

음성처리 평가기술 및 음성코퍼스 관련 국제활동 보고

이용주

원광대학교 컴퓨터공학과

Report on the international cooperative activities for speech I/O assessment technology and speech corpora

Yong-Ju Lee

Dept. of Computer Eng., Wonkwang University

1. 배경

음성언어정보 처리기술의 발전과 함께 최근 국내외에
서는 이를 상품화하기 위한 연구들이 활발히 진행되고
있다. 국내에서도 다양한 환경에서 음성인식을 실용화/
상품화하려는 움직임이 활발하다. 음성인식의 경우, 이
동단말기의 음성다이얼링, 전화음성에 의한 리모트 콘
트롤, 음성에 의한 정보검색, 자동차내의 음성명령, 가
전제품 및 게임기 등의 음성명령, 음성타자기(dictation)
등이 그 예가 될 것이다.

또한 합성의 경우도 이미 상당한 분야에서 실용화가
이루어지고 있고 품질개선 여하에 따라 그 응용의 폭은
확대되어 갈 것으로 보인다.

이러한 상품화 개발을 위해서는 이들 기술의 발전상
황 및 상호성능비교 등을 위한 객관적인 평가법의 필요
성이 대두되었다. 그러나 이러한 기술의 객관적인 평가
는 쉬운 문제가 아니다. 예를 들면 다양한 환경, 다양한
분야, 다양한 테스트 하에서 평가되어야 할 경우에 특히
그렇다. 더욱이 동일한 언어를 사용하지 않는 경우의
평가는 더욱 어렵고, 또한 인식이나 합성 모두, 어떤 언
어에 대해서 개발된 시스템과 다른 언어를 대상으로 개
발된 시스템간의 상호비교도 어려운 문제중의 하나이
다. 특히 음성인식에 있어서 시스템의 개발, 훈련, 테스
트 등에 필수적인 것이 많은 사람이 다양한 내용을 발

성한 음성으로 만들어진 음성 코퍼스(음성DB) 이다. 그
래서 여러 나라들에서는 공동으로 이용할 수 있는 음
성코퍼스를 국가주도 또는 연구자그룹이 공동으로 구축
하여 사용하고 있다. 이렇게 구축 초기부터 포맷이나
내용 그리고 레이블링 등에서 공동의 표준을 설정하여
구축한다고 하더라도 그러한 음성코퍼스의 상호 이용
에는 여러 가지 해결해야할 문제들이 있다 이러한 문제
는 정보화사회의 세계화 추세 속에서는 어떤 한 나라만
의 문제가 아니라 각국이 힘을 합하여 해결해나가야 한
다. 이러한 필요성에 따라 음성코퍼스 및 음성입출력
품질 평가법에 관한 국제적인 협력활동이 다음과 같은
회의들을 통해 꾸준히 논의되어왔다.

- o 1982년 3월18일~19일
미국 Gaithersburg (Dr. Dave Pallett 주관)
- o 1989년 9월20일~23일:
네델란드 Noordwijkerhout(ESCA ETRW, Prof.
Louis Pols 주관)
- o 1990년 11월23일~24일:
일본 코베 (ICSLP'90의 satellite event, prof. Fujisaki
주관)
- o 1991년 9월27일~28일:
이탈리아 Chiavari (EUROSPEECH'91의 satellite
event, Dr. Castagneri주관)

이러한 논의들을 통해 음성코퍼스 및 음성입출력 평가 방법에 관한 본격적인 조직인 COCOSDA(International Coordinating Committee on Speech Databases and Assessment)가 구성되었다. 구체적인 COCOSDA의 골격은 Chiavari workshop에서의 다음과 같은 결정에 의해 만들어졌다.

원래의 명칭은 “Committee for the international Coordination and Standardization of Speech Databases and Speech Input/Output Assessment Methods”이었다. 이 위원회의 목표는 음성언어 코퍼스와 음성입출력 평가법에 관한 국제적인 협력체의 설립 및 관련 활동의 지원이었다. 관심 있는 모든 나라들이 여기에 포함되었으며 이때 참여한 나라들은 미국(DARPA, LDC), 유럽(EuroCocosda, Eagles, ELRA), 호주, 중국, 프랑스, 독일, 한국, 일본, 네델란드, 스페인 캐나다 등이었다. 구조를 보면 COCOSDA는 central coordinating committee(CCC)와 음성인식, 음성합성, 음성corpora /레이블링 등 3개의 워킹그룹(WGs)으로 구성되었으며 각 그룹들은 북미, 유럽, 아시아태평양지역의 수명의 멤버로 구성되었다. CCC는 현재 7명의 멤버(M. Liberman, D. Pallet, J. Mariani, H. Fujisaki, A. Kurematsu, N. Campbell, B. Millar)로 구성되어 있으며 각 전문기관, 국내 및 국제적인 표준화기구 및 기금 출연기관들 사이의 공식적인 연락 및 조정 역할을 담당한다. COCOSDA 워크샵은 ICSLP와 EUROSPEECH와 연계하여 1년에 한 번씩 열린다.

