被妥善開發 [16]。圖五總和顯示目前語者辨識中適合使用的各種資訊 [1, 17]，最下層即是最具物理特性(Physical Traits)的頻譜特徵參數，愈往上層則是愈有個人特色的學習特性(Learned Traits)參數。雖說頻譜特徵參數容易受到環境變化干擾，但不同語者在頻譜上還

是能呈現其差異，更重要的是它使發聲器官的構造特性得以容易地自動擷取出來。反之，關於學習特性的特徵參數，如語意(Semantics)、用語(Diction)、發音(Pronunciation)等等，則和個人的成長環境、教育程度、家庭經濟等因數有較大的關係，可是儘管這類的特徵最能夠顯現出個人在言語上的特性，卻是相當不容易地自動擷取，因爲必須藉助較大量的訓練語料。在 [18] 專題研究計劃中，已特別針對語者辨識開發較高階層的特徵參數，並證實能夠輔助頻譜特徵爲基礎的辨識系統。



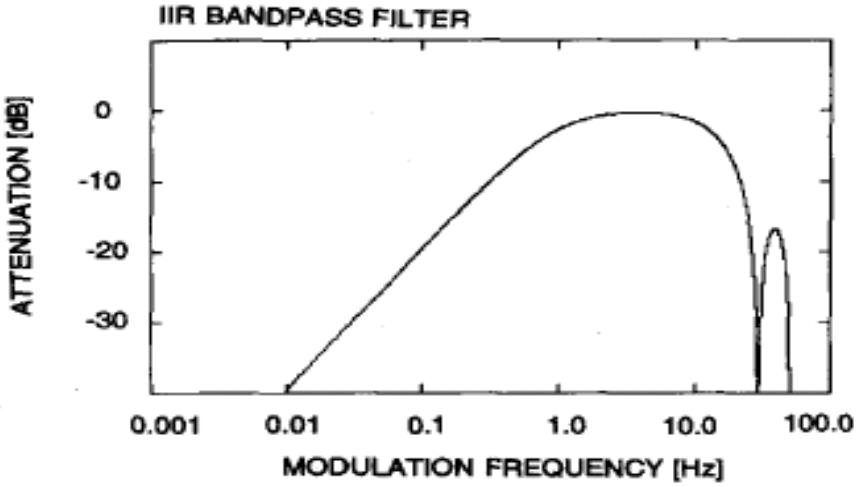
圖五、語者辨識使用資訊層級圖

6.2 特徵參數與分數正規化

在正規化方法中約略可分爲特徵參數與模型兩方面，其中特徵參數正規化最有名的莫過於早期由 S. Furui [19] 所提出的 Cepstral Mean Subtraction (CMS)方法，以及後來由 H. Hermansky 提出的 RelAtive SpecTrA(RASTA) filtering and spectral subtraction [20]。CMS 主要是將特徵參數減去其整體平均值(global mean)以消除線性通道效應和部份加乘性雜訊(Additive Noise)的邊際效應(Side Effects)，如下式子：

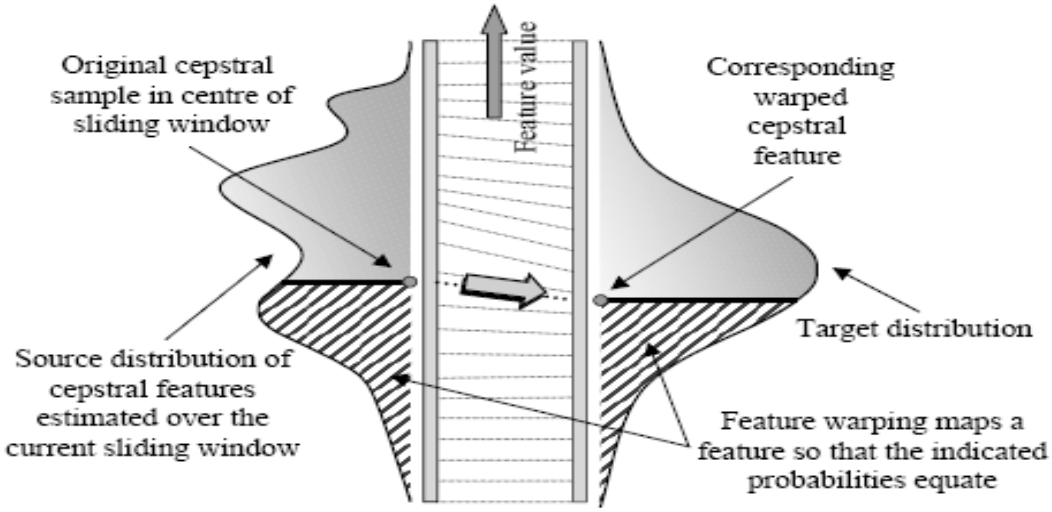
$$\vec{C}_{cms;i} = \vec{C}_{y;i} - \vec{C}_{y;avg} = \vec{C}_{y;i} - \frac{1}{N}\sum_{i=1}^{N}\vec{C}_{y;i} = \left(\vec{C}_{s;i} + \vec{C}_{h}\right) - \vec{C}_{h} = \vec{C}_{s;i}$$

