

Oriental COCOSDA : 音声データベースおよび音声入出力 システム評価国際協調委員会 東アジア部会について

板橋秀一^{†*} 中村 哲^{†+}

[†] 情報システム研究機構 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

*産業技術総合研究所 情報技術研究部門 〒305-8568 つくば市梅園 1-1-1

+情報通信研究機構 知識創成コミュニケーション研究センター † ATR 音声言語コミュニケーション研究所

〒169-0288 京都府相楽郡精華町光台 2-2-2

E-mail: † itabashi@nii.ac.jp, ‡ satoshi.nakamura@atr.jp

あらまし Oriental COCOSDA は、東アジア言語の話し言葉コーパスの構築、利用、普及に関する意見を交換し情報を共有して、地域的な問題や音声認識・合成システムの評価方法について議論するとともに、東アジア言語に関する音声研究を推進することを目的として設立された。1997年の準備会議以来、Oriental COCOSDA Workshop が「東アジア言語資源・評価国際ワークショップ」として毎年開催されている。開催地は香港での準備会以降、日本、台湾、中国、韓国、タイ、シンガポール、インド、インドネシア、マレーシアとなっている。Oriental COCOSDA は代表1名、アドバイザー3名、地域代表26名で運営されている。幾つかの国では既にコーパス関連の活動が行われているが、アジア地域における研究機関・コンソーシアムとのより広範な協力が必要である。最近は上記の開催地以外の地域においても音声研究の進展が見られるので、今後はこれらの地域での音声研究推進の一助となるようワークショップの開催を考えている。

キーワード 音声、データベース、コーパス、システム、評価

On Oriental COCOSDA: East-Asian Chapter of International Coordinating Committee on Speech Databases and Speech Input/Output Systems Assessment

Shuichi ITAHASHI^{†*} and Satoshi NAKAMURA^{†+}

[†] National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan

*National Inst. of Advanced Industrial Science and Technology 1-1-1 Umezono, Tsukuba, Japan

‡ ATR Spoken Language Communication Res. Labs. 2-2-2 Seika-cho, Souraku-gun, Kyoto Japan

+National Inst. of Inform. and Communications Tech. 3-5 Seika-cho, Soraku-gun, Kyoto, Japan

E-mail: † itabashi@nii.ac.jp, ‡ satoshi.nakamura@atr.jp

Abstract The purpose of Oriental COCOSDA is to exchange ideas, to share information and to discuss regional matters on creation, utilization, dissemination of spoken language corpora of oriental languages and also on the assessment methods of speech recognition/synthesis systems as well as to promote speech research on oriental languages. A series of International Workshop on East Asian Language Resources and Evaluation (EALREW) or Oriental COCOSDA Workshop has been held annually since the preparatory meeting held in 1997. After that, we have had a series of workshops every year in Japan, Taiwan, China, Korea, Thailand, Singapore, India, Indonesia, and Malaysia. The Oriental COCOSDA is managed by a convener, three advisory members, and 26 representatives from 13 regions in Oriental countries. We need much more Pan-Asia collaboration with research organizations and consortia, though there are some domestic activities in Oriental countries. We note that speech research has become popular gradually in Oriental countries where the Workshop has not been held, including Vietnam and Nepal. We plan to hold future Oriental COCOSDA meetings in these new places in order to promote speech research there.

Keyword Speech, Database, Corpus, System, Evaluation

1. はじめに

音声情報処理の研究・開発において、多様な種類の音声データを収集・保管し、誰でも音声処理システムの研究・開発や認識装置性能評価に使えるようにすることが必要であることは広く知られている。共通の音声データを用いることで、提案された手法の追試が可能となり音声研究の客観性が増す。一方、言語学的、文化的な面からは種々の言語、特に危機言語と呼ばれているものを保存することが必要かつ重要である。少数話者の言語や方言が日々消滅していることはしばしば報じられている。そのような言語の自然な記録を保存することは緊急の課題である。これは音声データベースのもう一つの重要な側面である。このような目的に利用される音声データの集積は音声データベースまたは音声コーパスと呼ばれている（図1参照）[1]。