1992년이후에 열린 워크샵은 다음과 같다.

- o 1992년 10월 12일~13일
캐나다 밴프(Banff) (ICSLP'92의 satellite event, Joseph Mariani 주관)
- o 1993년 9월 23일~25일
독일 베를린 (EUROSPEECH'93의 satellite event, Adrian Fourcin 주관)
- o 1994년 9월 22일~23일
일본 요코하마(ICSLP'94의 satellite event, Fujisaki & Fourcin 주관)
- o 1995년 9월 22일
스페인 마드리드(EUROSPEECH'95의 satellite event, Adrian Fourcin 주관)
- o 1996년 10월 7일
미국 필라델피아(ICSLP'96의 satellite event, Mark Liberman 주관)
- o 1997년 9월 26일~27일:
그리스 Rhodes (EUROSPEECH'97의 satellite event, Mark Liberman 주관)
- o 1998년 11월 30:
호주 시드니 (ICSLP'98의 satellite event, Mark Liberman 주관)

이러한 연례적인 워크샵에서는 관련 분야에 대한 각국의 진행상황의 발표와 토론 및 다음 연도까지의 활동계획 등을 논의하는 포럼 형식으로 진행된다.(<http://www.itl.atr.co.jp/cocosda>참조)

2. COCOSDA와 관련된 활동들

2.1 EuroCocosda

EuroCocosda는 main COCOSDA의 활동에 대한 유럽측에서의 contribution을 위해 조직된 것으로 EU가 지원하는 Linguistic Research Engineering(LRE)프로그램에 의해 1993년 부터 활동을 시작하였으며 영국, 프랑스, 네델란드, 이탈리아, 독일 등 5개국이 주 멤버로 활동하고 있다. 이곳에서는 한시적으로 다음과 같은 활동도 병행하였다.

- o COCOSDA가 주관하는 범세계적인 전화음성코퍼스 구축 계획인 POLYPHONE의 유럽측 담당
- o COCOSDA 음성합성코퍼스의 유럽측 framework구성
- o 신문에 대한 다국어 텍스트 및 음성코퍼스(NEWS) 구축을 위한 NL-Speech Initiative 유럽측 담당 등

2.2 Oriental COCOSDA

main COCOSDA에 대한 동양권 국가간의 연구협력의 목적으로 우리 나라, 일본, 중국, 대만 등이 주축이 되어 1998년부터 활동을 시작하였다. 이를 위한 준비모임이 1997년 3월 홍콩에서 있었고 (우리 나라에서는 저자참석), 이 모임에서 제1회 oriental COCOSDA workshop을 일본 쓰쿠바에서 1998년 5월에 개최하기로 하였다. 이 워크샵에서 주로 다룰 내용으로 다음과 같은 분야가 정해졌다.

- o Speech and Text Corpora
- o Speech Input systems assessment
- o Speech output systems assessment
- o Romanization of non-Roman characters
- o Orthographic System of oriental language
- o Prosodic and segmental notations

2.3 EAGLES(Expert Advisory Group on Language Engineering Standards)

1993년 2월에 EU의 지원으로 구성된 전문가 그룹으로 음성 및 언어를 포함하여 언어공학적인 응용을 위한 각종 spec의 가이드라인을 제시하는 것을 목적으로 하며 산하에 Text corpora, Lexicons, Formalizations, Evaluation, Speech 등 5개의 각종 워킹그룹이 있다. 주

된 역할은 다음과 같다.

- o 음성 및 언어과학에 대한 연구 및 응용에 대한 광범위한 컨설팅,
- o 기 구축된 음성언어자원의 평가 및 효율적인 구축을 위한 방법론
- o 음성언어자원 및 표준화의 관련된 기술분야의 정의
- o 이러한 활동의 가이드라인 제시를 위한 핸드북 작성
- o 음성연구자와 언어연구자간의 시설공동사용 및 협력 활동 강화
- o 국내외 관련기관과의 교류의 창구역할

그리고 Speech 워킹그룹 안에 음성언어코퍼스 관련 서브그룹이 활동한다. 이 그룹의 노력으로 음성언어시스템 관련 표준 및 resource에 관한 핸드북("Handbook of Standards and Resources for Spoken Language systems")이 발간되어 관련 연구자들에게 도움을 주고 있다. 다루고 있는 주요내용은 다음과 같다.