而 RASTA filtering 方法則是對那些相較於一般語音變化程度過快或是過慢的頻譜成份作抑制，降低加乘性雜訊和捲積性噪音(Convolutional Noise)造成的失真，其使用之濾波器如圖六所示。



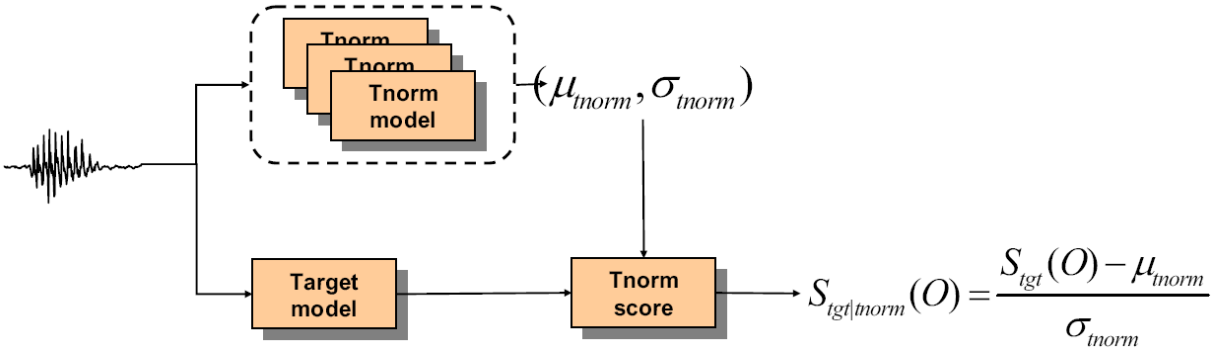
圖六、RASTA filtering 方法之帶通濾波器

近年來關於特徵參數正規化的研究陸續發表，且都一併考慮加乘性雜訊與捲積性噪音的影響，如 O. Viikki [21] 的 Cepstral Mean and Variance Normalization (CMVN)是合併特徵參數分佈範圍的正規化，改善 CMS 在通道效應作線性變化假設的缺失，後來更有 C. -P. Chen [22] 加入 ARMA filtering 使得這系列正規化方法有最完整的分析與進一步的改善。然而這類作法在語者辨識上也很可能破壞原有的語者資訊，因此 Jason Pelecanos [23] 所提出之 Feature Warping 可說是既能保留原先全部的資訊，亦達到對線性通道效應及小幅加乘性雜訊的強健性，如圖七所示。類似這種轉換特徵參數分佈的方法還有 Histogram Equalization (HEQ) [24-26] 和 short-time Gaussianization [27] 等等，而 Yanlu Xie [28, 29]的 Kurtosis Normalization 更加強 Gaussianization 方法的使用，讓語者驗證效果能再有一些進步。



圖七、Warping of features according to a target distribution shape.

