Limited Distribution

TR-I- 0230

# Research Activities of the Speech Processing Department

November 1989 through December 1990

ATR Interpreting Telephony Research Laboratories

October, 1991

ATR Interpreting Telephony Research Laboratories

Abstract

This volume is a complete collection of technical papers from Speech Processing Department, ATR Interpreting Telephony Research laboratories, published from November, 1989 through December, 1990. The research areas of the Speech Processing Department include:

1. Large Vocabulary Continuous Speech Recognition.
    (a) Feature-Based Approaches.
    (b) Hidden Markov Models.
    (c) Neural Network Approaches.
2. Speaker Adaptation and Noise-Robust Speech Recognition
3. Language Source Modeling.
4. Speech Synthesis by Rule.
5. Voice Conversion.
6. Speech Database

For additional information, contact:

Shigeki Sagayama, Head, Speech Processing Department
ATR Interpreting Telephony Research Laboratories
Inuidani, Seika-cho, Souraku-gun, Kyoto 619-02, JAPAN
Telephone: +81 7749 5 1311,
Telefax: +81 7749 5 1308
E-mail: saga@atr-la.atr.co.jp

# Preface: The Overview of the Research Activity of Speech Processing Department

## 1 Introduction

At ATR Interpreting Telephony Research Laboratories, speech research is being intensively pursued to realize automatic telephone interpretation which allows speech communication between telephone users speaking different languages. Automatic telephone interpretation, or *speech translation* according to more general terminology, is one of the technologies most eagerly awaited by people all over the world.

The utterance of one party will be translated into another language and heard by another. The basic constituent technologies are speech recognition, language translation, and speech synthesis, which are not only essential for realization of speech translation but also widely applicable in various other areas. Although this concept of automatic telephone interpretation is rather new and research is still in its early stages, it is rapidly attracting the interest of researchers and research organizations throughout the world, and is expected to become an area of common research interest.

## 2 Automatic Telephone Interpretation

### 2.1 History and general concept

ATR Interpreting Telephony Research Laboratories, a subsidiary of Advanced Telecommunications Research International, was established in 1986 to initiate basic research for telephone interpretation.

The basic components of an automatic telephone interpretation ("speech translation" in more general terminology) system are speech recognition, language translation, and speech synthesis. In more realistic applications, speaker adaptation for input, and speaker conversion for output, are also required.

Simple combination of these elements of current technology, however, does not give a satisfactory speech translation system. What is required of these components is somewhat different from a conventional application. The major differences are as follows.

- Particularly high performance both in speech recognition and translation is required. Unlike conventional language translation, there cannot be any "pre/post-editing", since both input and output are speech.
- Real-time operation is a strong requirement.
- Spoken language is quite different from written language. Though sentences are generally short and their structures are not particularly complicated in spoken language, spoken dialogues include elliptic and anaphoric expressions. They may also include many syntactically ill-formed expressions.

Thus, speech translation is a new area which requires significant innovations in these constituent technologies.

Although the ultimate goal of the telephone interpretation system will be universal dialogue in an unlimited domain, at present, the goal should be a more feasible one such as a system which is limited to specific, task oriented areas. ATR Interpreting Telephony Research Laboratories has selected an international conference registration task as a constrained task domain where the dialogue will be goal-directed and the vocabulary is expected to be limited to approximately 1,500 words.

### 2.2 ATR's Experimental Speech Translation System

Fig.1 illustrates the outline of the current experimental Japanese-English speech translation system at ATR Interpreting Telephony Research Laboratories. The input sentence speech uttered phrase by phrase is recognized with speaker adaptation capability under linguistic constraints and converted into several word sequence candidates. The number of candidates is narrowed down by dependency analysis and finally uniquely determined by linguistic analysis, followed by language translation consisting of linguistic analysis, Japanese-to-English language transfer, and English sentence generation. The final stage is to convert English sentences into speech with the original speaker's characteristics although this stage is not integrated in the system. Another direction, English-Japanese conversion, can be considered in the same way. As it seems inefficient to split our research efforts into Japanese and English aspects, our research topics mainly cover Japanese speech recognition, Japanese-English translation, and Japanese speech synthesis. Basically, the same approaches can be applied to the remaining parts.

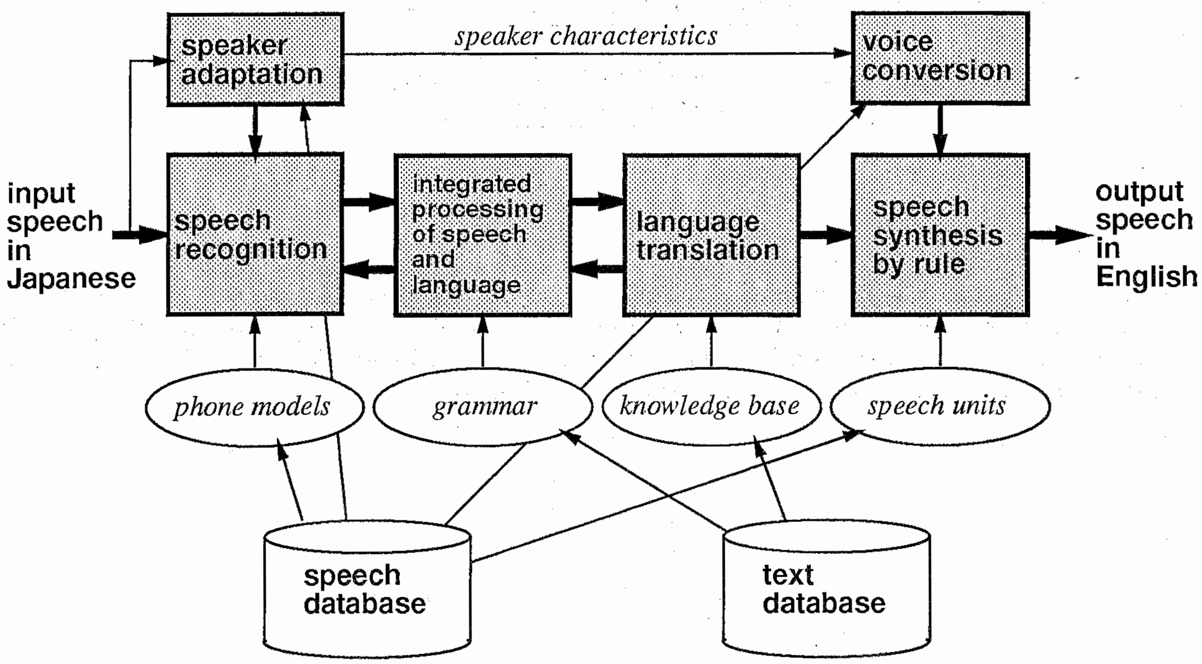



Figure 1: SL-Trans: an experimental speech translation system (Japanese-English)

