INTERSPEECH 2010 Satellite Workshop on

# "Second Language Studies: Acquisition, Learning, Education and Technology"


## L2WS'2010


## Technical Program and Abstracts

**Co-organized by AESOP, SLaTE, NICT & LASS**
**22 to 24 September 2010**
**The International Conference Center,**
**Waseda University, Tokyo, Japan**

This workshop is co-organized by
AESOP (Asian English Speech cOrpus Project),
SLaTE (the ISCA SIG on Speech and Language Technology in Education),
NICT (National Institute of Information and Communications Technology), and
LASS (Language and Speech Science Research Laboratories).

This workshop is sponsored by SCAT, NICT, CCDL (Cross-Cultural Distance Learning Research Center), and International Office of Waseda University.

# Messages from the General Chair

Welcome to L2WS. The aim of this workshop is for people working in speech science and engineering, linguistics, psycholinguistics, language education to get together and discuss L2 acquisition in the classroom settings: the problems that classroom researcher-teachers are facing in their daily practice and to what extent technology can help us improve our classroom practice. Therefore, this workshop is not only for speech engineers but also for researcher-teachers (applied linguists). We aim to enhance didactic competences among us.

We have also organized a pre-workshop symposium: An Open Symposium on "Primary School Education in Asia." English Education is going to be introduced in Japanese primary schools in 2011. We have invited the experts on primary school English Education from China, Korea, Taiwan and Japan. Their talk covers brief history of primary education in each country, unique teaching methodologies for young learners and their learning processes. They are also invited to hold special sessions on the second day of our workshop. During our workshop, we have arranged technical presentations and empirical investigations in an balanced way so that we can learn from each other.

I would like to thank the SLaTE committee, AESOP, NICT, LASS and CCDL for their support, and the reviewers for reading and assessing their allocated papers.

I would like to especially thank Nobuaki Minematsu and Mariko Kondo for organizing this wonderful workshop! It is well-planned and is full of interesting events, papers and demos. I hope you all enjoy this workshop and find many new ideas to take home with you.

Michiko Nakano,
Chair, L2WS'2010

# Messages from the Local Organizers

We are very happy to welcome you to the first workshop on Second Language Studies: Acquisition, Learning, Education and Technology. As the long name suggests, the workshop aims to be truly interdisciplinary. The idea for the workshop was initially conceived by researchers involved in second language acquisition and the second language speech corpus. The team was interested in collaborating with language education specialists and education software engineers to host an interdisciplinary workshop on second language studies. We are interested in various areas of L2 research and are keen to learn about new areas, but it is not always easy to step out of our own academic discipline. We tend to stick to our own areas and are shy to adventure into a new world. So the aim of this workshop is to provide an opportunity for all of us working in different fields to meet and learn what other people are doing.

Sixty five papers were submitted on all areas of second language research and we have selected 43 papers for presentation at the workshop. It was a very difficult task to select only 43 papers because the level of academic quality of submitted papers was very high, but we had to draw a line in order to avoid parallel sessions. Interestingly, there were almost equal numbers of theoretical papers and technical papers submitted and more or less equal numbers were finally selected. There are also 15 demonstrations, and a special session on Primary School English Language Education in Asia is scheduled on the 2nd day.

We are very lucky to be financially supported by a few institutions, namely NICT, SCAT, LASS and CCDL of Waseda University, and Waseda University International Office, and we are grateful to all of them for supporting the workshop. We would also like to thank all reviewers who had to read and assess the many papers in a short period of time.

I hope that all of you will enjoy the workshop.

Mariko Kondo
Local Organizer, L2WS'2010

Welcome to the L2WS'2010! We're very happy to see more than a hundred participants from all over the world and from many fields related to second language learning and teaching. Following SLaTE'2007 and SLaTE'2009, this workshop is the third ISCA-supported workshop on Speech and Language Technology in Education since 2007. This time, the workshop is co-organized by plural associations of AESOP, SLaTE, NICT, and LASS and this co-organization has added a new flavor to the workshop. How to support teachers and students in classrooms? A good answer to this question can be found only by good collaboration among teachers, researchers, and engineers. We hope that this workshop will be a place where different cultures meet and inspire each other.

In Japan, English education starts at every public primary school next year. In China, Korea, Taiwan, it was already introduced to primary schools some years ago. In my case, I started to learn English by reading a textbook when I was a middle school student. But primary school children today encounter the new language mainly by hearing and speaking. We organized a special session of Primary School English Language Education in Asia. We hope that Japanese teachers can learn many things from our neighbors and that speech scientists and engineers can find what kind of scientific and technical support is still needed for Asian children to learn English as *spoken* language.

You will see many interesting demonstrations. Teachers can learn what is possible and what is impossible by using the current speech technology and engineers can learn what kind of technologies are really needed in language learning environments. In the workshop, you will find many technical and scientific papers. In them, I found some new keywords, one of which is RALL (Robot-aided Language Learning). How can robots, embodied agents, help teachers and students? You will also find other new technical, scientific, and/or educational ideas. Why are these new ideas needed? For true understanding, you may have to learn a new language of a different culture. Give it a try!

Nobuaki Minematsu,
Local Organizer, L2WS'2010

## Local Organizing Committee

| | |
|---|---|
| Sylvain Detey | LASS & Waseda University |
| Tetsuo Harada | LASS & Waseda University |
| Miki Ikoma | LASS & Waseda University |
| Hiroaki Kato | NICT |
| Mariko Kondo | AESOP, LASS & Waseda University |
| Nobuaki Minematsu | SLaTE & The University of Tokyo |
| Michiko Nakano | AESOP, LASS & Waseda University |
| Hajime Tsubaki | LASS & Waseda University |

The organizers would like to thank Maxine Eskenazi (Chair of SLaTE) and Martin Russell (Organizer of SLaTE'2009) for giving us instructive advice to organize this workshop.
The organizers also appreciate Aki Kunikoshi's contribution to design a beautiful leaflet for the symposium and the front page of this book.

## Scientific Review Committee

The organizers would like to thank the following individuals who took part in the review of papers submitted to L2WS'2010.

| | | |
|---|---|---|
| Gregory Aist | Reiko Akahane-Yamada | Tomoyasu Akiyama |
| Abeer Alwan | Shigeaki Amano | Makiko Aoyagi |
| Anton Batliner | Kay Berkling | Jared Bernstein |
| Louis ten Bosch | Lei Chen | Jean-Pierre Chevrot |
| Martin Cooke | Catia Cucchiarini | Rodolfo Delmonte |
| Sylvain Detey | Ryan Downey | Donna Erickson |
| Maxine Eskenazi | Horacio Franco | María Luisa García Lecumberri |
| Björn Granström | Susan Guion | Carlos Gussenhoven |
| Tetsuo Harada | Mark Hasegawa-Johnson | Valerie Hazan |
| Yukari Hirata | Julia Hirschberg | Guoping Hu |
| Kaori Idemaru | Miki Ikoma | Akinori Ito |
| W Lewis Johnson | Hiroaki Kato | Hideaki Kikuchi |
| Mariko Kondo | Nuno Mamede | Dominic Massaro |
| Helen Meng | Hansjörg Mixdorff | Haruko Miyakoda |
| Hirohide Mori | Sachiho Mori | Satomi Mori |
| Kevin G. Munhall | Michiko Nakano | Jean-Luc Nespoulous |
| Mari Ostendorf | Takashi Otake | Kazuharu Owada |
| Masanori Oya | Kyung-Ja Park | Martin Russell |
| Stephanie Seneff | Shigeko Shinohara | Helmer Strik |
| Kazuyo Tanaka | Joseph Tepperman | Isabel Trancoso |
| Chiu-yu Tseng | Yasushi Tsubota | Chiharu Tsurutani |
| Tanya Visceglia | Kimiko Yamakawa | Kiyoko Yoneyama |

# Technical and Social Program at a Glance

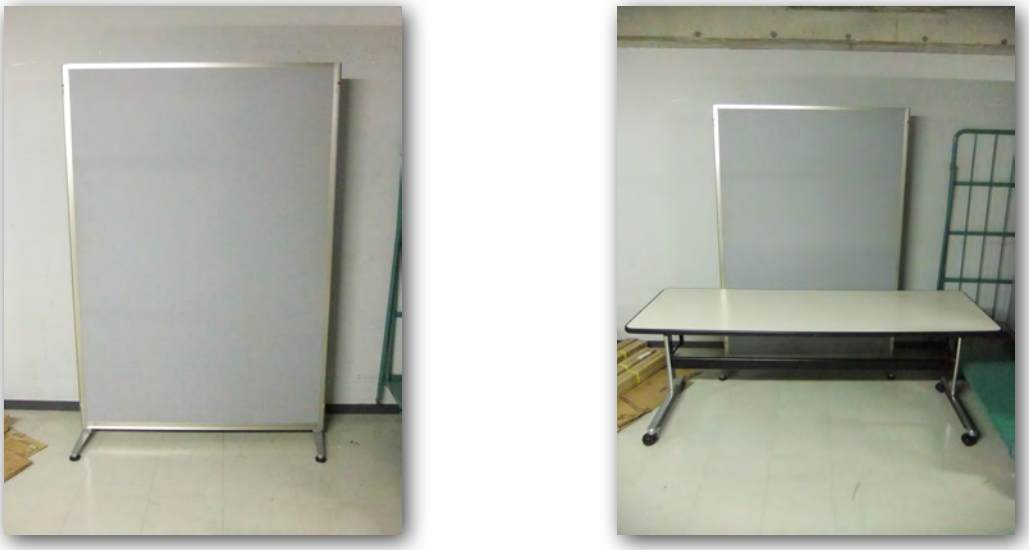
| Wednesday 22nd | Thursday 23rd | Friday 24th |
|---|---|---|
| 08:30 Registration | 08:30 Registration | 08:30 Registration |
| 09:20 Opening<br>*Welcome to L2WS'2010* | 08:50 Oral session 3<br>*Production of a second language* | 08:50 Oral session 4<br>*Prosodic training and corrective feedback* |
| 09:40 Oral Session 1<br>*Perception of a second language* | 10:40 Coffee break | 11:05 Coffee break |
| | 11:00 Demo session 1<br>*New technologies and methodologies help LL.* | 11:20 Demo session 2<br>*New technologies and methodologies help LL.* |
| 11:30 Lunch break | 12:40 Lunch break | 13:00 Lunch break |
| 13:00 Poster session 1<br>*Teaching and learning environment* | 14:00 Poster session 2<br>*Science and technology of speech and language for education* | 14:00 Closing |
| 14:40 Coffee break | 15:40 Coffee break | |
| 15:00 Oral session 2<br>*Automatic pronunciation assessment* | 16:00 Special session<br>*Primary school English education in Asia* | |
| 18:30 Reception | 18:30 Banquet | 15:00 Tour to Sugamo |

# Guidelines for Presentation

**Instructions for oral presentation:**
Each oral session starts with the session chair's overview remark on the papers of that session. Each presentation slot is limited to 25 min, 20 min for presentation and 5 min for Q&A and change of speakers. You're supposed to present your work by connecting your own PC or Mac to the projector pre-installed in the hall. Test and confirm that your machine can be properly connected to the projector no later than the starting time of the session.