- o system design
- o corpus design
- o recording and documentation
- o spoken language lexicography
- o language model and physical characterization of corpora
- o assessment of speech recognizers, speaker verification devices, speech synthesizers and interactive dialogue systems
- o experimental assessment methodology, software tools
- o reference materials

2.4 ELRA(European Language Resources Association)

ELRA는 유럽에서의 음성언어 관련 자원의 작성, 검증, 배포를 촉진할 목적으로 1995년에 설립되었으며 사무국은 룩셈부르크에 있다. ELRA는 비영리조직으로 회원제이며, 보급뿐만 아니라 음성언어자원의 개발을 지원하는 역할도 한다. ELRA에서는 음성 및 텍스트에 관한 코퍼스, 문법, 사전, 전문용어 등 공유 가능한 자원들을 모두 다루고 있다. 주로 EU(European Union)가 지원하는 연구개발 프로젝트에서 작성된 자원들을 모아서 배포하고 있다.

(참고, <http://www.icp.grenet.fr/elra/home.html>)

2.5. LDC(Linguistic Data Consortium)

LDC는 대학, 기업, 정부연구기관의 컨소시엄 형태로 구성된 조직으로, 음성언어에 관한 연구에 필요한 각종 코퍼스를 공유할 수 있도록 서비스하고 있는 조직이다. LDC에서는 음성 및 텍스트에 관한 코퍼스, 사전, 그 밖의 공유 가능한 자원에 대해서 데이터의 수집, 작성, 배포를 담당하고 있으며 미국의 DARPA의 지원으로 1992년에 설립되었다. 현재 사무국은 펜실베이니아대학교에 위치하며 필요한 사용계약, 지적재산권 등에 관한 업무를 담당하고 있다. LDC는 회원제로서, 회원에 등록되면 필요한 코퍼스를 제공받을 수 있다.

(참고: <http://www ldc.upenn.edu>)

3. 관련 국제활동에 대한 우리의 대응

3.1 oriental COCOSDA 활동의 참여

제1회 oriental COCOSDA 워크샵이 쓰쿠바 대학의 Itahashi 교수 주관으로 일본 쓰쿠바에서 1998년 5월18일부터 19일까지 개최되어 4개국 53명의 연구자가 참가하였고 우리 나라에서는 본인을 포함하여 5명이 참석하였다. 참고로 국내참가자들의 발표내용은 다음과 같다.

- o 조철우
Aspects on the collection and applications of Multimedia Emotional speech database
- o 이용주
The Current Status of Korean Speech Corpora
- o 이숙향
Prosodic and Segmental Labelling System for Korean
- o 문승재
The Orthographic System and the Romanization of Korean

제2회의 워크샵은 1999년 5월13일부터 14일까지 대만에서 Lin-shan Lee 교수 주관으로 개최 되었는데, 10개국의 126명이 참석하였으며, 우리 나라에서도 본인을 포함하여 5명의 연구자가 참가하였다.

국내 연구자의 발표내용은 다음과 같다.

- o 이용주, 최기선
Some activities of Korean COCOSDA and Korean text corpora
- o 조철우
Collection and analysis of Korean emotion speech
- o 조철우, Bruce Miller 외

Collection of the Korean auditory-visual speech data

- o 이상호, 오영환
CART-based modelling of Korean segmental duration
- o 박전규, 박동인 외
Corpus-based language model and speech database for Korean dictation.

아울러, 다음과 같은 주제의 패널 토의에도 본인이 참석하였다.

Next steps for promoting the speech research on East-Asian languages

- o common corpora, shared software, joint assessment, collaborated projects
- o Ideas from trends in new technology developments
- o What's the right model for East-Asia considering experiences in North America and Europe?

이상의 주제로 유익한 토론이 이루어졌으며 현재도 이를 구체화하기 위한 토의가 E-mail을 통해 계속 이루어지고 있다.

앞으로, 제3회 워크샵은 중국과 일본 공동 주관으로 ICSLP 2000과 연계하여 개최할 예정이며, 제4회는 2001년에 한국에서 개최할 예정이다.

3.2 Korean COCODSA의 결성

이상과 같은 국제적인 협력활동의 추세와 함께 국내에서도 관련 활동의 활성화가 시급히 요청되고 있다. 특히 음성인식 및 합성 등 음성정보처리 전반에 걸쳐서 관련기술의 객관적인 성능평가체계의 확립과 공동의 연구기반인 음성코퍼스의 효율적인 확보를 위한 표준화 활동, 그리고 이를 효과적으로 공동 이용할 수 있는 국내의 보급체계의 구축이 시급한 실정이다. 이를 위해서는 COCODSA 등 국제적인 단체들과 협력을 강화하고 지속적인 관심과 연구활동이 우리 나라대로 이루어져야 할 것이다.