# 3 Speech Research at ATR Interpreting Telephony Research Laboratories

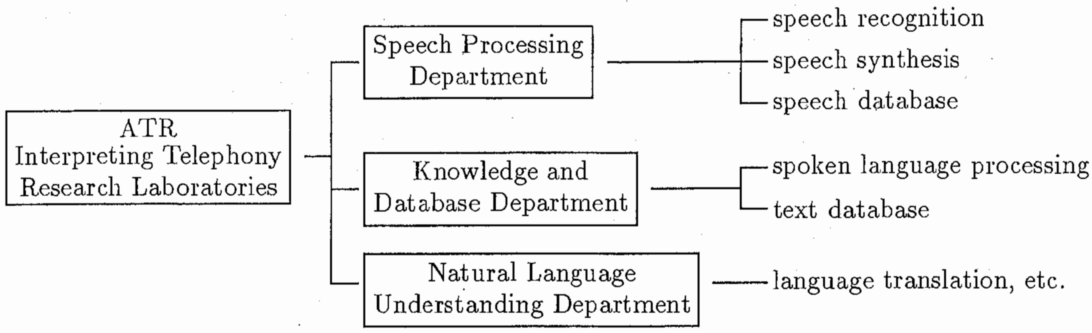
## 3.1 Research project

ATR Interpreting Telephony Research Laboratories (Akira Kurematsu, president) consists of three departments, Natural Language Understanding Department, Knowledge and Database Department, and Speech Processing Department. The laboratory was founded in April of 1986 with the support of the Japan Key Technology Center, ATR International, NTT, KDD, NHK, and other Japanese enterprises. The main target of our institute is fundamental research into speech and language processing, and integration of speech and language processing technologies to demonstrate the feasibility of an automatic telephone interpretation system. As of 1990, the laboratory had a research budget of 2.6 billion yen, a research staff of about 40, and 235 academic publications including 70 papers at international conferences.

## 3.2 Organization

ATR Interpreting Telephony Research Laboratories (Akira Kurematsu, President) consists of three departments, i.e., Speech Processing Department (Shigeki Sagayama, Head), Knowledge and Database Department (Tsuyoshi Morimoto, Head), and Natural Language Understanding Department (Hitoshi Iida, Head).

The outline of the organization is shown below.




## 4 Research Staff

The research staff is mainly composed of members seconded from research institutes, laboratories, and industrial companies which support ATR, and visiting scientists. The following table shows the past and current members of our research staff staying more than 6 months. In addition to the list, we have/had a number of intern students and working students from universities.

Table 1: Research Staff of the Speech Processing Department

| name & title | position | period |
|---|---|---|
| Akira Kurematsu, Dr | President | 1986.4 – |
| Kiyohiro Shikano, Dr | Department Head | 1986.6 –1990.1 |
| Shigeki Sagayama | Department Head | 1990.2– |
| Hisao Kuwabara, Dr | Supervisor | 1986.10 – 1989.7 |
| Yoshinori Sagisaka, Dr | Superviser | 1986.4 – |
| Tetsuo Umeda | Supervisor | 1989.7 – 1990.6 |
| Masahide Sugiyama, Dr | Supervisor | 1990.2– |
| Takeshi Kawabata, Dr | Senior Researcher | 1986.9 –1990.2 |
| Shin-ichi Tamura | Senior Researcher | 1986.9 –1990.2 |
| Hidefumi Sawai, Dr | Senior Researcher | 1988.4 –1991.3 |
| Satoshi Nakamura | Researcher | 1986.9 – 1989.8 |
| Kaichiro Hatazaki | Researcher | 1986.12 – 1989.3 |
| Kazuya Takeda | Researcher | 1986.8 – 1990.2 |
| Masanori Miyatake | Researcher | 1986.9 – 1989.8 |
| Toshiyuki Hanazawa | Researcher | 1987.3 –1990.2 |
| Masanobu Abe | Researcher | 1987.4 – 1991.2 |
| Katsuteru Maruyama | Researcher | 1987.3 –1990.2 |
| Masami Nakamura | Researcher | 1987.9 – 1990.8 |
| Katsuo Abe | Researcher | 1987.3 –1990.2 |
| Yasuhiro Komori | Researcher | 1988.9 – |
| Hiroaki Hattori | Researcher | 1989.5 – |
| Kazumi Ohkura | Researcher | 1989.9 – |
| Nobuyoshi Kaiki | Researcher | 1989.11 – |
| Jun-ichi Takami | Researcher | 1989.11 – |
| Kazuki Katagishi, Dr | Researcher | 1990.2 – |
| Akito Nagai | Researcher | 1990.3 – |
| Keiji Fukuzawa | Researcher | 1990.4 – |
| Katsuhiko Mimura | Researcher | 1990.5 – |
| Shingo Fujiwara | Researcher | 1990.5 – |
| Naoto Iwahashi | Researcher | 1990.10 – |
| Jin-ichi Murakami | Researcher | 1991.3 – |
| Yoshinaga Kato | Researcher | 1991.4 – |
| Kouichi Yamaguchi | Researcher | 1991.5 – |
| Yasunaga Miyazawa | Researcher | 1991.7 – |
| Tetsuo Kosaka | Researcher | 1991.9 – |
| Alex Waibel, Dr | Invited Researcher | 1987.5 – 1988.8, 1989.6 – 1989.9 |
| William Poser, Dr | Invited Researcher | 1988.9 – 1989.2 |
| Alain de Cheveigné, Dr | Invited Researcher | 1989.2 – 1990.2 |
| Wilhelm N. Campbell | Visiting Researcher | 1990.8 – |
| David Rainton, Dr | Visiting Researcher | 1990.4 – |
| Harald Singer | Visiting Researcher | 1991.3 – |
| Kouichi Murayama | Engineer | 1987.9 – 1988.8 |
| Takaharu Tanaka | Engineer | 1988.9 – 1989.8 |
| Kouji Kitagaito | Engineer | 1989.9 – 1990.8 |
| Shinobu Araki | Engineer | 1990.9 – 1991.8 |
| Yoshinori Ono | Engineer | 1991.9 – |
| H. Segot | Student | 1987.3 – 1987.11 |
| Patrick G. Haffner | Student | 1988.3 – 1988.11 |
| Jean-Claude Dang | Student | 1989.3 – 1989.11 |
| Yasuhiro Minami | Student | 1989.4 – 1990.3 |
| Alain Biem | Student | 1990.8 – 1991.1 |
| Romain Brunias | Student | 1991.8 – |
| Richard Lengagne | Student | 1991.8 – |