**Instructions for poster presentation:**
A poster panel of 114 cm wide and 168 cm high is provided for each presentation. Note that the panel is in a portrait format. Each poster session is held *after* a lunch break and it is 100 min long. You must put up your poster in the lunch break. Materials to fix your poster on the panel will be provided in the session room. You must remove your poster by the end of the entire sessions of that day. If you need a small table, a power supply, etc, please make a contact with the committee. Wireless internet connection is available in the room.

**Instructions for demonstration and exhibition:**
A poster panel of 114 cm wide and 168 cm high and a table is provided for each presentation. Note that the panel is in a portrait format. Wireless internet connection is available in the session room. Each demo session is held *before* a lunch break and it is 100 min long. You must set up your demo system and poster *before* the morning session or during a break before the demo session. Materials to fix your poster on the panel will be provided in the room. You must remove your demo system and poster during the lunch break so that a poster session will start smoothly. If you need a chair, a monitor, loud speakers, etc, please make a contact with the committee.

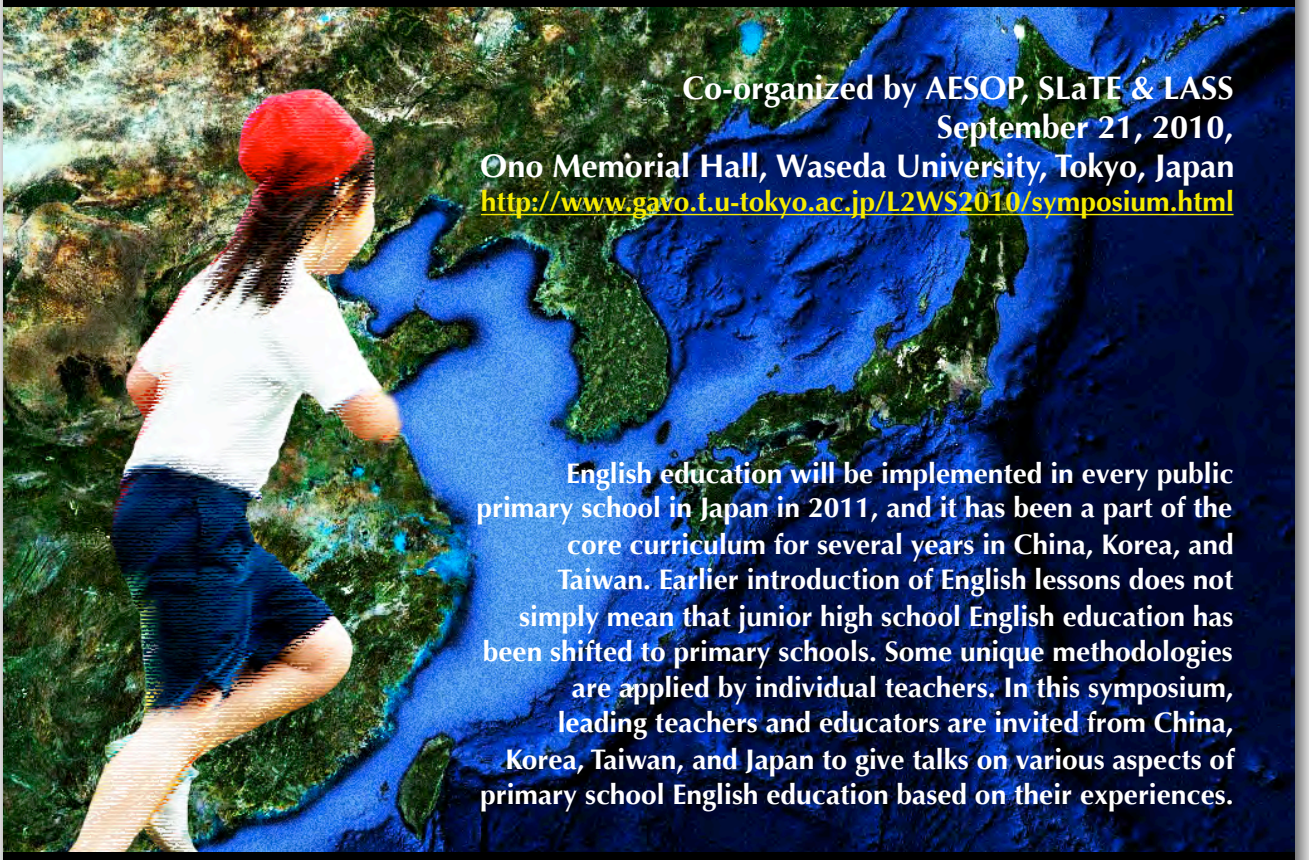


A poster panel provided for each poster presentation (left) and
a poster panel and a table provided for each demonstration (right)

## Open Symposium on "Primary School English Education in Asia"

An open symposium on

# "Primary School English Education in Asia"



**Co-organized by AESOP, SLaTE & LASS**
**September 21, 2010,**
**Ono Memorial Hall, Waseda University, Tokyo, Japan**
**http://www.gavo.t.u-tokyo.ac.jp/L2WS2010/symposium.html**

**English education will be implemented in every public primary school in Japan in 2011, and it has been a part of the core curriculum for several years in China, Korea, and Taiwan. Earlier introduction of English lessons does not simply mean that junior high school English education has been shifted to primary schools. Some unique methodologies are applied by individual teachers. In this symposium, leading teachers and educators are invited from China, Korea, Taiwan, and Japan to give talks on various aspects of primary school English education based on their experiences.**

### Program

12:00 Ono Memorial Hall opens.
13:00 **"Primary English Curriculum Reform in Beijing"**
Prof. Zehang Chen (Beijing Normal University, China)
14:00 **"Primary ELT in Korea: Start, Taxi, Take-off and Fly"**
Prof. WonKey Lee (Seoul National University of Education, Korea)
15:20 **"An Overview of English Language Education at Primary Level in Taiwan"**
Prof. Chiou-lan Chern (National Taiwan Normal University, Taiwan)
16:20 **"Elementary School English Education in Japan"**
Prof. Kyoko Kasuya (Tokyo Gakugei University, Japan)
Prof. Yuri Kuno (Chubu Gakuin University, Japan)
18:00 Ono Memorial Hall closes.

### Admission

It is free.

### Registration

Please notify your attendance to us by email (L2WS-org@list.waseda.jp). Note that the registration will be closed without notice when the number of the attendance reaches the capacity of the hall.

## Nearby Places of Interest

**Sugamo (巣鴨):** It is an area in the center of Tokyo that has a different atmosphere to most of the city. It has kept a feeling of old traditional Tokyo. It developed as a commercial center in the middle of the Edo era, on one of the five main roads linking Edo (old name of Tokyo) with the rest of Japan. Sugamo was the first stop on Nakasendo Road after departing from Nihonbashi near Ginza. It is also famous for Sugamo Togenuki Jizo temple. Jizo was a disciple of Buddha who is considered as a protector of children. In Sugamo there is a traditional shopping street leading to and from the Jizo temple with shops selling traditional sweets, snacks, dried seaweeds, dried fish, curio and fashion shops catering for old ladies. It is a very popular place with old people and is always mobbed with visitors young and old. There are special market days on the 4th, 14th and 24th of each month with many stalls which sell second-hand kimonos, potteries, bits and pieces etc.

**Kansenen Park (甘泉園公園):** There is a public park with a traditional Japanese garden called "Kansen-en" just a 3-4 minute walk from the workshop venue. "Kansen" means "sweet spring" and the name came from the spring water in the center of the park that was suitable for making Japanese green tea. It used to be much larger and was a garden of the Shimizu family who worked for the Tokugawa shogun. Waseda University acquired the garden in 1938 but later transferred the ownership to Shinjuku-Ward of Tokyo.
Opening hours: 7:00am - 7:00pm (free)

**Shin-Edogawabashi Park (新江戸川橋公園):** It is only a few minutes walk from the workshop venue, situated just beside Chinzanso where we will have the workshop dinner. It used to be one of houses of the Hosokawa family who had been an influential samurai family since the 13th century. The garden shows a dynamic landscape and it is famous for beautiful autumn leaves.
Opening hours: 9:00am - 5:00pm (last admittion: 4:30pm, free)

**Kandagawa River (神田川):** It is 24.6 km long starting in the west of Tokyo. It runs through the center of the city and then joins Sumida River and eventually flows into Tokyo Bay. The river around Waseda University is quite pretty and is famous for cherry blossoms in spring; the river is lined with about 2.5 km of cherry trees between Waseda and Edogawabashi - one of the best cherry blossom spots in Tokyo. The water around Waseda is amazingly clean and you can see lots of carp, tortoises, and many birds including herons and storks. It is a very good strolling and jogging spot.

Kansenen Park in early summer

Shin-Edogawabashi Park in late spring

Kandagawa River in spring

# Technical Program of Day 1

**Registration** **[8:30-9:20]**

**Opening** **[9:20-9:40]**

**Oral Session 1: Perception of a Second Language** **[9:40-11:30]**
**chaired by Chiu-yu Tseng**

Overview of the papers of this session (10min)

O1-1: Second-Language Experience and Speech-in-Noise Recognition: the Role of L2 Experience in the Talker-Listener Accent Interaction
*Melanie Pinet, Paul Iverson, Mark Huckvale*

O1-2: Inter- and Intra-L1 Differences in L2 Speech Perception
*Jeffrey J. Holliday*

O1-3: Categorizing Mandarin Tones into Japanese Pitch-Accent Categories: The Role of Phonetic Properties
*Connie K. So*

O1-4: A Pilot Study on Perception of Spanish Stress by Japanese Learners of Spanish
*Takuya Kimura, Hirotaka Sensui, Miyuki Takasawa, Atsuko Toyomaru, José Joaquín Atria*

**Lunch break** **[11:30-13:00]**

**Poster Session 1: Teaching and Learning Environment** **[13:00-14:40]**

P1-1: A Resource for Learning Swedish Oral Skills
*Linda Ösp Heimisdóttir, Cecilia Ovesdotter Alm, Ian Alden Coots, Kateri Krantz-Odendahl*

P1-2: Bridging a Gap between L2 Research and Classroom Practice (1): English as a Lingua Franca (ELF) in Asia and Some Assessment Based on Common European Framework of Reference for Languages(CEFR)
*Michiko Nakano, Eiichiro Tsutsui, Yusuke Kondo*

P1-3: Bridging the Gap between L2 Research and Classroom Practice (2): Evaluation of Automatic Scoring System for L2 Speech
*Yusuke Kondo, Eiichiro Tsutsui, Michiko Nakano*

P1-4: Bridging the Gap between L2 Research and Classroom Practice (3) -- Online Assessment and Practical Teaching
*Eiichiro Tsutsui, Yusuke Kondo, Michiko Nakano*

P1-5: Laying the Groundwork for Ongoing Learning: A Scaffolded Approach to Language Education in Japanese Elementary Schools and Beyond
*Francesco Bolstad, Toshiyuki Kanamaru, Akira Tajino*