이를 위해서는 소수의 연구자들만의 단편적인 활동을 탈피하여 체계적이고 지속적이며 대표성이 있는 담당조직의 구성이 필요하다. 이는 어떤 면에서는 국제적인 표준화활동과 유사한 점도 있으나 우리말을 대상으로 하는 것인 만큼 공통으로 활용 가능한 외국에서의 경험을 적극 받아들이는 창구로도 활용될 수 있을 것이며 우리말 연구의 국제화에도 기여하게 될 것이다. 따라서 위와 같은 국제 학술활동에 적극 대처하기 위하여 1999년 4월에 뜻을 같이 하는 연구자들을 중심으로 Korean

COCOSDA가 결성되어 그 활동을 시작하였다.

Korean COCODSA는 Central coordinating committee를 중심으로, 음성인식, 합성, 음성코퍼스 등의 3개의 sub-committee로 구성되었다. 물론, 목표는 국내에서의 음성정보처리기술의 객관적 성능평가를 위한 방법론의 토의와 평가법의 권고, 체계적인 음성코퍼스의 구축을 위한 방법론의 검토와 제안 등을 목표로 한다. 아울러 관련 분야의 활성화를 위한 연구계획의 수립, 연구환경 조성을 위한 제안 뿐만 아니라, 음성코퍼스의 경우는 적극적인 구축 및 보급을 위한 체계의 정립에도 노력하고자 한다.

아직은 초기단계인 상태이므로 관심있는 많은 연구자의 적극적인 참여가 요청된다.

참고문헌

(* 본 보고는 1998년도에 음성통신 및 신호처리 워크샵에서 보고된 것에 최근의 활동을 보완한 것임)

- [1] 이용주: 음성DB 및 음성입출력평가법 관련 국제협력활동 소개 및 제안, 한국음향학회 음성통신 및 신호처리 워크샵 논문집(1998)
- [2] H. Fujisaki, Overview of International Efforts towards Coordination and standardization of speech databases and speech Input/Output Assessment Methods, Proc. of Oriental COCODSA workshop'98 (May 1998)
- [3] 이용주, 음성언어코퍼스, 한국정보과학회지 제16권 제2호 (1998, 2.)
- [4] 조철우, "음성DB구축을 위한 국제간 활동현황", 한국음향학회 제12회 음성통신 및 신호처리워크샵 논문집 (1995, 6.)
- [5] Pallett, D.S. (ed.), Proceeding of the Workshop on Standardization for Speech I/O Technology, Gaithersburg, MD, U.S.A., March 18-19, 1992.
- [6] Pols, L.C.W. (ed.), Proceedings of ESCA Tutorial Day and Workshop on Speech Input/Output Assessment and Speech Databases, Noordwijkerhout, the Netherlands, September 20-23, 1989.
- [7] Fujisaki, H. (ed.), Preprints of the International Symposium on International Coordination and Standardization of Speech Databases and Assessment Techniques for Speech Input/Output, Kobe, Japan, November 23, 1990.
- [8] Fujisaki, H. (ed.), Preprints of the International Workshop on International Coordination and Standardization of Speech Databases and

- Assessment Techniques for Speech Input/Output, Kobe, Japan, November 23, 1990.
- [9] Castagneri, G. (ed.), Preprints of the Workshop on International Coordination and Standardization of Speech Databases and Speech I/O Assessment Methods, Chiavari, Italy, September 27-28, 1991.
- [10] Jones, K. and Mariani, J. (eds.), Proceedings of the 1992 Workshop of the International Coordinating Committee on Speech Databases and Speech I/O Systems Assessment, Banff, Canada, October 12, 1992.
- [11] Fourcin, A. (ed.), Report on the COCOSDA Workshop 93, Berlin, Germany, September 23-25, 1993.
- [12] Forucin, A. and Fujisaki, H. (eds.), Notes from the COCOSDA Workshop 94 Yokohama, Japan, September 22-23, 1994.
- [13] Fourcin, A. and Fujisaki, H. (eds.), Notes from the COCOSDA Workshop 95, Madrid, Spain, September 22, 1995.
- [14] Gibbon, D. Moore, R. and Winski, R. (eds.), Handbook of Standards and Resources for Spoken Language Systems, Mouton de Gruyter, Berlin & New York, 1997.
- [15] Choukri, K. (ed.), The ELRA Newsletter, vol 1, no. 1, March 1996.
- [16] Dafydd Gibbon, Roger Moore, Richard Winski, Handbook of Standards and Resources for Spoken Language Systems, Mouton de Gruyter 1997
- [17] 김봉완, 김종진, 김선태, 이용주, “공동이용을 위한 음성DB의 설계 및 구축에 관한 연구” 한국음향학회 논문지 제16권 제4호 1997. 6
- [18] Proceedings of First international workshop on east-asian language resources and evaluation - Oriental COCOSDA workshop 98 - (1998.5.)
- [19] Proceedings of second international workshop on east-asian language resources and evaluation - Oriental COCOSDA'99 - (1999.5.)