另一方面是分數正規化方法，這類方法是對語者驗證所得到的分數作正規化轉換，目的在於強化驗證門檻值(Threshold)之效用，因爲根據 GMM-UBM 爲架構的語者驗證系統來說，接受與否的處理是利用宣稱語者(Claimed Speaker)模型與通用背景模型所得之相似度比值(Likelihood Ratio)，和此驗證門檻值比較後才做其決策 [10]，且這個驗證門檻值會受到驗證分數的變化影響而難以抉擇，而此驗證分數的變化包含由註冊語料多寡造成語者模型品質上的差異，以及測試環境不匹配。D. A. Reynolds [30] 的 Znorm 方法試圖調整冒充者分數之間的分佈差異，藉由計算目標語者(Target Speaker)和一組冒充者語句的分數，估測其平均值與標準差(Standard Deviation)來正規化目標語者原有的分數。另外 D. A. Reynolds [31] 亦有針對通道差異的問題提出 Hnorm 的改善方法，主要是嘗試移除和語者關聯的尺度與不同通道或麥克風造成的偏移量。然而這兩種分數正規化法所使用的參數都必須透過事先已知的冒充者語句獲得，因此 R. Auckenthaler [32] 的 Tnorm 方法改以在測試時才開始估測其必需的正規化參數，這使得目標語者分數的調整更符合測試時的環境。不過在 D. A. Reynolds [33] 的 ATnorm 仍針對 Tnorm 中固定使用同一組的冒充者模型做了一些改良，採用相似語者(Cohort Speaker)的方式讓每個目標語者都有各自的一組冒充者模型，如下圖八之分數正規化系統。



圖八、System for normalizing score distributions.

6.3 特徵參數與模型補償

相較於前兩大類型的方法，特徵參數與模型補償的研究在文獻上是比較少的。模型補償方面最具代表性的實屬 R. Teunen [34] 所做的 Speaker Model Synthesis (SMS)，主要是學習語者模型參數在不同通道之間的改變情形，並透過各個通道相關(Channel Dependent)的背景模型，將此資訊合成(Synthesize)出沒有訓練語料的通道特性，如此便能讓系統內的模型參數與測試環境的通道特性較爲一致。而考慮到每位語者使用的通道特性轉換應有所不同，Wei Wu [35] 修正 SMS 共用背景模型的缺點，使驗證效果改善不少。在特徵參數補償方面則是以 D. A. Reynolds [36] 提出的 Feature Mapping 最具特色，這個方法和模型補償的 SMS 有異曲同工之妙，同樣地藉由觀察特徵參數在通道無關與通道相關這兩種模型間的空間分

佈變化，得到特徵參數的投影函數。而 FU Zhonghua [37] 則合併考慮 Feature Mapping 中不同高斯分佈的影響，讓此方法的構思更完善地呈現。

7. 結語

在這篇文章中，介紹了目前語者辨識的基本作法，並說明當今語者辨識技術所面臨的難題，以及簡介不同層面的補償方法，雖然沒有任何一種方法可以保證完全解決訓練與測試環境不匹配的問題，但對於語者辨識技術確實有不錯的貢獻，倘若這般環境不匹配的困境能克服的話，語者辨識系統進入人類真實生活中的距離就不遠了。

參考文獻

[1] http://www.caip.rutgers.edu/news/bios/Joseph%20Campbell%20slides%2011-2003.pdf

[2] J. P. Campbell, "Speaker Recognition: A Tutorial," *Proc. of the IEEE*, 85(9), pp. 1437-1462, 1997.

[3] http://www.ll.mit.edu/IST/pubs/020314_Reynolds.pdf

[4] http://www.sdrforum.org/pages/sdr04/1.5%20Cognitive%20Radio%20Neel/1.5-2ppt%20Lewandowski.pdf

[5] NIST Speaker Recognition Evaluation http://www.nist.gov/speech/tests/spk/index.htm

[6] Linguistic Data Consortium Website http://www.ldc.upenn.edu/

[7] http://www.elra.info/

[8] http://www.iscslp2006.org/

[9] D. A. Reynolds, "Speaker Identification and Verification using Gaussian Mixture Speaker Models," *Speech Communication*, 17, pp. 91-108, 1995.

[10] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker Verification Using Adapted Gaussian Mixture Models," *Digital Signal Processing*, 10, pp. 19-41, 2000.

[11] M. W. Mak and S. Y. Kung, "Estimation of elliptical basis function parameters by the EM algorithms with application to speaker verification," *IEEE Transaction on Neural Networks*, 11, pp. 961-969, 2000.