# 5 Research Facilities

In the year of 1990, we renovated our research facility. Each of research staff is provided a DECstation5000 (speed 24MIPS) or DECstation3100 (speed 16MIPS), with 24 - 32 Mbyte memory and a 600M - 3G byte disk, which is connected to an Ethernet. We introduced 2 DECsystem5000s (currently, 3 systems) with over 30G byte disk memory which are used as file servers containing a large scale speech database and are accessed from workstations through the Ethernet. Some workstations are equiped with speech input/output devices. Older computer systems including VAX8600, a VAX8700, a VAX8800 and an Alliant 9800 (eight-CPU array processor) are also connected to the Ethernet and accessed from workstations.

# 6 Speech Recognition Research

## 6.1 Phone Modeling

Reliable phoneme recognition and segmentation algorithms have been investigated leading to considerable improvements over conventional approaches. We have been pursuing three approaches: feature-based approach, hidden Markov model approach, and neural network approach. These improvements resulted in the successful implementation of a continuous speech recognition system such as HMM-LR, by combining HMM phoneme models with the generalized LR parsing algorithm. Hardware implementation has also been done.

### 6.1.1 Hidden Markov models

Hidden Markov models have been intensively studied for representing phone models in continuous speech recognition.

At the early stage of HMM research, the number of states, tied probabilities of transition and output, probability smoothing techniques, initial probability settings and duration control techniques were studied, based on the task of phoneme recognition using our large vocabulary speech database. Separate VQ (multiple codebook) and fuzzy VQ techniques improved the recognition rate from 86.5% without them, up to 95.7%. Major research topics cover the following items.

- discrete HMM phone models
- HMM phone models based on fuzzy vector quantization
- continuous output probability density HMMs
- allophonic (context dependent) HMM phoneme models
- concatenated training
- segment-based HMM
- HMM - neural network hybrid

HMM phone models are used in HMM-LR where the HMM phoneme models are combined with a generalized LR parser (a language source model) to efficiently recognize a Japanese phrase input, word spotting in continuous speech, phonetic typewriter, and English word recognition.

A recent study on context-dependent phone models represented by single Gaussian output probabilities has attained a speaker-dependent phoneme recognition rate of 98.3% for phonemes in word utterances. For phonemes in phrase utterances, they performed even better than the Gaussian mixture models.

### 6.1.2 Neural network approaches

ATR started neural network research in its early period and intensively investigated its application to speech recognition. The TDNN (time-delay neural network) is one of our developments. Major topics in this area cover:

- time delay neural network (TDNN)
- fast back-propagation algorithm
- neural network work bench
- phoneme spotting
- phoneme filter approach
- robustness across differences of speaking styles

( neighbor integration, KNIT ($k$-nearest neighbor interpolative training), pairwise discriminant TDNNs, fuzzy training )

- new architectures
- Boltzmann machine approach

TDNN attained a phoneme recognition rate of 98.6% for a /b/, /d/, /g/ task, and 96.7% for a full consonant task. Phoneme spotting by TDNN aimed at determining a continuous speech recognition approach by neural networks has been carried out and resulted a phoneme spotting rate of 98%.

Efforts to speed up the back-propagation algorithm resulted in a 1,000-fold speed-up of the TDNN. This speed-up and the use of a mini-supercomputer have made it possible to challenge larger scale neural networks.

Robustness across different speaking styles is one of most important issues in neural networks for speech recognition. Although neural networks show very high phoneme recognition performance for word speech when they have been trained using word utterances, the performance degrades drastically for different speaking styles such as continuous speech. To solve this problem, intensive studies have been made and new schemes of neural network training (e.g., KNIT method, fuzzy training) and new architectures (e.g., pairwise discriminant TDNNs) have been developed, and they perform much better than the original neural network approaches.

Other neural network approaches such as a deterministic Boltzmann machine and a neural prediction model are also being studied. A constructive neural network (CNN) to recognize words in bottom-up fashion using a TDNN phoneme spotting network is also studied.

### 6.1.3 Feature-based approaches

A feature-based approach is also taken for utilizing expert knowledge in spectrogram reading. Major topics in this area are:

- spectrum reading expert system
- hybrid system of knowledge based segmentation and neural network phoneme recognition
- feature based segmentation combined with HMM

Phoneme segmentation knowledge for consonants is described by ART. A phoneme segmentation expert is integrated with a TDNN neural network for consonant discrimination. The system attains a consonant segmentation rate of 94.5% and a consonant recognition rate of 88.8%. Moreover, spotting of vowels, semi-vowels and syllabic nasal has been studied using a vowel spotting TDNN.

## 6.2 Speaker Adaptation and Noise-robust Speech Recognition

Aiming at a general preprocessor for speaker normalization, speaker adaptation research is studied.

- codebook mapping for speaker adaptation
- supplemented codebook mapping
- speaker weighted HMM training
- HMM model modification
- speaker mapping by a neural network
- speaker independent phoneme recognition by neural networks

Codebook mapping is a technique to find a mapping between a pair of vector quantization codebooks of different speakers using fuzzy vector quantization and DTW(dynamic time warping)-based word speech matching. This HMM speaker adaptation algorithm attains a phrase recognition rate of 78.6%, where the phrase recognition rates of speaker independence and speaker dependence are 59.6% and 88.4%, respectively. A supplemented HMM phoneme model training approach has been initiated to cope with speaker coarticulation variations. In the codebook mapping algorithm, a speaker adaptation approach has been undertaken using the concept of fuzzy vector quantization and spectrum mapping.

Recently, instead of such a frame-wise approach, adaptation of continuous output density HMM models has been found to be very successful and has attained a phoneme recognition performance higher than 90% after speaker adaptation.