P1-6: Form-Focused Task-Oriented Dialogues for Computer Assisted Language Learning: A Pilot Study on German Dative
*Magdalena Wolska, Sabrina Wilske*

P1-7: Development of a System to Assist Simultaneous Interpretation and Shadowing
*Michiko Watanabe, Youichi Tokioka, Keikichi Hirose*

P1-8: ARTICULA - A Tool for Spanish Vowel Training in Real Time
*William R. Rodríguez, Oscar Saz, Eduardo Lleida*

P1-9: A Multilingual Platform for Building Speech-Enabled Language Courses
*Manny Rayner, Pierrette Bouillon, Nikos Tsourakis, Johanna Gerlach, Claudia Baur, Maria Georgescul, Yukie Nakao*

P1-10: Visual Articulatory Feedback for Phonetic Correction in Second Language Learning
*Pierre Badin, Atef Ben Youssef, Gérard Bailly, Frédéric Elisei, Thomas Hueber*

P1-11: Cognitive Effects of Robot-Assisted Language Learning on Oral Skills
*Sungjin Lee, Hyungjong Noh, Jonghoon Lee, Kyusong Lee, Gary Geunbae Lee*

**Coffee break** **[14:40-15:00]**

**Oral Session 2: Automatic Pronunciation Assessment** **[15:00-17:15]**
**chaired by Björn Granström**

Overview of the papers of this session (10min)

O2-1: Automatic Fluency Assessment by Signal-Level Measurement of Spontaneous Speech
*Suma Bhat, Mark Hasegawa-Johnson, Richard Sproat*

O2-2: A Simple Feature Normalization Scheme for Non-Native Vowel Assessment
*Mitchell Peabody, Stephanie Seneff*

O2-3: Pronunciation Proficiency Estimation Based on Multilayer Regression Analysis Using Speaker-Independent Structural Features
*Masayuki Suzuki, Yu Qiao, Nobuaki Minematsu, Keikichi Hirose*

O2-4: Automatic Pronunciation Error Detection in Repetitor
*Eric Sanders, Henk van den Heuvel*

O2-5: How Many Labellers? Modelling Inter-Labeller Agreement and System Performance for the Automatic Assessment of Non-Native Prosody
*Florian Hönig, Anton Batliner, Karl Weilhammer, Elmar Nöth*

**Welcome Reception** **[18:30-]**

# Technical Program of Day 2

**Registration** **[8:30-8:50]**

**Oral Session 3: Production of a Second Language** **[8:50-10:40]**
**chaired by Valerie Hazan**

Overview of the papers of this session (10min)

O3-1: Interaction of Lexical and Sentence Prosody in Taiwan L2 English
*Tanya Visceglia, Chiu-yu Tseng, Zhao-yu Su, Chi-Feng Huang*

O3-2: German Learners of Japanese - Perceptual and Prosodic Analysis of Utterances from a Meditative Setting
*Hansjörg Mixdorff, Ryoko Hayashi, Yoriko Yamada-Bochynek, Keikichi Hirose, Hiroya Fujisaki*

O3-3: A Cross-Language Study of Compensatory Response to Formant-Shifted Feedback
*Takashi Mitsuya, Ewen N. MacDonald, David W. Purcell, Kevin G. Munhall*

O3-4: Phoneme Errors in Read and Spontaneous Non-Native Speech: Relevance for CAPT System Development
*Joost van Doremalen, Catia Cucchiarini, Helmer Strik*

**Coffee break** **[10:40-11:00]**

**Demo Session 1: New Technologies and Methodologies Help Language Learning.** **[11:00-12:40]**

D1-1: A New Force in the Interpretation Technology
*Yoichi Tokioka, Kazuhiro Okuda*

D1-2: CALL Systems at Kyoto University
*Yasushi Tsubota, Tatsuya Kawahara, Masatake Dantsuji*

D1-3: CALL-SLT/Web, A Speech-Enabled Translation Game on the Internet
*Manny Rayner, Nikos Tsourakis, Pierrette Bouillon, Matthew Fuchs*

D1-4: Phonensínthesis: A System for Teaching French Pronunciation to Japanese Students of French
*Josafá de Jesus Aguiar Pontes, Sadaoki Furui*

D1-5: Automatic Assessment of Non-Native Prosody
*Florian Hönig, Anton Batliner*

D1-6: Pronunciation Assessment System Using Structural Features
*Masayuki Suzuki, Nobuaki Minematsu*

D1-7: Dialect (Regional Accent)-Based Speaker Classification of Chinese
*Xuebin Ma, Nobuaki Minematsu*

D1-8: JaFIX (Japanese as a Foreign Language with Integrative-Communicative Steps) -- JFL-Acquisiton through Relax and Gesticalization
*Yoriko Yamada-Bochynek*

**Lunch break (SLaTE meeting)** **[12:40-14:00]**

**Poster Session 2: Science and Technology of Speech and Language for Education** **[14:00-15:40]**

P2-1: A Study of Pitch Patterns of Sentence Utterances by Japanese Speakers of English in Comparison with Native Speakers of English
*Tomoko Nariai, Kazuyo Tanaka*

P2-2: Towards a Computer-Aided Pronunciation Training System for German Learners of Mandarin - Prosodic Analysis
*Hussein Hussein, Hansjörg Mixdorff, Hue San Do, Si Wei, Shu Gong, Hongwei Ding, Qianyong Gao, Guoping Hu*

P2-3: Effects of Pitch Cues on the Identification of Vowel Length in L2 Japanese
*Izumi Takiguchi*

P2-4: Examination of the Relationship between L2 Perception and Production: An Investigation of English /r/-/l/ Perception and Production by Adult Japanese Speakers
*Kota Hattori, Paul Iverson*

P2-5: Consonant Cluster Production in Japanese Learners of English
*Yoshiho Shibuya, Donna Erickson*
P2-6: Can We Predict Who Will Benefit from Computer-Based Phonetic Training?
*Valerie Hazan, Yoon Hyun Kim*
P2-7: Attention to Critical Acoustic Features for L2 Phonemic Identification and Its Implication on L2 Perceptual Training
*Yoon Hyun Kim, Jung-Oh Kim*
P2-8: Speech Analysis for Automatic Evaluation of Shadowing
*Dean Luo, Yutaka Yamauchi, Nobuaki Minematsu*
P2-9: Synthesizing Expressive Speech to Convey Focus Using a Perturbation Model for Computer-Aided Pronunciation Training
*Fanbo Meng, Helen Meng, Zhiyong Wu, Lianhong Cai*
P2-10: Multimodal Learning of Words: A Study on the Use of Speech Synthesis to Reinforce Written Text in L2 Language Learning
*Kevin Dela Rosa, Gabriel Parent, Maxine Eskenazi*
P2-11: Automatic Generation of Cloze Question Distractors
*Rui Correia, Jorge Baptista, Nuno Mamede, Isabel Trancoso, Maxine Eskenazi*
P2-12: Automatic Selection of Collocations for Instruction
*Adam Skory, Maxine Eskenazi*
P2-13: Toward a Chanting Robot for Interactively Teaching English to Children
*Ryo Nagata, Tomoya Mizumoto, Kotaro Funakoshi, Mikio Nakano*
P2-14: Applications of the Buckeye GTA Corpus for L2 Teaching and Research
*Jocelyn B. Hardman, Elizabeth McCullough*

**Coffee break** **[15:40-16:00]**

**Special Session: Primary School English Education in Asia** **[16:00-17:40]**
**chaired by Nobuaki Minematsu**

S-1: Primary English Curriculum Reform in Beijing
*Zehang Chen, Lingdi Shen*
S-2: Primary ELT in Korea: Start, Taxi, Take-off and Fly
*WonKey Lee*
S-3: An Overview of English Language Education at Primary Level in Taiwan
*Chiou-lan Chern*
S-4: Elementary School English Education in Japan -- Its History and the Sound of Its Teaching Materials --
*Kyoko Kasuya, Yuri Kuno*

**Banquet** **[18:30-]**

# Technical Program of Day 3

**Registration** **[8:30-8:50]**

**Oral Session 4: Prosodic Training and Corrective Feedback** **[8:50-11:05]**
**chaired by Maxine Eskenazi**

Overview of the papers of this session (10min)

O4-1: MusicSpeak: Capitalizing on Musical Rhythm for Prosodic Training in Computer-Aided Language Learning
*Hao Wang, Peggy Mok, Helen Meng*

O4-2: Lexical Tones Learning with Automatic Music Composition System Considering Prosody of Mandarin Chinese
*Siwei Qin, Satoru Fukayama, Takuya Nishimoto, Shigeki Sagayama*

O4-3: Practicing Syntax in Spoken Interaction: Automatic Detection of Syntactical Errors in Non-Native Utterances
*Helmer Strik, Janneke van de Loo, Joost van Doremalen, Catia Cucchiarini*

O4-4: Simicry - A Mimicry-Feedback Loop for Second Language Learning
*Preben Wik, Björn Granström*

O4-5: The Role of Corrective Feedback in Second Language Learning: New Research Possibilities by Combining CALL and Speech Technology
*Bart Penning de Vries, Catia Cucchiarini, Helmer Strik, Roeland van Hout*

**Coffee break** **[11:05-11:20]**

**Demo Session 2: New Technologies and Methodologies Help Language Learning.** **[11:20-13:00]**

D2-1: Changyan Interactive English Learning System
*Guoping Hu*

D2-2: A Multi-Player Vocabulary Game that Teaches While It Learns
*Adam Skory, Maxine Eskenazi*

D2-3: NTU Chinese -- A Chinese Language Pronunciation Learning Software
*Yow-Bang Wang, Lin-Shan Lee*

D2-4: REAP.PT
*Rui Pedro dos Santos Correia*

D2-5: Development of an Automatic Evaluation System of ESL/EFL Learners' Skills of Shadowing
*Dean Luo, Yutaka Yamauchi, Nobuaki Minematsu*

D2-6: Automatic Tests of Spoken Spanish, Arabic, and Chinese; and 4-Skills Testing in English
*Jared Bernstein, Jian Cheng, Elizabeth Rosenfeld*

D2-7: A Vowel Training System for All
*Toshiko Isei-Jaakkola, Takatoshi Naka, Keikichi Hirose*

**Lunch break** **[13:00-14:00]**

**Closing** **[14:00-15:00]**

**Tour to Sugamo** **[15:00-]**

# Exhibitions

**E-1: CHIeru Digital Language Lab System CaLabo EX**
***CHIeru Co., Ltd.***
Chieru is dedicated to the whole integration of technology, content and communication for ICT in Education to build schools for the future. In this workshop, we are introducing CaLabo EX, which is the most popular Computer-Assisted Language Lab system brand in Japan. With CaLabo, institutions can create a computer-assisted language learning environment that includes high quality audio, video and multimedia educational resources.

**E-2: The English Learning Materials You Can Learn Anytime and Anywhere**
**-- English Learning for Entrance Examinations with Nintendo DS Machines --**
***Educational Network Inc.***
We are a publisher of educational materials for prep schools called "juku" and private junior and high schools. In this time, we are introducing our Nintendo DS software for entrance examinations to the high school and university. You can learn English anytime and anywhere you want.