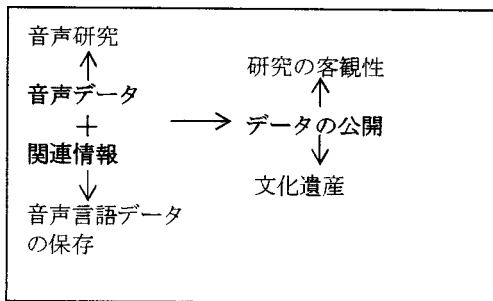


図1 音声コーパスの必要性

音声コーパス構築の国際協力と音声入出力システムの評価方法に関する国際協調を推進するためにCOCOSDA（音声データベースおよび音声入出力システム評価国際協調委員会）が1991年に設立された[2][3]。これはInternational Coordinating Committee on Speech Databases and Speech Input/Output Systems Assessmentの略称である。音声言語処理国際会議(ICSLP)とヨーロッパ音声通信・工学国際会議(EUROSPEECH)（現在は統合されてINTERSPEECH）の会議にあわせてワークショップを開催している。1993年にはその下部組織としてEuroCOCOSDAが設立され、後にOriental COCOSDA（以下O-Cと略称することがある）が組織されたが、それについて以下に

詳しく述べる[4]。以下、第2節ではO-Cの経緯を述べ、第3節でO-Cの組織、第4節ではこれまでの年次大会を概観し、第5節で将来展望を行い、第6節でまとめを行う。

2. これまでの経緯

1994年に横浜で開催されたCOCOSDAワークショップの際に、東アジア地域において、音声言語処理に関する意見交換、情報共有、地域問題の議論のための組織が必要であることを著者が提案した。東アジアの言語は多様な特徴を有しているため、以下のようにヨーロッパの諸言語とは異なる問題を有している。

- (1) 種々の語族に属する言語が多いため多様性が大きい。例えば、

オーストロネシア語族：マレー語、インドネシア語
 シナ・チベット語族：中国語、チベット語、ビルマ語
 オーストロ・アジア語族：クメール語、ベトナム語
 タイ・カダイ語族：タイ語、ラオ語
 ドラヴィダ語族：タミル語、テルグ語
 アルタイ語族：チュルク語、モンゴル語、(朝鮮語)
 インドヨーロッパ語族：ヒンディー語、ウルドゥー語

- (2) 漢字、ハングル、かな文字など表記法が多様。
 固有文字：漢字、かな、ハングル、タイ文字、ビルマ文字、クメール文字等。
 ローマ字：インドネシア語、マレー語、ベトナム語等。
- (3) 表記において単語間に空白を置かない言語がある（中国語、日本語、タイ語など）。
- (4) 音調言語（中国語、ビルマ語、ラオ語など）。
- (5) 多様な語順（SOV, SVO, VSO, VOS、ただし、S:主語、V:述語、O:目的語）。
- (6) ローマ字による表記法にも多様性がある。

これらの特徴を持つ言語を処理するためには、ヨーロッパの諸言語の処理に適したものとは別の、より適した方法があると考えるのは自然なことであろう。

研究・開発の過程および種々の音声システムの性能評価における利用のために、共通利用可能な種々の音声コーパスを構築しその利用システムを設立することが必要なことは東アジア地域でも認識されていた。

しかしながら、この「認識」を具体化する活動は東アジア地域では規模が小さく、なかなかまとまらない状況であった。我々は、学問的重要性のみならず産業面での国際協調の点からも、これらの努力をまとめていく必要性を感じた。

東アジアの幾つかの国々では音声研究が次第に進展しており、音声（コーパス）に関する組織がすでにあったが、互いの連携は殆どなかった。これらのことを考慮すると、将来の音声および関連分野の研究の進歩のために、音声データを収集・構築、保存、配付、共有するための共通の枠組みを徐々に準備することが必要であった。

このような組織を設立して音声コーパス、音声認識・合成、音声入出力システム評価方法に関連した問題の調整にあたることに中国、韓国、日本の研究者が合意した。これが Oriental COCOSDA である。当初は Asian COCOSDA とすることを考えていたが、Asian とすると中東まで広がって範囲が広すぎるということから Oriental とすることになった。

Oriental COCOSDA の目的は、東アジア言語の話し言葉コーパスの構築、利用、普及に関する意見を交換し、情報を共有して地域的な問題や音声認識・合成システムの評価方法について議論するとともに、東アジア言語に関する音声研究を推進することにある。O-C の準備会は香港大学で 1997 年に行われた。それ以来、日本、台湾、中国、韓国、タイ、シンガポール、インド、インドネシア、マレーシアで毎年ワークショップを開催してきた。

3. 組織

Oriental COCOSDA は代表者 1 名および中国、日本、韓国からの 3 人のアドバイザーと 13 の地域（中国、香港、インド、インドネシア、日本、韓国、マレーシア、モンゴル、ネパール、シンガポール、台湾、タイ、ベトナム）からの 26 人の地域代表で運営されている。初代代表は本稿第一著者の板橋秀一（1997-2005）、第 2 代代表は台湾科学アカデミーの Chiu-yu Tseng 博士（2006 年以降）であり、アドバイザーは、藤崎博也博士（東京大学名誉教授）、Jialu Zhang 教授（元中国科学院）、韓国の Sougil Ann 博士（ソウル大学）である。