## 6.3 Language Source Modeling

Phone models are connected under the control of a language model to recognize continuous speech. We are taking both syntactic and stochastic approaches to language modeling. Some of our activities in this area are as follows.

- generalized predictive LR parser
- HMM-LR continuous speech recognition
- phoneme context dependent LR parsing algorithm
- word/syllable/phoneme trigram models
- phonetic typewriter using phoneme sequence statistics
- word spotting based on HMM phone models
- two-stage LR parser for inter-*bunsetsu* syntax
- unknown word detection in HMM-LR
- hardware implementation of the HMM-LR scheme
- the use of pitch patterns for phrase boundary detection
- NETgram: neural network approach for word category prediction
- TDNN-LR continuous speech recognition
- PD-TDNN-LR, fuzzy-TDNN-LR continuous speech recognition schemes

A generalized LR parser has been combined with HMM phone models to search for the optimal grammatical path. Fig.2 shows an outline of the HMM-LR scheme for continuous speech recognition which

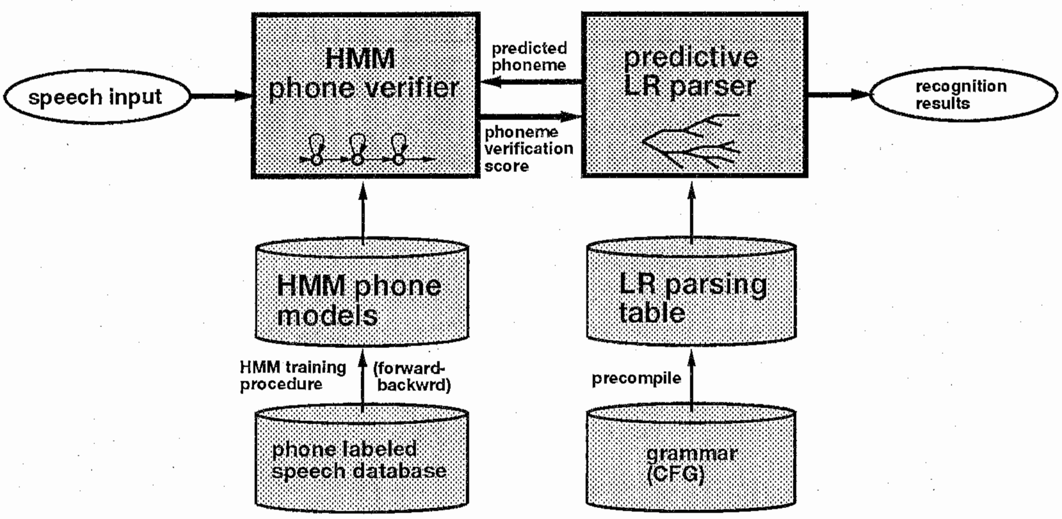



Figure 2: HMM-LR continuous speech recognition system

attained a phrase recognition rate of 88% for a task with a phoneme perplexity of 5.9 with vocabulary of 1,000 words. The cooccurrence of context-free rewriting rules is used to produce a better language source model. The LR parser is also successfully combined with the Sphinx system at Carnegie-Mellon University (CMU) under a research collaboration between ATR and CMU. A recent enhancement of the LR parser has attained a word accuracy of 97.5% and a sentence accuracy of 91.2%.

An experiment to recognize continuous utterances using TDNN phoneme spotting results has been tried by summing up the phoneme spotting output by means of a DTW algorithm with an LR parser. Recently, it is connected to both a PD-TDNN and a fuzzy trained TDNN and is attaining better results.

The NETgram approach for word category prediction using a neural network is also proposed and proved to be efficient when the training data size is relatively small compared with $n$-gram (bigram, trigram, etc.) approaches.

# 7 Speech Synthesis Research

## 7.1 Speech Synthesis by Rule

A high quality speech synthesis system by rule should be realized as an important part of the interpreting telephony system. A synthesis system based on flexible synthesis units has been developed. Moreover, the speech synthesis algorithm is evaluated to realize a better synthesis system. A unit selection algorithm from a large-scale speech database and a unit concatenation algorithm have been studied. The prosody control algorithm has been also studied. Also, rules between speaking styles and prosodic control have been studied to realize various kinds of synthesized voices. Our research interests are:

- non-uniform synthesis units
- optimal selection algorithm of units from a database for speech synthesis
- prosody control models using neural networks
- cepstrum-based speech waveform generation
- optimal segmental duration control using a statistical method
- optimal amplitude control using a statistical method
- analysis of segmental durations in English speech
- automatic phoneme segmentation for speech synthesis using HMM

A prototype speech synthesis-by-rule system has been implemented.

## 7.2 Voice Conversion

Voice conversion from one speaker to another is an important aspect in realizing an automatic telephone interpreting system. Formerly, a voice conversion algorithm based on codebook mapping by vector quantization was developed. Recently, our research interest has focus on two aspects. One is the application to cross language voice conversion, where speech data from speaker $S_1$ of source language $L_1$ and from speaker $S_2$ of destination language $L_2$ are given, and the latter is to be converted into speaker $S_1$'s voice. The other is a segment-based approach to voice conversion, where the input speech is segmented using HMM phone models and segment-wisely converted into anther speaker's voice. Our research subjects cover:

- frame-wise voice conversion using VQ codebook mapping
- cross-language voice conversion
- segment-based voice conversion using HMM

# 8 Speech Database

A large-scale speech database with phonetic transcription, mainly for research in speech recognition and speech synthesis, has been developed with considerable support from other speech research laboratories. This database includes three categories as follows.

(1) set-A: Large-scale database consisting of 5,240 Japanese common words, 216 phonetically-ballanced words, 101 Japanese syllables, 25 numerics, 35 alphabetical letters, 9 foreign words, and 115 sentences in a conference registration task with 3 different speaaking styles, uttered by 20 professional broadcast announcers

(2) set-B: 503 phonetically balanced sentences uttered by 10 professional speakers

(3) set-C: a database of a large number of speakers (mainly for speaker-independent/speaker-adaptive speaker recognition researches), containing 520 words (a subset of 5240 words in set-A), 216 phonetically-ballanced words, 15 numerics, and 150 sentences (a subset of set-B), with an ultimate goal of 200 speakers, half of which will be phonetically labeled.

These are recorded and sampled at 20 kHz, digitized in 16 bits, and phonetically labeled by well-trained workers by hand.