**E-3: A New Method of Second Language Studies Linked to Simultaneous Interpretation and Translation**
***Powershift and Straightword, Inc.***
1. "is" (A system to assist simultaneous interpretation and shadowing.)
2. "Nowgaku" (An English learning system)
3. "Interlingua" (Translation system on twitter).
4. "Realscript" (Simultaneous Interpretation on twitter)
5 "Eigo Onkan Lesson" (Written by Yoichi Tokioka / Published by Asahi Press 2010. 09.22)

**E-4: Application of Pronunciation Recognition to CALL -- Hatsuon-ryoku (Power of Pron.) --**
***Prontest Inc.***
We developed the technology to judge the condition of the phonic-organ automatically using the precise phonetic-analysis technique. It is possible to guide specifically in the level of the phonic-organ and also possible to systematize knowledge having to do with a phonetics and the know-how of the pronunciation guide directly.
Very effective teaching method of pronunciation. It establishes a road to build good teaching materials.

# Abstracts

## Oral Session 1: Perception of a Second Language

**O1-1: Second-Language Experience and Speech-in-Noise Recognition: the Role of L2 Experience in the Talker-Listener Accent Interaction**

*Melanie Pinet, Paul Iverson, Mark Huckvale,*
*Department of Speech, Hearing and Phonetic Sciences, University College London, UK*

This study investigated how L2 experience modulates L1-L2 talker-listener intelligibility. L1 southern British English (SE) and L1 French listeners with varying L2 experience (Inexperienced 'FI', Experienced 'FE' and Bilinguals) were tested on their speech-in-noise recognition of English sentences that were spoken with a range of accents (SE, FI, FE, Northern Irish and Korean-accented English). Results showed that while the FI listeners had graded sensitivity for the accents, the SE listeners' recognition processes were selectively tuned to their own accent. Overall, this suggests that L2 experience affects talker-listener accent interactions, altering both accent intelligibility and selectivity of accent processing.

**O1-2: Inter- and Intra-L1 Differences in L2 Speech Perception**

*Jeffrey J. Holliday,*
*Department of Linguistics, The Ohio State University, USA*

In a perception experiment, L1 Mandarin and L1 Japanese novice learners of Korean classified non-tense /s/- or tense /s*/- initial Korean CV tokens. A mixed effects logistic regression model with acoustic cues as predictor variables was built for each L1 group, and each individual's regression coefficients were interpreted to be the cue weighting used in identifying Korean /s/ and /s*/. We propose that the weighting of L2 perceptual cues is influenced by the weighting of the same cues in the L1 perception of acousticaly similar contrasts, but that intra-L1 individual variation is great enough that the expected inter-L1 differences may appear less well defined.

**O1-3: Categorizing Mandarin Tones into Japanese Pitch-Accent Categories: The Role of Phonetic Properties**

*Connie K. So,*
*MARCS Auditory Laboratories, University of Western Sydney, Australia*

This study examined how native Japanese speakers, who were naïve to Mandarin, categorized Mandarin tones (in citation form) into their native pitch–accent categories. Results showed that Japanese listeners categorized the nonnative Mandarin tones into their native pitch accent categories, in ways that were consistent with the phonetic features of listeners' native language. The findings support the new assumption of PAM for suprasegmentals [14] that non-native prosodic categories (e.g., lexical tones) will be assimilated to the categories of listeners' native prosodic system.

**O1-4: A Pilot Study on Perception of Spanish Stress by Japanese Learners of Spanish**

*Takuya Kimura[1], Hirotaka Sensui[2], Miyuki Takasawa[3], Atsuko Toyomaru[4], José Joaquín Atria[5]*
*[1]Department of Spanish Language and Literature, Seisen University, Japan*
*[2]Department of Spanish and Latin American Studies, Nanzan University, Japan*
*[3]School of Education, Waseda University, Japan*
*[4]Faculty of Political Science and Economics, Takushoku University, Japan*
*[5]Faculty of Foreign Studies, Tokyo University of Foreign Studies, Japan*

Japanese learners of Spanish sometimes fail to perceive the stresses when listening to Spanish utterances. Results of a perceptual experiment with 270 stimuli and 64 informants (43 Spanish and 21 Japanese) reveal that Spanish speakers perceive the stresses correctly almost every time, whereas Japanese speakers tend to fail to do so when the word is pronounced with rising intonation. The cause of this is the difference in phonetic realizations of Spanish stresses and those of Japanese accents. Japanese learners should be taught that the Spanish stressed with a high pitch.

## Poster Session 1: Teaching and Learning Environment

**P1-1: A Resource for Learning Swedish Oral Skills**

*Linda Ösp Heimisdóttir[1], Cecilia Ovesdotter Alm[2], Ian Alden Coots[3], Kateri Krantz-Odendahl[3]*
*[1]Department of Linguistics, Cornell University, USA*
*[2]Department of German Studies, Cornell University, USA*
*[3]Cornell University, USA*

This paper discusses an interactive CALL (Computer-Assisted Language Learning) resource for the learning of Swedish oral skills. We present a linguistically sound model of oral skills divided into four levels and a sharable Swedish exercise database that reflects this model. We also discuss corresponding Swedish speech data that capture several dimensions of sociolinguistic speech variations (e.g., gender, native/heritage native/non-native speakership, and region). We also report on an evaluation of this new resource using a methodology that considers learning gain in terms of oral perception and listening skills, as well as four categories of user satisfaction. The paper concludes with thoughts for future directions of CALL technologies.

**P1-2: Bridging a Gap between L2 Research and Classroom Practice (1): English as a Lingua Franca (ELF) in Asia and Some Assessment Based on Common European Framework of Reference for Languages (CEFR)**

*Michiko Nakano[1], Eiichiro Tsutsu[2], Yusuke Kondo[3]*
*[1]Faculty of Education and Integrated Arts and Sciences, Waseda University, Japan*
*[2]International Center, Hiroshima International University, Japan*
*[3]Language Education Center, Ritsumeikan University, Japan*

Since English has become a lingua franca in the world, English programs need to be based on International standards such as Common European Framework of Reference (CEFR). (1) We report validation experiment by can-do self-check survey in relation to the CEFR levels, in order to find out the cut-off scores and ranges of our computer adaptive placement test called Web-based Test of English Communication (WeTEC), using logistic regression analysis. (2) We discuss Oral self-introduction data among Asian users of English based on the CEFR, using the multi-faceted Rasch Model.

**P1-3: Bridging the Gap between L2 Research and Classroom Practice (2): Evaluation of Automatic Scoring System for L2 Speech**

*Yusuke Kondo[1], Eiichiro Tsutsui[2], Michiko Nakano[3]*
*[1]Language Education Center, Ritsumeikan University, Japan*
*[2]International Center, Hiroshima International University, Japan*
*[3]Faculty of Education and Integrated Arts and Sciences, Waseda University, Japan*

This paper introduces the construction, the implementation, and the evaluation of an automated scoring system for read-aloud speech of L2 learners'. In this system, evaluation scores given by trained human raters are predicted, based on the speech characteristics of learners in read-aloud speech.

**P1-4: Bridging the Gap between L2 Research and Classroom Practice (3) -- Online Assessment and Practical Teaching**

*Eiichiro Tsutsui[1], Yusuke Kondo[2], Michiko Nakano[3]*
*[1]International Center, Hiroshima International University, Japan*
*[2]Language Education Center, Ritsumeikan University, Japan*
*[3]Faculty of Education and Integrated Arts and Sciences, Waseda University, Japan*

Recent technological advancements have been changing our educational environments. Because of learning management systems and online communication tools, educators can create and tailor virtual learning environments relatively easily, and they can incorporate new-dimensional approaches into their practical teaching. The inevitable consequence is that L2 learners need to get more accustomed to new types of learning environments. Not only should they know the effective use of computers for their learning, but they have to know how to become more responsible learners. This is because they have more chances to learn English independently in e-learning environments. Many educators emphasize the importance of student-centered and student-oriented learning. However, there are hardly any educational projects or methods that can facilitate the transition from teacher-centered to student-centered approaches. Against this background, the aim of this study is two-fold. First, we will show how we support a new generation of language learners that should become independent learners of English. Secondly, we will present survey results on L2 learners' impressions of our supporting methods.

**P1-5: Laying the Groundwork for Ongoing Learning: A Scaffolded Approach to Language Education in Japanese Elementary Schools and Beyond**

*Francesco Bolstad[1], Toshiyuki Kanamaru[1], Akira Tajino[2]*
*[1]Graduate school of Human and Environmental Studies, Kyoto University, Japan*
*[2]Center for the Promotion of Excellence in Higher Education, Kyoto University, Japan*

This paper is based on a twofold argument. Firstly, it argues for the need to view English language learning as a long-term undertaking and to link elementary school English programs more strongly with junior and senior high school programs in order to support students' success across the term of their lives as English language learners. Secondly, the results of a classroom investigation into the feasibility of using meaning-based word order (IMIJUN) as a framework for achieving this greater congruity of programs by scaffolding Japanese elementary school students learning as they are challenged to go beyond memorizing English to creating their own English sentences is reported.

**P1-6: Form-Focused Task-Oriented Dialogues for Computer Assisted Language Learning: A Pilot Study on German Dative**

*Magdalena Wolska, Sabrina Wilske*
*Computational Linguistics, Saarland University, Saarbrücken, Germany*

We report on a pilot experiment conducted in order to investigate whether computer-based conversational focused tasks promote acquisition of forms. The structure we targeted was the German dative case in prepositional phrases. The goal of the task we designed was two-fold: First, learners should improve their overall communicative skills in the scenario and, second, expand their mastery of the target structure. In this paper, we present an evaluation of learners' progress on the latter.

**P1-7: Development of a System to Assist Simultaneous Interpretation and Shadowing**

*Michiko Watanabe[1], Youichi Tokioka[2], Keikichi Hirose[3]*
*[1]Graduate School of Frontier Sciences, The University of Tokyo, Japan*
*[2]Straightword, Inc., Japan*
*[3]Graduate School of Information Science and Technology, The University of Tokyo, Japan*

We have been developing a system to assist simultaneous interpretation on site by introducing merits of consecutive interpretation. While recording a speaker's voice, it can replay the voice at various speeds. Interpreters can stop and restart the replay at any given point in time so that they can concentrate on giving interpretation when necessary. This approach will free interpreters from simultaneously engaging in multiple tasks, such as listening to the original speech, comprehending it, translating it into another language and speaking out. We conducted experiments to examine the effects of the system on English-Japanese interpretation using four interpreters. With the system, the accuracy and the fluency of interpretations improved. However, it took interpreters more than twice as long as the interpretations in traditional methods. We report our efforts to reduce the time delay and discuss applying the system to shadowing.

**P1-8: ARTICULA - A Tool for Spanish Vowel Training in Real Time**

*William R. Rodríguez, Oscar Saz, Eduardo Lleida*
*Communications Technology Group (GTC), Aragon Institute for Engineering Research (I3A), University of Zaragoza, Spain*

This paper describes a free tool for the training of Spanish vowels called ARTICULA. This tool shows an accurate approximation in real time to the position and movements of the tongue, jaw and lips during vowel utterances independent of user's characteristics as age and gender. At the same time, the tool displays acoustic parameters as intensity, pitch, formants and spectrum, thus making ARTICULA a good alternative for vocalic articulation in voice therapy and training in Spanish language studies. The tool uses a formant normalization through the vocal tract length in order to reduce the high variability between speakers. A preliminary study in voice therapy in children with voice disorders shows the adequate biofeedback provided by the tool and, the improvement in specific vowels after ten weeks of therapy.