4. 年次大会

Oriental COCOSDA の準備会は中国、香港、日本、韓国、台湾の 5 地域から 11 名が集まり 1997 年に香港大学で開催された。藤崎博也東京大学名誉教授が COCOSDA の活動を概観し、コーパス研究における一般のおよび地域的問題を指摘した。藤崎教授はまた Oriental ということが地域と言語の二重の意味を持つことを指摘した。O-C のメンバーは東アジア地域に住んでいて、その言語を話しかつ研究している人か、東アジア言語のコーパスや音声入出力システムの標準化に興味を持つ人ということになる。O-C は COCOSDA に出席してその活動を報告し議論に参加するという意味で、COCOSDA の下部組織であることが了解されている。また、第 1 回大会を 1998 年 5 月につくばで開催することが合意された。これまで開催された各大会の論文数、参加者数、参加地域数の概略は表 1 の通りである。5 回目以降は会議期間中に地域代表者の会議が開かれ、次回開催地の決定などを行っている。表 1 で見るように、会議の規模は大きくはないが全員参加のもとに質疑応答が活発に行われ、また代表者会議では次回開催の誘致活動が行われるなど、参加者の関心の高さが見られる。

表 1 年次大会の概要

開催地	論文数	参加者	参加地域
1. つくば（日本）	30	54	4
2. 台北（台湾）	44	120	7
3. 北京（中国）	8	20	6
4. デジョン（韓国）	11	25	6
5. ホアヒン（タイ）	24	96	8
6. シンガポール（シンガポール）	28	60	8
7. ニューデリー（インド）	55	150	10
8. ジャカルタ（インドネシア）	24	65	9
9. ペナン（マレーシア）	33	60	9

5. 将来の展望

第 10 回大会は 2007 年 11 月にベトナムのハノイで開催予定である。その翌年の 2008 年は第 1 回大会から丁度 10 年目であることから日本で開催することになり、

2008年11月に京都で開催される予定である。ネパール、新疆ウイグル自治区等で音声研究が次第に活発になってきていることから、東アジアにおける音声研究の推進を図るために、それらの地域でのO-C開催を計画している。O-Cの振興および音声言語処理に関する研究推進を図るために、他の関連会議との合同セッションや特別セッションを計画している。これまでもタイ、シンガポール、インドの会議では、関連する会議と共催の形で開催した実績がある。さらに、2006年12月に開催された中国語音声言語処理国際シンポジウム(ISCSP)では特別セッション「多言語音声コーパス開発」の開設に協力した。また、設立10周年を記念して、音声コーパスと標準化に関する書籍の出版が計画されている。

音声コーパスに関連する組織として、日本では言語資源協会(GSK)が1999年に設立され[5]、韓国では音声情報工学・産業振興センター(SITEC)が2001年に[6]、中国LDCが2002年に設立された[7]。中国には中国コーパスコンソーシアム(CCC)もある。日本ではさらに国立情報学研究所(NII)が2006年に音声資源コンソーシアム(NII-SRC)を立ち上げた[8]。GSKはテキストコーパスを主に扱い、NII-SRCは音声コーパスを扱うことになっている。今後これらの組織間のネットワークの構築が課題となろう。また、このような組織間協力の具体的な目標としては、意味的に類似した内容の多言語音声コーパスの構築が上げられる。

6. おわりに

本稿ではOriental COCOSAの概要を紹介した。関連する組織として、アメリカの言語資源コンソーシアム(LDC)およびヨーロッパ言語資源協会(ELRA)は音声コーパスの構築、収集、配付に多大な貢献をしてきた。GSK, SITEC, Chinese LDC, NII-SRCはアジアにおけるLDC, ELRAの役割を果たすことが期待されている。アジア諸言語の連続音声コーパス特に自発音声コーパスはまだこれからというところである。アジアの幾つかの言語では音調やピッチアクセントが重要な音声学的特徴であり、多数の言語がローマ字以外の文字を表記に用いている。このような特徴をもつアジアの諸言語の音声・言語コーパスの収集と利用を図ることは今後

ますます重要になるものと考えられる。

最近、日本の国立情報学研究所がNII-SRCを立ち上げて、GSKと協力しながら音声・言語コーパス事業に着手したことは、日本のみならず東アジアにおける音声・言語資源共有と標準化活動を強化することになる。Oriental COCOSAが東アジアにおいてその目標を追求してさらに重要な役割を果たすことを念願している。