The former two, (1) and (2), have been already completed and are gradually being put to public availability.

The speech database has effectively been used to perform speech research at ATR and other laboratories. Data of several speakers are already available in CD-ROMs. This speech database is now being accepted as the standard in Japan.

# 9 Techical Publications from November 1989 through December 1990

Folowing pages are the list of technical publications from the Speech Processing Department, ATR Interpreting Telephony Research Laboratories, published during the period from November 1989 through December 1990, and reprints of these publications.

Abbreviations used in the list are as follows:

- ASJ: the Acoustical Society of Japan
- IEICE: the Institute of Electronics, Information and Communications Engineers (Japan)

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Abe,K 89-11 | 音韻環境に応じた音声合成素片の接続方法の検討<br><br>[On the Concatenation of Speech Synthesis Units according to Unit Extraction Context] | 安部　勝雄<br>武田　一哉<br>匂坂　芳典<br>[K. Abe<br>K. Takeda<br>Y. Sagisaka] | 電子情報通信学会技術研究報告 SP89-66, pp.17-22, 1989.<br><br>[IECIE Technical Report, SP89-66, pp.17-22, 1989] | 1 |
| Abe,K 90-2a | 波形重ね合わせ法による合成音の品質について<br><br>[Quality Evaluation for Synthesized Speech Using Wave Overlap Adding] | 安部　勝雄<br>匂坂　芳典<br>桑原　尚夫<br>[K. Abe<br>Y. Sagisaka<br>H. Kuwabara] | ATR Technical Report TR-I-0135 | - |
| Abe,K 90-2b | 音韻環境に応じた音声合成素片の接続方法の検討<br><br>[On the Concatenation of Speech Synthesis Units According to Unit Extraction Context] | 安部　勝雄<br>武田　一哉<br>匂坂　芳典<br>[K. Abe<br>K. Takeda<br>Y. Sagisaka] | ATR Technical Report TR-I-0136 | - |
| Abe,K 90-3 | 規則合成におけるパワー制御の検討<br><br>[A Study on Power Control of Speech Synthesis Units] | 安部　勝雄<br>武田　一哉<br>匂坂　芳典<br>[K. Abe<br>K. Takeda<br>Y. Sagisaka] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 1-4-12, pp.201-202, 1990.<br>[Proc. of ASJ Spring Meeting, 1-4-12, pp.201-202, 1990] | 7 |
| Abe,M 89-5 | **Voice Conversion Through Vector Quantization** | M. Abe<br>H. Kuwabara<br>S. Nakamura<br>K. Shikano | The Journal of the Acoustical Society of Japan, pp.71-76, 1989. | 9 |
| Abe,M 89-11 | **Fundamental Frequency Database with Linguistic and Phonetic Information** | M. Abe<br>Y. Sagisaka<br>H. Kuwabara | The Journal of the Acoustical Society of America, Supplement 1, Vol.86, O8, pp.S36, Fall, 1989. | 15 |
| Abe,M 89-12 | **Cross-Language Voice Conversion** | M. Abe | ATR Technical Report TR-I-0126 | - |
| Abe,M 90-2 | 言語間にわたる声質変換<br><br>[Cross-Language Voice Conversion] | 阿部　匡伸<br><br>[M. Abe] | 電子情報通信学会技術研究報告 SP89-123, pp.31-38, 1990.<br>[IECIE Technical Report, SP89-123, pp.31-38, 1990] | 16 |
| Abe,M 90-3a | 声質変換の研究<br><br>[A Study on Voice Conversion] | 阿部　匡伸<br><br>[M. Abe] | ATRジャーナル, No.7, pp.8-11, 1990.<br>[ATR Journal, No.7, pp.8-11, 1990] | 24 |
| Abe,M 90-3b | **Cross-Language Voice Conversion** | M. Abe | 日本音響学会 平成2年度春季研究発表会講演論文集, 3-4-11, pp.289-290, 1990.<br>[Proc. of ASJ Spring Meeting, 3-4-11, pp.289-290, 1990] | 28 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Abe,M 90-4 | **Cross-language Voice Conversion** | M. Abe<br>K. Shikano<br>H. Kuwabara | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S6a.14, pp.345-348, Albuquerque, 1990. | 30 |
| Abe,M 90-6 | **Voice Conversion for an Interpreting Telephone** | M. Abe<br>K. Shikano<br>H. Kuwabara | Proceedings of the Tutorial and Research Workshop on Speaker Characterization in Speech Technology, pp.40-45, Edinburgh, 1990. | 34 |
| Abe,M 90-9a | 音声セグメントを変換の単位とする声質変換<br>[A Segment Model Based Approach to Voice Conversion] | 阿部　匡伸<br>嵯峨山　茂樹<br>梅田　哲夫<br>[M. Abe<br>S. Sagayama<br>T. Umeda] | 日本音響学会 平成元年度秋季研究発表会講演論文集, 3-6-11, pp.287-288, 1990.<br>[Proc. of ASJ Fall meeting, 3-6-11, pp.287-288, 1990] | 40 |
| Abe,M 90-9b | 研究用日本語音声データベース利用解説書(連続音声データ編)<br>[Speech Database User's Manual] | 阿部　匡伸<br>匂坂　芳典<br>梅田　哲夫<br>桑原　尚夫<br>[M. Abe<br>Y. Sagisaka<br>T. Umeda<br>H. Kuwabara] | ATR Technical Report TR-1-166 | - |
| Abe,M 90-11 | **Statistical Study on Voice Individuality Conversion across Different Languages** | M. Abe<br>S. Sagayama | Proceedings of 1990 International Conference on Spoken Language Processing, 5.8, pp.157-160. Kobe,1990. | 42 |
| Campbell 90-11a | **Evidence for a Syllable-based Model of Speech Timing** | Nick Campbell | Proceedings of 1990 International Conference on Spoken Language Processing, 1.3, pp.9-12, Kobe, 1990. | 46 |