[12] D. A. Reynolds, "HTIMIT and LLHDB: Speech corpora for the study of handset transducer effects," *Proc. ICASSP'97*, Ⅱ, pp. 1535-1538, 1997.

[13] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Transaction on Speech and Audio Processing*, 3, pp. 72-83, 1995.

[14] R. Mammone, X. Zhang and R. Ramachandran, "Robust speaker recognition – a feature-based approach," *IEEE Signal Processing Magazine*, pp. 58-71, 1996.

[15] H. A. Murthy, F. Beaufays, L. P Heck and M. Weintraub, "Robust Text-Independent Speaker Identification over Telephone Channels," *IEEE Transaction on Speech and Audio Processing*, 7, pp. 554-568, 1999.

[16] J. P. Campbell, D. A. Reynolds and R. B. Dunn, "Fusing high- and low-level features for speaker recognition," *Proc. Eurospeech'03*, pp. 2665-2668, 2003.

[17] M. Faundez-Zanuy and E. Monte-Moreno, "State-of-the-Art in Speaker Recognition," *IEEE Aerospace and Electronic Systems Magazine*, 20, pp. 7-12, 2005.

[18] SuperSID Website http://www.clsp.jhu.edu/ws2002/groups/supersid/

[19] S. Furui, “Cepstral analysis technique for automatic speaker verification,” *IEEE Transaction on Acoustics, Speech and Signal Processing*, 29, pp. 254-272, 1981.
[20] H. Hermansky and N. Morgan, “RASTA processing of speech,” *IEEE Transaction on Speech and Audio Processing*, 2, pp. 578-589, 1994.
[21] O. Viikki and K. Laurila, “Cepstral domain segmental feature vector normalization for noise robust speech recognition,” *Speech Communication*, 25, pp. 133-147, 1998.
[22] C. -P. Chen and J. Bilmes, “MVA Processing of Speech Features,” *IEEE Transaction on Audio, Speech and Language Processing*, 15, pp. 257-270, 2007.
[23] Pelecanos, Jason / Sridharan, Sridha (2001): "Feature warping for robust speaker verification", In *ODYSSEY-2001*, 213-218.
[24] S. Molau, M. Pitz and H. Ney, “Histogram Based Normalization in the Acoustic Feature Space,” *Proc. Automatic Speech Recognition and Understanding*, 2001, pp. 21-24, 2001.
[25] F. Hilger and H. Ney, “Quantile based histogram equalization for noise robust speech recognition,” *Proc. Eurospeech’01*, pp. 1135-1138, 2001.
[26] A. de la Torre et al., “Histogram equalization of the speech representation for robust speech recognition,” *IEEE Transaction on Speech and Audio Processing*, 13, pp. 355-366, 2005.
[27] B. Xing et al., “Short-time Gaussianization for robust speaker verification,” *Proc. ICASSP’02*, 1, pp. 681-684, 2002.
[28] Yanlu Xie, Beiqian Dai, Zhiqiang Yao and Minghui Liu, “Kurtosis Normalization in Feature Space for Robust Speaker Verification,” *Proc. ICASSP’06*, 1, pp. I-117 – I-120, 2006.
[29] Yanlu Xie, Beiqian Dai and Jun Sun, “Kurtosis Normalization after Short-Time Gaussianization for Robust Speaker Verification,” Proc. Intelligent Control and Automation, 2, pp. 9463-9467, 2006.
[30] D. A. Reynolds, “Comparison of background normalization methods for text-independent speaker verification,” *Proc. Eurospeech’97*, pp. 963-966, 1997.
[31] D. A. Reynolds, “The Effects of Handset Variability on Speaker Recognition Performance Experiments on the Switchboard Corpus,” *Proc. ICASSP’96*, I, pp. 113-117, 1996.
[32] R. Auckenthaler, M. Carey and H. Lloyd-Thomas, “Score normalization for text-independent speaker verification systems,” *Digital Signal Processing*, 10, pp. 42-54, 2000.
[33] D. E. Sturim and D. A. Reynolds, “Speaker adaptive cohort selection for Tnorm in text-independent speaker verification,” *Proc. ICASSP’05*, pp. I-741 – I-744, 2005.
[34] R. Teunen, B. Shahshahani and L. Heck, “A Model-based Transformational Approach to Robust Speaker Recognition,” *Proc. ISCSLP’00*, 2000.
[35] Wei Wu, T. F. Zheng and Mingxing Xu, “Cohort-based speaker model synthesis for channel robust speaker recognition,” *Proc. ICASSP’06*, pp. I-893 – I-896, 2006.
[36] D. A. Reynolds, “Channel robust speaker verification via feature mapping,” *Proc. ICASSP’03*, 10, pp. Ⅱ-53 - Ⅱ-56, 2003.
[37] Fu Zhonghua, Xie Lei and Zhao Rongchun, “Channel robust speaker verification via extended feature mapping,” *Proc. ICSP’04*, pp. 2417-2410, 2004.