付録として音声入出力関連の標準化の状況とOriental COCOSA年次大会の概要を付けたので参照されたい。なお、本稿は文献4の一部の和訳をもとに加筆・修正したものである。

文 献

- [1] S. Itahashi, "Overview of the Asian Activities on Speech Corpora and Standardization," Proc. iSTRANS-2004 and Oriental COCOSA 2004, Delhi, India, pp. 3-11(Nov. 2004).
- [2] 板橋秀一, "COCOSA: 音声データベース及び音声入出力システム評価の国際協調委員会について" 日本音響学会誌, 53巻, 12号. pp. 949-954 (Dec. 1997).
- [3] N. Campbell, "COCOSA - A Progress Report," Proc. LREC 2000, Athens, Greece, pp. 73-76 (May 2000).
- [4] S. Itahashi, C-Y Tseng, S. Nakamura, "Oriental COCOSA: Past, Present and Future," Proc. LREC 2006, Genoa, Italy, pp. 753-756 (May 2006).
- [5] 田中徳積, "言語資源協会(GSK)について" <http://www.gsk.or.jp/>
- [6] Y-J Lee, "Speech Information Technology & Industry Promotion Center (SITEC)," Proc. 18th ICA, Kyoto, Japan, pp. I-397 - I-400 (2004).
- [7] Bo Xu, "Chinese Activities on Sharing Linguistic Data," Proc. 18th ICA, Kyoto, Japan, pp. I-545 - I-548 (2004).
- [8] S. Itahashi, T. Ohsuga, "Introduction of NII-Speech Resources Consortium," Proc. Oriental COCOSA 2006, Penang, Malaysia, pp. 38-43 (Dec. 2006).
- [9] J-L Zhang, S-W Dong, "Guidelines to Assessment of Speech Synthesis System for Chinese," Proc. Oriental COCOSA'98, Tsukuba, Japan, pp. 46-50 (May 1998).
- [10] S. Itahashi, M. Akabane, Y. Kitahara, K. Tanaka, "Standard for Japanese Text-to-Speech Speech Synthesizer," Proc. Oriental COCOSA 2004, Delhi, India, pp. 58-63 (Nov. 2004).
- [11] K. Tanaka, M. Akabane, T. Minowa, S. Itahashi, "JEIDA Standard of Symbols for Japanese Text-to-Speech Synthesizer," Proc. Oriental COCOSA 2000, Beijing, China, pp. 27-32 (Oct. 2000).
- [12] R-H Wang, "National Performance Assessment of Speech Recognition System for Chinese," Proc. Oriental COCOSA'99, Taipei, Taiwan, pp. 41-44 (May 1999).

[13] M-W Koo, "An Assessment Tool for Speech recognition System over Telephone Network," Proc. Oriental COCOSDA 2001, Taejon, Korea, pp. 136-141 (Aug. 2001).

[14] B-S Lin, L-S Lee, "Early Stage Evaluation of Spoken Dialogue System Performance Based on Quantitative Simulations," Proc. Oriental COCOSDA 2001, Taejon, Korea, pp. 126-135 (Aug. 2001).

付録1：音声関連標準化の状況

音声合成装置の性能評価については、中国科学院の報告[9]、日本の電子情報技術産業協会(JEITA)の音声入出力方式標準化委員会音声合成関連の規格等が上げられる[10][11]。音声認識関連では、JEITAの音声認識関連の規格、中国科学技術大学[12]や韓国[13]、台湾[14]の報告等があげられる。また、公開ソフトウェアとしては、日本の音声認識エンジン Julius/Julian や音声合成ツールキット HTS、擬人化エージェント GALATEAなどが上げられる。

A1-1 電子情報技術産業協会規格等

1) JEITA IT-4001：音声合成システム性能評価方法 (2003年2月)

2) JEITA IT-4002：日本語テキスト音声合成用記号 (2005年3月)

3) JEITA IT-4003：日本語音声認識用読み記号 (2005年3月)

4) JEITA IT-4005：日本語音声認識エンジン性能評価方法 (2007年3月)

5) JEITA ガイドライン：日本語カーナビゲーション音声コマンド設計ガイドライン (2007年3月)

6) JEITA TT-6004：ITS 車載器用音声合成記号 (2007年3月)

A1-2 公開ソフトウェア

1) <http://julius.sourceforge.jp/>

2) <http://hil.t.u-tokyo.ac.jp/~galatea/index-jp.html>

3) <http://hts.ics.nitech.ac.jp/>

付録2：Oriental COCOSDA 年次大会の概要

A2-1 つくば大会(1998)