**P1-9: A Multilingual Platform for Building Speech-Enabled Language Courses**

*Manny Rayner[1], Pierrette Bouillon[1], Nikos Tsourakis[1], Johanna Gerlach[1], Claudia Baur[1], Maria Georgescul[1], Yukie Nakao[2]*
*University of Geneva, TIM/ISSCO, Switzerland*
*LINA, Nantes University, France*

We present CALL-SLT, a generic multilingual speech-enabled Open Source CALL system based on the "translation game" idea of Wang and Seneff, focussing on recent enhancements which allow the instructor to define a structured language course divided up into a set of lessons. Each lesson picks out a subset of the corpus using a combination of semantic and syntactic constraints. We describe how the "structured lesson" framework interacts with the spoken help facilities offered by the system, and outline the initial sets of lessons we have constructed for the English, French and Japanese versions of CALL-SLT.

**P1-10: Visual Articulatory Feedback for Phonetic Correction in Second Language Learning**

*Pierre Badin, Atef Ben Youssef, Gérard Bailly, Frédéric Elisei, Thomas Hueber*
*GIPSA-lab (Département Parole & Cognition / ICP), UMR 5216 CNRS – Grenoble University, France*

Orofacial clones can display speech articulation in an augmented mode, i.e. display all major speech articulators, including those usually hidden such as the tongue or the velum. Besides, a number of studies tend to show that the visual articulatory feedback provided by ElectroPalatoGraphy or ultrasound echography is useful for speech therapy. This paper describes the latest developments in acoustic-to-articulatory inversion, based on statistical models, to drive orofacial clones from speech sound. It suggests that this technology could provide a more elaborate feedback than previously available, and that it would be useful in the domain of Computer Aided Pronunciation Training.

**P1-11: Cognitive Effects of Robot-Assisted Language Learning on Oral Skills**

*Sungjin Lee, Hyungjong Noh, Jonghoon Lee, Kyusong Lee, Gary Geunbae Lee*
*Department of Computer Science and Engineering,*
*Pohang University of Science and Technology (POSTECH), South Korea*

This study introduces the educational assistant robots that we developed for foreign language learning and explores the effectiveness of robot-assisted language learning (RALL). To achieve this purpose, a course was designed in which students have meaningful interactions with intelligent robots in an immersive environment. A total of 24 elementary students, ranging in age from 9 to 13, were enrolled in English lessons. A pre-test/post-test design was used to investigate the cognitive effects of the RALL approach on the students' oral skills. No significant difference in the listening skill was found, but the speaking skills improved, with a large effect size at the significance level of 0.01.

## Oral Session 2: Automatic Pronunciation Assessment

**O2-1: Automatic Fluency Assessment by Signal-Level Measurement of Spontaneous Speech**

*Suma Bhat[1], Mark Hasegawa-Johnson[1], Richard Sproat[2]*
*[1]Department of ECE, University of Illinois, Urbana-Champaign, USA*
*[2]Center for Spoken Language Understanding, Oregon Health and Science University, USA*

In its narrow sense, the term fluency connotes fluidity of speech. This study is a step in the quest for objective language assessment methods one of which is rating for oral fluency in a second language. In particular, we seek to find what measures obtained from a spontaneous utterance can be used as predictors of fluency and, to assess the utility of a set of acoustic measures obtained by signal-level measurements towards predicting fluency automatically. Experiments done using an ESL data set of spontaneous speech show that articulation rate and phonation-time ratio are good predictors of fluency, in line with earlier findings.

Our contribution is to use signal-level measurements as quantifiers of perceived fluency in a logistic regression framework and to show the existence of an alternate approach to ASR-based fluency assessment, which, owing to unacceptable levels of recognition accuracies, have limited use in real testing scenarios. Our results have implications in developing fluency assessment systems for language-resource scarce settings as well as for a wide variety of testing scenarios.

**O2-2: A Simple Feature Normalization Scheme for Non-Native Vowel Assessment**

*Mitchell Peabody, Stephanie Seneff*
*Spoken Language Systems, CSAIL, MIT, USA*

We introduce a set of speaker dependent features derived from the positions of vowels in Mel-Frequency Cepstral Coefficient (MFCC) space relative to a reference vowel. The MFCCs for a particular speaker are transformed using simple operations into features that can be used to classify vowels from a common reference point. Classification performance of vowels using Gaussian Mixture Models (GMMs) is significantly improved, regardless of which vowel is used as the target among /ɑ/, /i/, /u/, or /ə/. We discuss how this technique can be applied to assess pronunciation with respect to vowel structure rather than agreement with absolute position in MFCC space.

**O2-3: Pronunciation Proficiency Estimation Based on Multilayer Regression Analysis Using Speaker-Independent Structural Features**

*Masayuki Suzuki[1], Yu Qiao[2], Nobuaki Minematsu[1], Keikichi Hirose[1]*
*[1]The University of Tokyo, Japan*
*[2]Shenzhen Institutes of Advanced Technology, China*

Teachers can assess the pronunciations of students independently of extra-linguistic features such as age and gender observed in the students' utterances. This capacity is, however, difficult to realize on machines because linguistic differences and extra-linguistic differences change acoustic features commonly. Therefore, the performance of automatic pronunciation assessment is inevitably affected by the extra-linguistic features. Recently, we proposed acoustic features that are independent of extra-linguistic factors, called structural features and realized a technique for pronunciation proficiency estimation that is extremely robust to these factors. In this paper, we extend this technique with multilayer regression analysis, where supervised learning is done at each layer by using teachers' scores of that layer. Experiments of estimating the proficiency show that higher correlations between teachers and machines are obtained compared to our previous structure-based assessment.

**O2-4: Automatic Pronunciation Error Detection in Repetitor**

*Eric Sanders, Henk van den Heuvel*
*Centre for Language and Speech Technology, Radboud University Nijmegen, the Netherlands*

This paper describes a pronunciation error detection method for Repetitor, a pronunciation training computer program for second language learners of Dutch. A database of L2-speech was constructed and a selection of relevant pronunciation errors for Repetitor was made. Our error detection method is based on a weighted variant selection using forced alignment. Tested on the database, the results show that our method achieves satisfactory detection performance for most pronunciation errors yielding a precision of correct rejects of over 85% for most errors, and scoring accuracies between 85% and 100%.

**O2-5: How Many Labellers? Modelling Inter-Labeller Agreement and System Performance for the Automatic Assessment of Non-Native Prosody**

*Florian Hönig[1], Anton Batliner[1], Karl Weilhammer[2], Elmar Nöth[1]*
*[1]Pattern Recognition Lab, Universität Erlangen-Nürnberg, Germany*
*[2]Digital publishing, München, Germany*

On a database of non-native English productions annotated by 60 native English speakers as for their quality w. r. t. intelligibility, non-native accent, melody and rhythm, we study how inter-labeller correlation and performance of a regression system change when varying the number of labellers used for training. This depends highly on the difficulty of the labelling task, the features used by the regression system and the type of regression used. We propose a model that parametrises these dependencies and is able to predict the system's performance when increasing the number of labellers. This can provide a valuable basis for decision-making when trying to improve an existing regression system as efficiently as possible. We show the plausibility of our approach by experimental evaluation.

## Oral Session 3: Production of a Second Language

**O3-1: Interaction of Lexical and Sentence Prosody in Taiwan L2 English**

*Tanya Visceglia[1], Chiu-yu Tseng[2], Zhao-yu Su[2], Chi-Feng Huang[2]*
*[1]Department of Applied English, Ming Chuan University, Taiwan*
*[2]Phonetics Lab, Institute of Linguistics, Academia Sinica, Taipei, Taiwan*

This study investigates the effect of sentence-level prosody on production of English lexical stress, comparing L1 English and L1 Taiwan Mandarin speaker groups. 4 L1 North American English speakers and 9 L1 Taiwan Mandarin speakers were asked to produce a set of 20 disyllabic and multisyllabic words embedded in three different prosodic contexts: neutral broad focus, at a phrase/sentence boundary, and in narrow focus. Results suggest that production of the prosodic cues to mark lexical stress (F0, duration and amplitude) becomes much more difficult for L2 speakers when disyllabic and multisyllabic words are embedded in higher-level prosodic contexts.

**O3-2: German Learners of Japanese - Perceptual and Prosodic Analysis of Utterances from a Meditative Setting**

*Hansjörg Mixdorff[1], Ryoko Hayashi[2], Yoriko Yamada-Bochynek[3], Keikichi Hirose[4], Hiroya Fujisaki[5]*
*[1]Dept. of Computer Science and Media, Beuth University of Applied Sciences, Berlin, Germany*
*[2]School of Intercultural Studies, University of Kobe, Japan*
*[3]Institute of Japanese Studies, Free University Berlin, Germany*
*[4]Department of Information and Communication Engineering, the University of Tokyo, Japan*
*[5]Professor Emeritus, the University of Tokyo, Japan*

This study examines how closely Japanese utterances produced by German learners match those of Japanese natives, especially with respect to word and sentence prosody. Utterances were rated perceptually by Japanese native listeners as well as evaluated with respect to timing and F0 contours. We found that the German learners speak at a similar speech rate, but their syllabic durations are much more at variance than those of the Japanese controls. Altogether their rhythmic patterns are more similar than those by the Japanese. Contrary to our expectations perceptually prominent so-called pseudo-accent syllables are not lengthened, although the timing of tonal transitions at these is slightly different from the Japanese norm. Some transitions were not at all realized due to "deaccenting" phenomena.

**O3-3: A Cross-Language Study of Compensatory Response to Formant-Shifted Feedback**

*Takashi Mitsuya[1], Ewen N. MacDonald[1], David W. Purcell[2], Kevin G. Munhall[1]*
*[1]Queen's University, Kingston, ON, Canada*
*[2]University of Western Ontario, London, ON, Canada*

Learning new sounds in a second language requires the acquisition of new motor routines and new sensorimotor planning systems needed to ensure coordination. Auditory feedback is an important part of the planning and control system required for fluent speech production. ESL vowel production was studied using a real-time formant perturbation technique to modify auditory feedback. Three groups of subjects (Native English, Japanese ESL, and Korean ESL) produced tokens of the English word "Head" with the first formant (F1) shifted either up or down in frequency. When F1 was shifted up, compensations by Native English speakers were larger than either ESL group. The F1 lowering perturbations produced more similar compensations by all three groups. This direction asymmetry in magnitude of compensation is discussed in relation to differences in native vowel inventories and the nature of auditory feedback processing.