| Campbell 90-11b | **Duration, Pitch and Diphones in the CSTR TTS System** | Nick Campbell<br>S.D. Isard<br>A.I.C.Monaagham<br>J.Vechoeveen | Proceedings of 1990 International Conference on Spoken Language Processing, 19.16, pp.825-828, Kobe, 1990. | 50 |
| Cheveigné 90-2a | **The MapSignal Remote Speech Editor** | Alain de Cheveigné | ATR Technical Report TR-I-0137 | - |
| Cheveigné 90-2b | **Experiments in Pitch Extraction** | Alain de Cheveigné | ATR Technical Report TR-I-0138 | - |
| Cheveigné 90-2c | **Auditory Nerve Fiber Spike Generation Model** | Alain de Cheveigné | ATR Technical Report TR-I-0139 | - |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Dang 90-1a | シフト不変型決定論的ボルツマンマシンによる音声認識<br><br>[Shift-invariant Deterministic Boltzmann Machines for Phoneme Recognition] | Jean-Claude Dang<br>田村　震一<br>沢井　秀文<br>[S. Tamura<br>H. Sawai] | 電子情報通信学会技術研究報告 SP89-98, pp.1-6.1990.<br><br>[IECIE Technical Report, SP89-98, pp.1-6, 1990] | 54 |
| Dang 90-1b | シフト不変型決定論的ボルツマンマシンによる音声認識<br><br>[Shift-invariant diterministic Boltzmann Machines for Phoneme Recognition] | Jean-Claude Dang<br>田村　震一<br>沢井　秀文<br>[S. Tamura<br>H. Sawai] | ATR Technical Report TR-I-0130 | - |
| Fukuzawa 90-9 | ニューラルネットワークによる恒等写像を用いた話者適応<br>[Speaker Adaptation Using Identity Mapping] | 福沢　圭二<br>沢井　秀文<br>杉山　雅英<br>[K. Fukuzawa<br>H. Sawai<br>M. Sugiyama] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 1-8-16, pp.31-32, 1990.<br>[Proc. of ASJ Fall Meeting, 1-8-16, pp.31-32, 1990] | 60 |
| Fujiwara 90-7 | ホルマントを用いる音声合成のためのルールインタプリタの開発<br><br>[Development of a Rule Interpreter for Speech Synthesis by Rule using the Formant Synthesizer ] | 藤原　宏之<br>野村　康雄<br>海木　延佳<br>鬼頭　惇悟<br>山下　洋一<br>溝口　理一郎<br>[H. Fujiwara<br>Y. Nomura<br>N. Kaiki<br>A. Kito<br>Y. Yamashita<br>R. Mizoguchi] | 電子情報通信学会技術研究報告, SP90-29, pp.9-16, 1990.<br><br>[IECIE Technical Report, SP90-29, pp.9-16, 1990] | 62 |
| Gurgen 90-11 | **Line Spectrum Pair-Frequency-based Distance Measures for Speech Recognition** | Fikret Gurgen<br>S. Sagayama<br>S. Furui | Proceedings of 1990 International Conference on Spoken Language Processing, 13.1, pp.521-524, Kobe, 1990 | 70 |
| Hanazawa 89-12 | HMM-LR音声認識システムの性能評価<br><br>[HMM-LR Speech Recognition System Performance] | 花沢　利行<br>川端　豪<br>北　研二<br>中村　哲<br>鹿野　清宏<br>[T. Hanazawa<br>T. Kawabata<br>K. Kita<br>S. Nakamura<br>K. Shikano] | 電子情報通信学会 技術研究報告, SP89-94, pp.63-70, 1989.<br><br>[IECIE Technical Report, SP89-94, pp.63-70, 1989] | 74 |
| Hanazawa 90-2a | Hidden Markov Model による音韻認識実験の結果<br><br>[Phoneme Recognition Using Hidden Markov Models] | 花沢　利行<br>川端　豪<br>鹿野　清宏<br>[T. Hanazawa<br>T. Kawabata<br>K. Shikano] | ATR Technical Report TR-I-0147 | - |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Hanazawa<br>90-2b | HMM音韻認識における音韻連鎖統計情報の利用<br><br>[HMM Phoneme Recognition Using Syallable Trigrams] | 花沢　利行<br>川端　　豪<br>伊藤　克亘<br>鹿野　清宏<br>[T. Hanazawa<br>T. Kawabata<br>K. Ito<br>K. Shikano] | ATR Technical Report<br>TR-I-0148 | - |
| Hanazawa<br>90-3 | HMM音韻認識における音節連鎖統計情報の利用<br><br>[HMM Phoneme Recognition Using Syllable Trigrams] | 花沢　利行<br>川端　　豪<br>伊藤　克宣<br>鹿野　清宏<br>[T. Hanazawa<br>T. Kawabata<br>K. Itoh<br>K. Shikano] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 3-3-9, pp.87-88, 1990.<br><br>[Proc. of ASJ Spring Meeting, 3-3-9, pp.87-88, 1990] | 82 |
| Hanazawa<br>90-4 | **ATR HMM-LR Continuous Speech Recognition System** | T. Hanazawa<br>K. Kita<br>T. Kawabata<br>S. Nakamura<br>K. Shikano | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S2.4, pp.53-56, Albuquerque,1990 | 84 |
| Hanazawa<br>90-10 | HMM-LR音韻認識システムの性能評価<br><br>[HMM-LR Speech Recognition System Performance] | 花沢　利行<br>北　　研二<br>中村　　哲<br>川端　　豪<br>鹿野　清宏<br>[T. Hanazawa<br>K. Kita<br>S. Nakamura<br>T. Kawabata<br>K. Shikano] | 日本音響学会 誌, 46巻, 10号, pp.817-823, 1990.<br><br>[The Journal of the Acoustical Society of Japan, Vol.46, 10, pp.817-823, 1990] | 88 |
| Hatazaki<br>90-1 | スペクトログラムリーディング知識を用いた音韻セグメンテーションエキスパートシステム<br>[Phoneme Segmentation Expert System Using Spectrogram Reading Knowledge] | 畑崎香一郎<br>小森　康弘<br>川端　　豪<br>鹿野　清宏<br>[K. Hatazaki<br>Y. Komori<br>T. Kawabata<br>K. Shikano] | 電子情報通信学会論文誌, D-II, Vol.J73-D-II, No.1, pp.1-9, 1990.<br><br>[Trans. of IEICE, D-II, Vol.J73-D-II, No.1, pp.1-9, 1990] | 95 |
| Hattori<br>89-12 | 話者重畳型HMMによる分節認識<br><br>[Speech Recognition Using Supplemented HMM] | 服部　浩明<br>中村　　哲<br>鹿野　清宏<br>[H. Hattori<br>S. Nakamura<br>K. Shikano] | 電子情報通信学会 技術研究報告, SP89-90, pp.31-38, 1989.<br><br>[IECIE Technical Report, SP89-90, pp.31-38, 1989] | 104 |