招待講演

藤崎博也：International Efforts toward Coordination of

Speech Databases and Speech Input/Output Systems Assessment

鈴木英一：The Use of Linguistic Database for Linguistic Research

セッション：コーパス(合成, 認識, 対話・テキスト), 評価(合成, 認識), 正書法・ローマ字化, 韻律表記

A2-2 台北大会(1999)

招待講演

1. 藤崎博也 "Information Retrieval based on human-machine dialogue"

2. B-H Juang "Experiment Design for Speech Recognition and Understanding"

3. K. Choukri "European Language Resources Association - History and Recent Developments"

4. 鹿野清宏 "Volunteer-based IPA Japanese Dictation Free Software Project"

セッション：音声コーパス, 合成・認識評価, 音声学, セグメンテーションとラベリング, モデルとシステム, コーパス関連トピックス

A2-3 北京大会(2000)

テーマ：中国, 韓国, タイ, 台湾における音声コーパスと評価活動, 日本語音声合成用記号の JEIDA 規格, モンゴル語音声データベース, 中国語における強調のラベリング, COCOSDA の新体制

A2-4 デジョン大会(2001)

テーマ：音声情報工学・産業振興センター(SITEC)の紹介, 日本語文献検索音声対話コーパス, 中国語会話音声コーパス, オーストラリア, タイ, 台湾における音声コーパスの状況, Korterm の紹介, 感情音声プロジェクト, 音声処理ツールの比較, 音声対話システム性能評価, 電話系音声認識システム評価

A2-5 ホアヒン大会(2002)

第5回自然言語処理シンポジウム(SNLP)と共催

招待講演 (音声関係のみ)

1. 藤崎博也：Toward an Intelligent System for Academic Information Retrieval Through

Human-Machine Spoken Dialogue

2. Fanxing Chen, "Issues in Speech Synthesis for Tonal Languages"
3. 藤崎博也: Modeling in the Study of Tonal Feature of Speech with Applications to Multilingual Speech Synthesis

セッション: 音声処理 (認識・合成・セグメンテーション・韻律), コーパス

A2-6 シンガポール大会(2003)

第 17 回言語・情報・計算太平洋アジア会議(PACLIC17)と共催

招待講演

1. 板橋秀一: Overview of the East-Asian Activities on Speech Corpora and Assessment
2. Jie Xu: Focus Marking in Chinese and Malay – A Comparative Perspective
3. T. Fang Zheng: Making Full Use of Speech Corpora
4. K-S Choi: Virtual Linked Lexical Knowledge Base for Causality Reasoning

セッション: 音声技術評価, コーパス, 音声入出力, 東アジア言語の音声体系

A2-7 ニューデリー大会(2004)

音声工学・処理システム国際シンポジウム(iSTEPS-2004)および機械翻訳, 自然言語処理, 翻訳サポートシステム国際シンポジウム(iSTRANS-2004)と共催

招待講演 (音声関係のみ)

1. 板橋秀一: Overview of the Asian Activities on Speech Corpora and Standardization
2. 藤崎博也: Problems and Prospects of Spoken Language Processing
3. P. V. S. Rao: An Approach towards Automatic Evaluation of Accent and Style
4. Tanja Schultz: Independent and Language Adaptive

Speech Processing

5. Kuldip Paliwal: Importance of Phase Spectrum in Speech Processing
6. P. K. Saxena: Securing Speech Communication
7. Y. Sagisaka: Spoken Language Processing as Computational Human Modeling

テーマ等: 音声・言語工学の現状と将来展望, 音声言語による人間・機械通信, 産・官・学相互協力, インド言語の音声コーパス, 雑音下音声認識性能評価, 音声合成装置性能評価, 各種音声コーパス, 音声・楽器・歌声の識別

A2-8 ジャカルタ大会(2005)

招待講演

1. 中村 哲: Corpus and Technologies for ATR Speech-to-Speech Translation Research
2. A. A. Arman: Characteristics of Indonesian Language from the Perspective of Language Technology

セッション: コーパスと技術, 音声認識, 話者認識・感情検出, 音声合成, 音声対話システム

A2-9 ペナン大会(2006)

招待講演

1. Y-J Lee: Creation and Distribution of Language Resources in Korea
2. 広瀬啓吉: Corpus-based Synthesis of Fundamental Frequency Contours Using Generation Model and Automatic Preparation of training Corpora
3. S. S. Agrawal: Issues in the Development, Design and Standardization of Speech Corpora for Indian Languages

セッション: コーパスと技術, 感情と音声認識, 音声合成