中華民國計算語言學學會

出版品價目表及訂購單

編號	書　目	會　員	非會員	冊數	金額
1.	no.92-01, no. 92-04 (合訂本)　ICG 中的論旨角色 與 A conceptual Structure for Parsing Mandarin--its Frame and General Applications--	NT$ 80	NT$ 100	_____	_____
2.	no.92-02, no. 92-03 (合訂本) V-N 複合名詞討論篇 與V-R 複合動詞討論篇	120	150	_____	_____
3.	no.93-01　新聞語料庫字頻統計表	120	130	_____	_____
4.	no.93-02　新聞語料庫詞頻統計表	360	400	_____	_____
5.	no.93-03　新聞常用動詞詞頻與分類	180	200	_____	_____
6.	no.93-05　中文詞類分析	185	205	_____	_____
7.	no.93-06　現代漢語中的法相詞	40	50	_____	_____
8.	no.94-01　中文書面語頻率詞典（新聞語料詞頻統計）	380	450	_____	_____
9.	no.94-02　古漢語字頻表	180	200	_____	_____
10.	no.95-01　注音檢索現代漢語字頻表	75	85	_____	_____
11.	no.95-02/98-04　中央研究院平衡語料庫的內容與說明	75	85	_____	_____
12.	no.95-03　訊息爲本的格位語法與其剖析方法	75	80	_____	_____
13.	no.96-01　「搜」文解字－中文詞界研究與資訊用分詞標準	110	120	_____	_____
14.	no.97-01　古漢語詞頻表 (甲)	400	450	_____	_____
15.	no.97-02　論語詞頻表	90	100	_____	_____
16.	no.98-01　詞頻詞典	395	440	_____	_____
17.	no.98-02　Accumulated Word Frequency in CKIP Corpus	340	380	_____	_____
18.	no.98-03　自然語言處理及計算語言學相關術語中英對譯表	90	100	_____	_____
19.	no.02-01　現代漢語口語對話語料庫標註系統說明	75	85	_____	_____
20.	no.05-01　意義與詞義系列-中文意義與詞義第二版	400	500	_____	_____
21.	no.06-01　意義與詞義系列-中文意義與詞義第三版	600	800	_____	_____
22.	意義與詞義系列-中文意義與詞義 Working Paper	80	100	_____	_____
23.	意義與詞義系列-中文意義與詞義- 中文詞彙意義的區辨與描述原則	80	100		
24.	交談系統暨語境分析研討會講義 （中華民國計算語言學學會1997第四季學術活動）	130	150	_____	_____
25.	中文計算語言學期刊 (一年四期)　年份：_____ （過期期刊每本售價500元）	免費	2,500	_____	_____
26.	Readings of Chinese Language Processing	675	675	_____	_____
27.	剖析策略與機器翻譯 1990	150	165	_____	_____
28.	論文集 ROCLING　年份：_____	100	200	_____	_____
29.	論文集 COLING 2002 紙本+光碟片	300	400	_____	_____
30.	論文集 ISCSLP 2002 光碟片	100	200	_____	_____
31.	計算語言學學會通訊專刊《計算語言學領域介紹(一)》	100	200	_____	_____
			合　計	_____	_____

※　此價格表僅限國內使用

劃撥帳戶：中華民國計算語言學學會
劃撥帳號：19166251
聯絡電話：(02) 2788-3799 轉1502
E-mail: aclclp@hp.iis.sinica.edu.tw
聯絡人：何婉如　小姐 或 黃琪 小姐