**O3-4: Phoneme Errors in Read and Spontaneous Non-Native Speech: Relevance for CAPT System Development**

*Joost van Doremalen, Catia Cucchiarini, Helmer Strik*
*Department of Linguistics, Radboud University Nijmegen, The Netherlands*

For the purpose of pronunciation assessment and training in a second language both read and spontaneous speech are employed. In this paper we present the results of a study on the nature of phoneme errors in Dutch read and spontaneous non-native speech and discuss the possible consequences and relevance of these findings for the purpose of developing Computer Assisted Pronunciation Training systems.

## Demo Seeion 1: New Technologies and Methodologies Help Language Learning.

**D1-1: A New Force in the Interpretation Technology**

*Yoichi Tokioka[1], Kazuhiro Okuda[2]*
*[1]Straightword Inc., Japan*
*[2]Powershift Inc., Japan*

"is" is a newly developed interpretation-support software. Since interpretation has long depended on the personal ability of interpreter, improvement in interpretation technique by IT technology has been thought to be impossible. On the other hand, while voice recognition technology is being developed, the study of machine interpretation and translation has started and drawn many researchers' interests. However, we are not able to access to a system of practical use. Our "is" is an innovative system, which helps interpreters understand more accurately and interpret quickly what speakers say. "is" has good functions to reduce the cognitive load of interpreters, such as controller to take advantage of a little bit time lag occurring in the interpretations.
"is" has the following three advantages:
1. Reducing the cognitive load of interpreters will improve the quality of interpretations.
2. Recording device can produce bilingual voice proceedings.
3. "is" system is applicable to all languages.

**D1-2: CALL Systems at Kyoto University**

*Yasushi Tsubota, Tatsuya Kawahara, Masatake Dantsuji*
*Kyoto University, Japan*

We will demonstrate systems for learning second language pronunciation using speech processing technology developed at Kyoto University. The target languages are English, Chinese and Japanese. One system assists earners by automatically subtitling their videos made during class with the relevant dialogue. Another system pinpoints pronunciation errors. We have used these systems in CALL classrooms at Kyoto University and universities in China.

**D1-3: CALL-SLT/Web, A Speech-Enabled Translation Game on the Internet**

*Manny Rayner, Nikos Tsourakis, Pierrette Bouillon, Matthew Fuchs*
*University of Geneva, TIM/ISSCO, Switzerland*

CALL-SLT/Web is a multilingual CALL platform for intermediate language students who wish to practice their spoken fluency, based on the "translation game" idea of Wang and Seneff and implemented on top of the Nuance Toolkit and the Open Source Regulus platform. The system prompts the student using either a telegraphic L1 text or a graphical representation of it. The student responds using with a spoken sentence in the L2. The system performs speech recognition and translation to compare the response with the prompt, and gives informative feedback; there is also an elaborate online help component. Linguistic content exists for English, Japanese, French, German and Swedish in a tourist restaurant domain. CALL-SLT uses a distributed client/server architecture, with the greater part of the processing carried out on the server. Versions previously demonstrated (http://www.issco.unige.ch/pub/lrec2010_callslt.pdf) used either a Java client running on a laptop or a C++ client running on a handheld device. Here, we will demo a new version, where the client is a Flash application running inside a browser. Remote internet-enabled speech processing is implemented using the file-based protocol similar to the one used for the recently deployed Internet game Minion Dominion (http://www.miniondominion.com).

**D1-4: Phonensínthesis: A System for Teaching French Pronunciation to Japanese Students of French**

*Josafá de Jesus Aguiar Pontes, Sadaoki Furui*
*Tokyo Institute of Technology, Japan*

Our system allows to produce high quality pronunciations for standard French language. Basically, we target at liaison in French, which is known to be one of the most challenging issues for French grapheme-to-phoneme converters. We have developed a technique that enables our system to predict most of the obligatory and forbidden liaisons precisely and accurately. Our results are comparable to the state-of-the-art systems. In addition, we utilize one of the best voice ever developed for French, in order to produce high quality speech. You will be able to input your own French text and listen to the synthesized speech. Alternatively, you can also search for a sentence from our bilingual corpus of Japanese/French sentences, and have it synthesized. Please come to see our demonstration, test our system by yourself and kindly leave your precious feedback.

**D1-5: Automatic Assessment of Non-Native Prosody**

*Florian Hönig, Anton Batliner*
*Pattern Recognition Lab, Universität Erlangen-Nürnberg, Germany*

Non-native prosodic, especially rhythmic traits are a main source for low intelligibility of the speech of non-native L2 speakers of English - and any other language. Thus, they should be addressed in CAPT applications. To this aim, we present a demonstrator that uses a comprehensive set of prosodic features and linear regression to automatically assess the quality of the learner's productions with respect to intonation and rhythm. The system is trained with annotations obtained from 60 American and British natives, and surpasses the average performance of a single human rater.

**D1-6: Pronunciation Assessment System Using Structural Features**

*Masayuki Suzuki, Nobuaki Minematsu*
*The University of Tokyo, Japan*

Pronunciation assessment systems have to be able to ignore speaker individualities like age or gender ( the size and length of the vocal tube). Our system ignore the speaker individualities by using approximately speaker-invariant features called structural features. The demo system can

1. record or log a history of vowel pronunciation training of each learner.
2. show which vowel to correct first to become like a teacher.
3. classify all the registered learners only with regard to pronunciation proficiency by ignoring gender, age, etc. very effectively.

After the session, we may show a classification result of all the participants in our demonstration.

**D1-7: Dialect (Regional Accent)-Based Speaker Classification of Chinese**

*Xuebin Ma, Nobuaki Minematsu*
*The University of Tokyo, Japan*

In China, different dialects often have different sets of vowels and each dialect comes to have its unique vowel space. In this demo, a participant is asked to read about forty Chinese letters by his/her own dialect and the vowel segments are automatically detected. Using these vowel segments, their vowel space (vowel structure) is estimated based on speech structure theory, where the extra-linguistic factors such as gender and age are effectively removed from speech acoustics and only the linguistic aspect is extracted. Using the vowel structures, Chinese speakers are classified purely based on their dialectal characteristics.

**D1-8: JaFIX (Japanese as a Foreign Language with Integrative-Communicative Steps) -- JFL-Acquisiton through Relax and Gesticalization**

*Yoriko Yamada-Bochynek*
*Institute of Japanese Studies, Free University Berlin, Germany*

Free from the prototypical classroom constraints, i.e., desks and chairs, the JaFIX Method involves the Relaxation with a verbal induction and Gesticalization, acting-out of linguistic expressions. This corporeal approach realizes thus a "quasi-L1 acquisition" of the JFL, resulting in the learners' internalization of a "native-like" prosody.

## Poster Session 2: Science and Technology of Speech and Language for Education

**P2-1: A Study of Pitch Patterns of Sentence Utterances by Japanese Speakers of English in Comparison with Native Speakers of English**

*Tomoko Nariai, Kazuyo Tanaka*
*Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan*

This paper describes statistical analyses for identifying certain inherent ambiguities on pitch patterns of sentence utterances in English spoken by Japanese (Japanese English, henceforth). Statistical significance of pitch pattern differences between Japanese English and native English speakers is evaluated depending on the word position in a sentence and the word class, such as content word and function word. Results suggest that in Japanese English, sentences have lower pitch at the beginning and higher pitch at the end than sentences uttered by English speakers. Also, pitch ranges in sentences in Japanese English are narrower than those for English speakers. These indicate that intonation pattern in Japanese English is rather flat. Additionally, the results suggest that function words in Japanese English have higher pitch than English speakers.

**P2-2: Towards a Computer-Aided Pronunciation Training System for German Learners of Mandarin - Prosodic Analysis**

*Hussein Hussein[1], Hansjörg Mixdorff[1], Hue San Do[1], Si Wei[2], Shu Gong[2], Hongwei Ding[3], Qianyong Gao[2], Guoping Hu[2]*
*[1]Department of Computer Sciences and Media, Beuth University of Applied Sciences, Berlin, Germany*
*[2]Department of EEIS, University of Science and Technology of China, Hefei, Anhui, P.R.China*
*[3]School of Foreign Languages, Tongji University, Shanghai, China*

This paper reports on the continued activities towards the development of a computer-aided language learning system for German learners of Mandarin. In this experiment we used a complex corpus which consists of whole sentences and read from German students from three different years of language education and native speakers of Mandarin. A contrastive analysis of prosodic features (rhythmic and intonational) of the Mandarin tones between native speakers and German learners of Mandarin was performed to identify the differences and similarities. We aimed to study the effect of learning time of Mandarin on the development of learner's level. Therefore, the rhythmic and intonational features of tones were compared between German learners according to every year of language education. German students tend to exaggerate the F0 contours to discriminate the tones better and learn to adapt these to the tones of native speakers with increasing learning time. The syllable duration depending on the tone by German learner is longer than by native speakers and the changes of F0 parameter of Mandarin tones by German students are greater than by native speakers of Mandarin.

**P2-3: Effects of Pitch Cues on the Identification of Vowel Length in L2 Japanese**

*Izumi Takiguchi*
*Graduate School of Foreign Studies, Sophia University, Japan / JSPS Research Fellow*

This study explored the effects of pitch cues on the identification of the word-final Japanese vowel length, which is primarily cued by vowel duration. Native speakers of English (NE), Chinese (NC) and Japanese (NJ) participated in the experiment. Learners, who do not use duration distinctively in their L1, utilize duration as a cue for the contrast and they can approximate boundary location to NJ's. In addition, pitch cues did not affect NE's perception but it did affect NC's identification. These results suggest that the role of cues in learners' L1 relates to the use of cues in their L2.

**P2-4: Examination of the Relationship between L2 Perception and Production: An Investigation of English /r/-/l/ Perception and Production by Adult Japanese Speakers**

*Kota Hattori, Paul Iverson*
*Division of Psychology and Language Sciences, University College London, London, UK*

This study took an individual differences approach to examine the relationship between L2 speech perception and production, with the aim of examining whether they share common underlying representations. All Japanese speakers were assessed in terms of their /r/-/l/ identification, discrimination, best exemplars, and production. The results demonstrated that, although there was a moderate correlation between English /r/-/l/ identification and production, all other perceptual behaviors poorly related to /r/-/l/ production, suggesting that L2 speech perception and production processes and representations may be somewhat autonomous.

**P2-5: Consonant Cluster Production in Japanese Learners of English**

*Yoshiho Shibuya[1], Donna Erickson[2]*
*[1]Department of General Education, Kanazawa Medical University, Japan*
*[2]Department of Linguistics, Showa Music University, Japan*

Japanese speakers often face difficulty in producing complex syllable onsets in English and insert an extra vowel. We examined whether the vowel inserted by Japanese speakers was epenthetic (phonological) or excrescent (phonetic). The acoustic data suggested that an L1 phonological process was involved in vowel insertion by Japanese speakers with lower-level English competency, because the inserted vowels were similar to vowels in Japanese. More advanced speakers' results, on the other hand, suggested that phonetics may be involved. The articulatory data from this pilot study with one speaker supported the findings of the acoustic data, suggesting that both phonetics and phonology affect Japanese speaker's vowel insertion in a complex way. This paper is based on work previously reported in [1].