List of Technical Publications from Speech Processing Department
(November 1989 through December 1990)
**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Hattori 90-3 | 話者適応における複数標準話者への重み付け<br><br>[Speaker Weighted Supplemented HMM for Compensation of Speaker Articulatory Variations] | 服部　浩明<br>中村　　哲<br>鹿野　清宏<br>[H. Hattori<br>S. Nakamura<br>K. Shikano] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 2-3-1, pp.51-52, 1990.<br>[Proc. of ASJ Spring Meeting, 2-3-1, pp.51-52, 1990] | 112 |
| Hattori 90-4 | **Supplementation of HMM for Articulatory Variation in Speaker Adaptation** | H. Hattori<br>S. Nakamura<br>K. Shikano | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S3.6, pp.153-156, Albuquerque, 1990 | 114 |
| Hattori 90-9 | コード遷移確率に基づく学習データ重み付けによる話者適応化<br>[Speaker Adaptation with Training Sample Weighting Based on Malkov Representation of Speaker] | 服部　浩明<br>嵯峨山茂樹<br><br>[H. Hattori<br>S. Sagayama] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 1-8-15, pp.29-30, 1990.<br>[Proc. of ASJ Fall Meeting, 1-8-15, pp.29-30, 1990] | 118 |
| Hattori 90-11 | **Speaker Weighted Training of HMM Using Multiple Reference Speakers** | H. Hattori<br>S. Nakamura<br>K. Shikano<br>S. Sagayama | Proceedings of 1990 International Conference on Spoken Language Processing,<br>5.6, pp.149-152, Kobe, 1990. | 120 |
| Hirato 90-3 | 種々の発話様式における韻律制御の検討<br>[Prosody Controls for Various Speaking Styles] | 平戸　毅<br>匂坂　芳典<br>[T. Hirato<br>Y. Sagisaka] | ATR Technical Report<br>TR-I-0149 | - |
| Honma 90-3 | 音素環境クラスタリングに基づいた音素単位HMMによる単語認識<br>[Word Recognition by Phoneme HMM Based on Phoneme Environment Clustering] | 本間　茂<br>嵯峨山　茂樹<br><br>[S. Honma<br>S. Sagayama] | 日本音響学会 平成2年春季研究発表会講演論文集, 1-3-13, pp.25-26, 1990.<br>[Proc. of ASJ Fall Meeting, 1-3-13, pp.25-26, 1990] | 124 |
| Huber 90-11 | **Prosodic Transfer in Spoken Language Interpretation** | Dieter Huber | Proceedings of 1990 International Conference on Spoken Language Processing, 12.7, pp.509-512, Kobe, 1990. | 126 |
| Inagaki 90-8 | 品詞情報による音韻継続時間長の分析<br><br>[Analysis of Segmental Duration Using Grammatical Information] | 稲垣　英浩<br>匂坂　芳典<br>海木　延佳<br>[H. Inagaki<br>Y. Sagisaka<br>N. Kaiki] | ATR Technical Report<br>TR-I-172 | - |
| Iwai 90-3 | $F_0$パターンと韻律構造上の『単位』について<br>[A Study on $F_0$ Patterns and Prosodic Units] | 岩井　康雄<br>匂坂　芳典<br>[Y. Iwai<br>Y. Sagisaka] | ATR Technical Report<br>TR-I-160 | - |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Kaiki 90-3 | 統計的手法を用いた文音声における音韻継続時間設定<br><br>[Phoneme Duration Setting in Sentence Utterances using Statistical Method] | 海木　延佳<br>安部　勝雄<br>武田　一哉<br>匂坂　芳典<br>[N. Kaiki<br>K. Abe<br>K. Takeda<br>Y. Sagisaka] | 日本音響学会平成2年度春季研究発表会講演論文集, 1-4-13, pp.203-204, 1990.<br><br>[Proc. of ASJ Spring Meeting, 1-4-13, pp.203-204, 1990] | 130 |
| Kaiki 90-5 | 文音声における音韻継続時間長の設定<br><br>[Phoneme Duration Setting in Sentence Utterances ] | 海木　延佳<br>武田　一哉<br>匂坂　芳典<br>[N. Kaiki<br>K. Takeda<br>Y. Sagisaka] | 電子情報通信学会技術研究報告, SP90-2, pp.9-16, 1990.<br><br>[IECIE Technical Report, SP90-2, pp.9-16, 1990] | 132 |
| Kaiki 90-9a | 文音声における子音継続長の設定<br><br>[Consonant Duration Setting in Sentence Utterances] | 海木　延佳<br>匂坂　芳典<br><br>[N. Kaiki<br>Y. Sagisaka] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 2-6-20, pp.259-260, 1990.<br>[Proc. of ASJ Fall Meeting, 2-6-20, pp.259-260, 1990] | 140 |
| Kaiki 90-9b | **The Control of Segmental Duration in Speech Synthesis Using Linguistic Properties** | N. Kaiki<br>K. Takeda<br>Y. Sagisaka | Proceedings of the ESCA Workshop on Speech Synthesis, pp.165-168, Autrans, France, 1990. | 142 |
| Kaiki 90-11 | **Statistical Analysis for Segmental Duration Rules in Japanese Speech Synthesis** | N. Kaiki<br>K. Takeda<br>Y. Sagisaka | Proceedings of 1990 International Conference on Spoken Language Processing, 1.5, pp.17-20, Kobe, 1990. | 146 |
| Kawabata 89-11a | 構成的ニューラルネットワークによる音声認識<br>[Constructive Neural Network for Speech Recognition] | 川端　　豪<br><br>[T. Kawabata] | ATR Technical Report<br>TR-I-0122 | - |
| Kawabata 89-11b | 音韻モデルと文法を融合した音声認識<br>[Speech Recognition by Combination of Phoneme Models and Grammar] | 川端　　豪<br>北　　研二<br>[T. Kawabata<br>K. Kita] | ATR ジャーナル, No.6, pp.6-9, 1989.<br>[ATR Journal, No.6, pp.6-9, 1989] | 150 |
| Kawabata 90-1 | HMM音韻認識における音節連鎖統計情報の利用<br><br>[HMM Phone Recognition Using Syllable Trigrams] | 川端　　豪<br>花沢　利行<br>伊藤　克亘<br>鹿野　清宏<br>[T. Kawabata<br>T. Hanazawa<br>K. Itoh<br>K. Shikano] | 電子情報通信学会 技術研究報告, SP89-110, pp.7-12, 1989.<br><br>[IECIE Technical Report, SP89-110, pp.7-12, 1989] | 154 |