**P2-6: Can We Predict Who Will Benefit from Computer-Based Phonetic Training?**

*Valerie Hazan, Yoon Hyun Kim*
*Speech, Hearing and Phonetic Sciences, UCL, UK.*

This study investigated whether specific auditory or cognitive skills were linked to initial sensitivity to a novel phonetic contrast (Korean lenis-aspirated contrast) or to the degree of learning following computer-based phonetic training. Correlations between auditory or cognitive skills and phonetic perception were generally fairly weak, with measures of frequency acuity and attentional switching most often correlated with phonetic ability. The ability to sort stimuli according to a particular acoustic cue was also correlated with performance on the syllable and word identification tests. However, rate of learning was not correlated with any of the auditory or cognitive skills tested.

**P2-7: Attention to Critical Acoustic Features for L2 Phonemic Identification and Its Implication on L2 Perceptual Training**

*Yoon Hyun Kim[1], Jung-Oh Kim[2]*
*[1]Interdisciplinary Program in Cognitive Science, Seoul National University, Seoul, Korea*
*[2]Department of Psychology, Seoul National University, Seoul, Korea*

This study examined whether native speakers of Japanese could attend to critical acoustic features while identifying lenis and aspirated among Korean alveolar stops. Most Japanese participants had studied Korean for more than one year at a university language center in Korea. Native speakers of Korean were also tested with the same task for comparison. Korean participants discriminated the phonemic contrast according to both VOT and F0 or just F0. In contrast, Japanese participants identified lenis and aspirated mostly based on VOT information. They correctly identified stimuli of the phonemes which a speaker produced distinctively in terms of VOT. When stimuli weren't noticeably different in VOT, they confused the two phonemes. Unlike Korean participants, they hardly considered F0 information. This result suggests that some training materials, although they were produced by native speakers, can't lead L2 learners to catch critical acoustic information of L2 phonemes. If learners can identify L2 sounds accurately without attention to critical features, they may stick to wrong information in the sounds.

**P2-8: Speech Analysis for Automatic Evaluation of Shadowing**

*Dean Luo[1], Yutaka Yamauchi[2], Nobuaki Minematsu[1]*
*[1]The University of Tokyo, Tokyo, Japan*
*[2]Tokyo International University, Saitama, Japan*

This paper presents acoustic analysis for the purpose of automatic evaluation of shadowing speech. We use self-checked scores of understanding, manual prosodic scores, and TOEIC scores as reference scores of learners' shadowing speech, and compare these scores with automatic scores based on acoustic features that can reflect phoneme intelligibility and prosodic fluency in terms of intonation, and rhythm. We also examine the differences of personal-best shadowing, shadowing after the transcription is shown and reading-aloud of the same contents. Experimental results show that learners' understanding of contents in shadowing affects segmental intelligibility and prosodic fluency of their shadowing productions. A multiple regression model that combines different features can better reflect learners' understanding of the contents of shadowing and other reference scores, and thus suitable for automatic evaluation of shadowing.

**P2-9: Synthesizing Expressive Speech to Convey Focus Using a Perturbation Model for Computer-Aided Pronunciation Training**

*Fanbo Meng[1], Helen Meng[2,3], Zhiyong Wu[2,3], Lianhong Cai[1,3]*
*[1]Department of Computer Science and Technology, Tsinghua University, Beijing, China*
*[2]Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong SAR, China*
*[3]Graduate School at Shenzhen, Tsinghua University, Shenzhen, China*

We present a perturbation model that can modify the acoustic features of neutral speech in order to synthesize focus for certain words. In doing so, we can generate expressive speech output that highlights important speech segments to attract the listener's attention. The ultimate objective is to synthesize corrective feedback in a computer-aided pronunciation training (CAPT) system. This work involves the design and collection of a speech corpus, whose text prompts contain focus words. Each prompt is recorded twice – a neutral production followed by an expressive one where specific words are highlighted with focus. The phones in these recordings are modeled in six different classes, based on their relations with stressed syllables in focus words. Phone boundaries are obtained automatically by forced alignment with an automatic speech recognizer. Acoustic features of the phones, relating to f0, energy and duration, are extracted. Features that have highest correlation with the phone classes, as well as low variances, are incorporated into the perturbation model. The model is applied to neutral recordings of 20 test sentences. Results from a listening test show that the 13 subjects can identify the focus words with an accuracy of over 98%. The perceived degree of focus in the identified words achieves a mean score of 4.5 in a five-point Likert scale.

**P2-10: Multimodal Learning of Words: A Study on the Use of Speech Synthesis to Reinforce Written Text in L2 Language Learning**

*Kevin Dela Rosa, Gabriel Parent, Maxine Eskenazi*
*Language Technologies Institute, Carnegie Mellon University, USA*

Past research has shown that the use of multimedia, such as pictures, audio narration, and video, can be beneficial in computer aided instruction. We propose that spoken words generated by speech synthesis can be used to reinforce written text during L2 language instruction, and can lead to a more robust learning experience than providing written language input alone. Two in-vivo studies were conducted with ESL (English as a second language) students to investigate the effect of providing spoken language produced by speech synthesis during different instructional events in REAP, a computer based vocabulary tutor. Our results show that students benefit from spoken language input, particularly when they are strongly encouraged to listen to words. Furthermore, our studies seem to suggest that on demand English text-to-speech synthesis may be good enough to provide added value during computer based L2 language instruction.

**P2-11: Automatic Generation of Cloze Question Distractors**

*Rui Correia[1], Jorge Baptista[2], Nuno Mamede[1], Isabel Trancoso[1], Maxine Eskenazi[3]*
*[1]INESC-ID Lisboa / IST, Portugal*
*[2]Universidade do Algarve, Portugal*
*[3]Language Technologies Institute, Carnegie Mellon University, USA*

This paper presents a technique to generate distractors for cloze questions in the context of a Computer-Assisted Language Learning tutoring system. The document will focus on an evaluation process used to measure the quality of the distractors that were automatically generated. The main goal of the present study is to be able to include this feature in the tutoring system.

**P2-12: Automatic Selection of Collocations for Instruction**

*Adam Skory, Maxine Eskenazi*
*Language Technologies Institute, Carnegie Mellon University, USA*

For teaching of collocations no resource exists that comprehensively ranks collocations in terms of usefulness for learners. Towards developing a method to produce such a resource, we define a collocation's utility in terms of its unpredictability; the inability of a student to derive the meaning of the collocation from her semantic knowledge of its constituent words. We conduct an experiment comparing knowledge of phrasal verb collocations to familiarity with each collocation's verb constituent in order to have empirical measures of predictability. We then investigate corpus-based methods to approximate collocation predictability and find statistically significant correlations between a subset of these methods and the experimental data. This demonstrates that automated statistical approaches can significantly approximate the predictability of phrasal verbs according to our measures. We intend for this research to lead to development of resources for automated content selection in CALL.

**P2-13: Toward a Chanting Robot for Interactively Teaching English to Children**

*Ryo Nagata[1], Tomoya Mizumoto[2], Kotaro Funakoshi[3], Mikio Nakano[3]*
*[1]Faculty of Intelligence and Informatics, Konan University, Japan*
*[2]Graduate School of Information Science, Nara Institute of Science and Technology, Japan*
*[3]Honda Research Institute Japan Co., Ltd., Japan*

To acquire a second language, one must develop an ear and tongue for the correct stress and intonation patterns of that language. In English education, there is a rhythmic teaching method called Jazz Chants. This paper proposes a new application for second language education which combines Jazz Chants with a companion robot, and reports our technical investigations toward realizing such a robot. Investigated were two key technologies: predicting stresses in Jazz Chants and synthesizing chant speech. Experiments show promising results and reveal requirements for further improvement.

**P2-14: Applications of the Buckeye GTA Corpus for L2 Teaching and Research**

*Jocelyn B. Hardman[1], Elizabeth McCullough[2]*
*[1]Department of Languages, University of Dayton, USA*
*[2]Department of Linguistics, Ohio State University, USA*

The Buckeye GTA Corpus contains 9,664 L1 and L2 sentence productions by 89 talkers (27 American English, 19 Hindi, 23 Mandarin, & 20 Korean). A total of 5,696 sentences were read in English, with each talker contributing 64 sentences. Hindi, Mandarin, and Korean talkers also read 64 sentences each in their native languages, contributing a total of 3,968 sentences. Potential uses of the corpus are illustrated by research projects on classroom communication and acoustic phonetic patterns. These projects demonstrate how investigations in different disciplines can make use of the same corpus and provide converging data on second language phonological acquisition.

## Special Session: Primary School English Education in Asia

**S-1: Primary English Curriculum Reform in Beijing**

*Zehang Chen[1], Lingdi Shen[2]*
*[1]School of Foreign Languages and Literatures, Beijing Normal University, Beijing, China*
*[2]Research Center of Fundamental Education & Teaching, Beijing, China*

This paper presents an overview of the Beijing government's policy and implementation on primary English curriculum reform. TEYL in Beijing is undergoing an important phase of innovation along with challenges featured by differences in quality of teaching and learning between urban and rural areas due to shortage of qualified teachers and high quality materials and technology. Teachers' professional development is obviously the key for the success of TEYL (Teaching English to Young Learners). The strategies adopted by the government have effectively helped the implementation although there are still tasks to be fulfilled.

**S-2: Primary ELT in Korea: Start, Taxi, Take-off and Fly**

*WonKey Lee*
*Seoul National University of Education, Korea*

I would like to discuss the developmental stages of primary ELT in Korea, by using the analogy of an airplane's flying procedure: start, taxi, take-off and fly.

**S-3: An Overview of English Language Education at Primary Level in Taiwan**

*Chiou-lan Chern*
*Department of English, National Taiwan Normal University, Taiwan*

English is the major foreign language taught at schools in Taiwan (Crawford, 2003; Su, 2000). It is also the most commonly studied foreign language and the language used for wider communication in business and scholarly exchange. It had been traditionally taught beginning at Year 7 until 2001 when the Nine-year Integrated Curriculum was implemented and English was introduced to the Grade 5 curriculum (Chang, 2007; Chern, 2002). English was later lowered to Grade 3 curriculum in 2005. To accommodate this change, many policies were stipulated and implemented in the first decade of the 21st century. The aim of this paper is to provide an overview and discuss some pertinent issues of English language teaching (ELT) in elementary schools in Taiwan.

**S-4: Elementary School English Education in Japan -- Its History and the Sound of Its teaching Materials --**

*Kyoko Kasuya[1], Yuri Kuno[2]*
*[1]Department of Education, Tokyo Gakugei University, Japan*
*[2]Chubugakuin University, Japan*

In this symposium, we would like to outline the past and present of elementary school English education in Japan. Its history is briefly described in section 1. We also would like to examine some teaching methodologies and materials for children. What role sound plays in language is shown in section 2. It is important for teachers to select good teaching materials which do no harm on children's second language acquisition. We hope this symposium will be a fruitful opportunity for all the participants to recognize the crucial role of sound in second language education.

## Oral Session 4: Prosodic Training and Corrective Feedback

**O4-1: MusicSpeak: Capitalizing on Musical Rhythm for Prosodic Training in Computer-Aided Language Learning**

*Hao Wang[1], Peggy Mok[2], Helen Meng[1]*
*[1]Dept. of Systems Engineering and Engineering Management, The Chinese University of Hong Kong*
*[2]Dept. of Linguistics and Modern Languages, The Chinese University of Hong Kong*

This paper presents a system named MusicSpeak, which strives to capitalize on musical rhythm for prosodic training in second language acquisition. The system targets for Chinese (L1) speakers learning English (L2). Their speech rhythms are considered to be syllable-timed and stress-timed respectively. Hence, language transfer creates a challenge for Chinese learners in acquiring English rhythm. We develop an automatic procedure that can be applied to any English sentence, to cast rhythmic patterns in speech (based on alternating stressed and unstressed syllables) into rhythmic patterns in music (based on musical bars and beats). We collected speech recordings from 9 speakers uttering 15 English sentences, first in natural style and then in synchrony with the generated musical rhythm. Comparison between the two styles based on rhythm metrics suggests that the latter has higher variability and better approximates stress-timed rhythm.

**O4-2: Lexical Tones Learning with Automatic Music Composition System Considering Prosody of Mandarin Chinese**

***Siwei Qin, Satoru Fukayama, Takuya Nishimoto, Shigeki Sagayama***
***Graduate School of Information Science and Technology, the University of Tokyo, Japan***

Recent research has found that there is an overlap in the processing of music and speech in certain aspects. This research focuses on the relationship between the pitch of tones in language and the melody of songs. We present an automatic music composition system based on the prosody rules of Mandarin and we hypothesize that songs generated with our proposed system can help non-native Mandarin speakers to learn the tones of Mandarin Chinese more easily. To verify this hypothesis, twelve non-Chinese speakers from Japan were asked to identify and pronounce the Mandarin sentence they heard in the experiments with three different learning methods. The result shows that participants got higher accuracies of performances in tone3 with the teaching method of "speech + music" and the teaching method of "music only" is not more effective than "speech only" in some particular tones.

**O4-3: Practicing Syntax in Spoken Interaction: Automatic Detection of Syntactical Errors in Non-Native Utterances**

***Helmer Strik, Janneke van de Loo, Joost van Doremalen, Catia Cucchiarini***
***Department of Linguistics, Radboud University Nijmegen, The Netherlands***

In the current paper we present a new method, called SynPOS: Syntactic analysis using POS-tags. SynPOS is applied to a corpus of spoken human-machine interactions. The results show that language learners of Dutch often make syntactical errors, that there are many different types of syntactical errors, and that their frequencies vary a lot. This information can be used next to select errors and develop exercises for CALL systems.

**O4-4: Simicry - A Mimicry-Feedback Loop for Second Language Learning**

***Preben Wik, Björn Granström***
***KTH, Centre for Speech Technology CTT, Stockholm, Sweden***

This paper introduces the concept of Simicry, defined as similarity of mimicry, for the purpose of second language acquisition. We apply this method using a computer assisted language learning system called Ville on foreign students learning Swedish. The system deploys acoustic similarity measures between native and non-native pronunciation, derived from duration syllabicity and pitch. The system uses these measures to give pronunciation feedback in a mimicry-feedback loop exercise which has two variants: a 'say after me' mimicry exercise, and a 'shadow with me' exercise.

The answers of questionnaires filled out by students after several training sessions spread over a month, show that the learning and practicing procedure has a promising potential being very useful and fun.

**O4-5: The Role of Corrective Feedback in Second Language Learning: New Research Possibilities by Combining CALL and Speech Technology**

***Bart Penning de Vries, Catia Cucchiarini, Helmer Strik, Roeland van Hout***
***Centre for Language and Speech Technology, Radboud University Nijmegen, the Netherlands***

The role of corrective feedback (CF) is debated in second language acquisition (SLA). It has not been unequivocally shown that CF is effective in SLA, in particular not in the case of on-line processing, as in oral second language (L2) proficiency. This might be because, to date, it has not been feasible to create appropriate research conditions. We claim that these problems can be alleviated by resorting to a computer assisted language learning (CALL) environment in which learners receive CF individually, on spoken output. Learner output is analyzed using automatic speech recognition (ASR). In the project FASOP (Feedback on Syntax in Oral Proficiency) we intend to use an ASR-based CALL system to generate different types of CF and to test their effectiveness. The central question in this project is whether CF on syntax contributes to the development of oral proficiency when it is provided under near-optimal conditions.

## Demo Seeion 2: New Technologies and Methodologies Help Language Learning.

**D2-1: Changyan Interactive English Learning System**

*Guoping Hu*
*iFly Speech Lab, University of Science and Technology of China, Hefei, China*

Changyan interactive English learning system is a system that can help students to practice their spoken English in an interactive mode. The system has four major functions, which are word learning, paragraph reading aloud, scenario responding and game playing. Subjective evaluation from Chinese students indicates that this system is attractive and helpful.

**D2-2: A Multi-Player Vocabulary Game that Teaches While It Learns**

*Adam Skory, Maxine Eskenazi*
*Language Technologies Institute, Carnegie Mellon University, USA*

We will demonstrate one game from a set of multi-player webgames targeted at advanced English language learners, and particularly at those preparing for standardized English proficiency tests (such asTOEIC, TOEFL, GRE). These games will be made available to play for free online.

The suite of games not only uses state-of-the-art student models, but introduces the use of content models. One primary difficulty in text-based language games is that of content creation. Language technologies exist to find content from resources such as corpora and the Internet, however these technologies do not have the discriminative power of human content authors. Active learning of content models is achieved by building implicit crowdsourcing techniques into the mechanics of the games. As more people play, higher quality game content is identified and favored, and targeting of that content is improved. In other words, the players themselves become the content authors.

The first game within this framework is a fast-paced, cooperative game. Two players interact through hinting and guessing to cooperatively solve fill-in-the-blank questions, in the process building and sharing contextual and semantic knowledge of vocabulary.

**D2-3: NTU Chinese -- A Chinese Language Pronunciation Learning Software**

*Yow-Bang Wang[1,2], Lin-Shan Lee[1,2]*
*[1]Institute of Information Science, Academia Sinica, Taiwan, R.O.C*
*[2]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.*

NTU Chinese is a successfully operating online Chinese pronunciation learning software, specifically designed for giving the students opportunities to practice both their listening and speaking skills anytime and anywhere. It was a joint effort between National Taiwan University and some industry partners. The first version of NTU Chinese has been completed and made available on-line at http://chinese.ntu.edu.tw/.

NTU Chinese is able to evaluate the utterance produced by an individual learner from four different aspects: pronunciation, pitch, timing and emphasis. For those phonemes with scores below a threshold, a 3-dimensional video will show on the screen to demonstrate the actions of the vocal tract shape, including the relative positions among the lip, tongue and other articulators.

The scoring algorithm was trained with the scores given by real professional Chinese teachers, over a corpus produced by a group of real learners whose mother tongues are not Chinese. Both the above training corpus and course content currently used in this software were contributed by the International Chinese Language Program of National Taiwan University. The system to be demonstrated will have an improved scoring algorithm trained with a much more complete set of training corpus, as compared to its earlier version.

**D2-4: REAP.PT**

*Rui Pedro dos Santos Correia*
*INESC-ID Lisboa / IST, Portugal*

REAP.PT (REAder-specific Practice PorTuguese) is a tutoring system in the Computer Assisted Language Learning area that aims to teach vocabulary to students of Portuguese as a Second Language. REAP.PT is based on the importance of reading activities as a way to become proficient in a new language and, from the standpoint of the student, the learning method can be summarized in two main phases: text reading and question answering. REAP.PT innovates by using real texts (collected from the web) with the words that are being taught highlighted (questions about these target words are generated automatically, and are presented after each reading). Each text is classified according to topic and readability which allows the system to cross this information with the student's interests and level. An oral comprehension module provides the user with the possibility of synthesizing any word sequence, hearing audio books while seeing their text, and watching yesterday's broadcast news (audio plus video), in which each story is classified with topic and readability level, with an automatic text transcription on the side. REAP.PT also provides a teacher interface where, if in a class environment, the teacher can control and manage the students' interaction with the system.

**D2-5: Development of an Automatic Evaluation System of ESL/EFL Learners' Skills of Shadowing**

*Dean Luo[1], Yutaka Yamauchi[2], Nobuaki Minematsu[1]*
*[1]The University of Tokyo, Tokyo, Japan*
*[2]Tokyo International University, Saitama, Japan*

The CALL system developed in our project can enhance ESL/EFL learners' skills of shadowing by automatically evaluating these skills in terms of pronunciation, prosodic features and overall proficiency levels. Learners are required to record their shadowing into the computer while listening to passages read by a native speaker of English. After recording, they can listen to their voices and observe the sound waves of their own recording and the model one. Through auditory and visual comparison of the two recordings, they can understand the shortcomings of their performances and where they should practice more. Learners' shadowed speech is automatically analyzed and evaluated by the computer using speech information processing technology like GOP (goodness of pronunciation), fundamental frequency (F0), power and length of pauses. Their English proficiency levels measured by TOEIC (Test of English as International Communication) are also predicted and presented. Based on the results of automatic scoring, the learners can understand how well they have conducted shadowing objectively and also grasp their own proficiency levels. From the viewpoint of material development, this CALL system enables instructors to choose any speech data obtained from CDs, DVDs, Web sites, etc, and use them as practice materials. Instructors' selection of speech data suitable for the learners' interests and proficiency levels can increase student motivation and continuous use of this system, hence improving both aural and oral skills.

**D2-6: Automatic Tests of Spoken Spanish, Arabic, and Chinese; and 4-Skills Testing in English**

*Jared Bernstein, Jian Cheng, Elizabeth Rosenfeld*
*Pearson KT, Palo Alto, California, USA*

A fully automatic system administers proficiency tests for several languages. Scoring of speech is based on ASR and other speech processing technology. Tests are administered on-demand by telephone or computer anywhere in the world, then the test-takers' responses are automatically processed and diagnostic and overall scores are returned on the web. Spoken English, Spanish and Arabic tests are commercially available; a Spoken Chinese test will be demonstrated, but is not yet fully operational. These tests last about 15 minutes and offer 60-70 listen-speak items scored for grammar, vocabulary, fluency, and pronunciation. The system derives scores from response content and the timing and spectral information in the spoken productions. An automated 4-skill test of English (Versant Pro) measures listening, speaking, reading and writing skills. It assesses spoken and written English as used in many workplace settings including phone conversations, discussions, negotiations, note-taking at meetings, writing summaries, and responding to emails. Versant-Pro covers speaking skills, listening and reading comprehension, as well as as grammar, vocabulary, organization, and tone in writing. An oral reading fluency test called ORF measures English fluency in reading aloud. All six tests are available for demonstration. Ample evidence of score reliability and validity is available.

**D2-7: A Vowel Training System for All**

*Toshiko Isei-Jaakkola[1], Takatoshi Naka[2], Keikichi Hirose[3]*
*[1]Chubu University, Japan*
*[2]Chukyo University, Japan*
*[3]The University of Tokyo, Japan*

A vowel training system is being developed basically for all kinds of foreign language learners who can train by themselves on a three-dimensional vowel chart and computer graphic articulatory movements.

# Memo

# Memo