## List of Technical Publications from Speech Processing Department
(November 1989 through December 1990)
**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Kawabata 90-3 | k-近傍内挿学習による音韻認識<br>[Generalization Effects of k-Neighbor Interpolation Training] | 川端　豪<br>[T. Kawabata] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 2-P-21, pp.161-162, 1990.<br>[Proc. of ASJ Spring Meeting, 2-P-21, pp.161-162, 1990] | 160 |
| Kitagawa 90-1 | HMM 音韻認識における音韻連鎖統計情報の利用<br>[On the Use of Statistical Information of Phoneme Sequences in HMM-based Phoneme Recognition] | 北川　英一郎<br>伊藤　克亘<br>川端　豪<br>鹿野　清宏<br>[E. Kitagawa<br>K. Itoh<br>T. Kawabata<br>K. Shikano] | ATR Technical Report<br>TR-I-0131 | - |
| Komori 89-12 | 音韻認識エキスパートシステムにおける知識とTDNNの融合法<br>[Combining Time Delay Neural Networks (TDNN) and Spectrogram Reading Knowledge into a Phoneme Recognition Expert System] | 小森　康弘<br>畑崎　香一郎<br>川端　豪<br>鹿野　清宏<br>[Y. Komori<br>K. Hatazaki<br>T. Kawabata<br>K. Shikano] | 電子情報通信学会 技術研究報告, SP89-84, pp.63-70. 1989.<br>[IECIE Technical Report, SP89-84, pp.63-70, 1989] | 162 |
| Komori 90-1 | スペクログラムリーディング知識とニューラル·ネットワークを用いた音韻認識エキスパート·システム<br>[Phoneme Recognition Expert System Using Spectrogram Reading Knowledge and Neural Networks] | 小森　康弘<br>畑崎　香一郎<br>川端　豪<br>鹿野　清宏<br>[Y. Komori<br>K. Hatazaki<br>T. Kawabata<br>K. Shikano] | 電子情報通信学会論文誌, D-II, Vol.J73-D-II, No.1, pp.10-18, 1990.<br>[Trans. of IEICE, Vol.J73-D-II, No.1, pp.10-18, 1990] | 170 |
| Komori 90-3a | 音韻認識エキスパート·システムにおける母音認識-TDNNによる母音スポッティング-<br>[Integrating Vowel-spotting TDNN into a Phoneme Recognition Expert System Using Spectrogram Reading Knowledge] | 小森　康弘<br>畑崎香一郎<br>川端　豪<br>鹿野　清宏<br>[Y. Komori<br>K. Hatazaki<br>T. Kawabata<br>K. Shikano] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 2-P-18, pp.155-156, 1990.<br>[Proc. of ASJ Spring Meeting, 2-P-18, pp.155-156, 1990] | 179 |
| Komori 90-3b | 時間構造を考慮したニューラル·ネットワークによる音韻認識<br>[Phoneme Identification Neural Networks Concerning Phonetic Temporal Structure] | 小森　康弘<br>南　泰浩<br>鹿野　清宏<br>[Y. Komori<br>Y. Minami<br>K. Shikano] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 2-P-19, pp.157-158, 1990.<br>[Proc. of ASJ Spring Meeting, 2-P-19, pp.157-158, 1990] | 181 |
| Komori 90-4 | **Combining Phoneme Indentification Neural Networks into an Expert System Using Spectrogram Reading Knowledge** | Y. Komori<br>K. Hatazaki<br>T. Tanaka<br>T. Kawabata | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S10.5, pp.505-508, Albuquerque, 1990. | 183 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Komori 90-11 | **A Fuzzy Training Approach for Phoneme Classification Neural Networks** | Y. Komori<br>S. Sagayama<br>A. Waibel | ATR Technical Report<br>TR-I-190 | - |
| Kurematsu 89-10 | ATRにおける自動翻訳の概要<br><br>[Overview of ATR Researches into Telephone Interpretation] | 榑松　明<br>鹿野　清宏<br>川端　豪<br>飯田　仁<br>森元　逞<br>[A. Kurematsu<br>K. Shikano<br>T. Kawabata<br>H. Iida<br>T. Morimoto] | ATR Technical Report<br>TR-I-184 | - |
| Kurematsu 90-1 | 自動翻訳電話の可能性<br><br>[On the Feasibility of Automatic Interpreting Telephony] | 榑松　明<br><br>[A. Kurematsu] | 日本ビジネスレポート『技術予測シリーズ』pp.165-174, 1990.<br>[Technical Forecast Series, Japan Business Report, pp.165-174, 1989] | 187 |
| Kurematsu 90-3 | **ATR Japanese Speech Database as a Tool of Speech Recognition and Synthesis** | A. Kurematsu<br>K. Takeda<br>Y. Sagisaka<br>S. Katagishi<br>H. Kuwabara<br>K. Shikano | Speech Communication, Vol.9, pp.357-363, 1990.North-Holland. | 197 |
| Kurematsu 90-5 | ATRにおける音声認識の研究<br><br>[Research on Speech Recognition at ATR] | 榑松　明<br><br>[A. Kurematsu] | 東北大学シンポジウム『音声の自動認識の現状と将来』pp.1-7, 1990.<br>[Proceedings of Tohoku University Symposium on Automatic Speech Recognition Present and Future, pp.1-7, 1990] | 204 |
| Kurematsu 90-9 | 音声言語の理解における概念形成の課題<br>[Problems on Concept Formulation in Spoken Language Understanding] | 榑松　明<br><br>[A. Kurematsu] | 電子情報通信学会第二種研究会,「言語獲得.概念形成」LA90-6, pp.1-10, 1990. | 211 |
| Kurematsu 90-10 | 自動翻訳電話--異言語間のコミュニケーションを目指して<br>[Automatic Telephone Interpretation -Towards a Communication Between Different Languages] | 榑松　明<br><br>[A. Kurematsu] | 電子情報通信学会技術研究報告, HC90-15, pp.1-8, 1990.<br><br>[IECIE Technical Report, HC90-15, pp.1-8, 1990] | 221 |
| Kurematsu 90-11 | **A Perspective of Telephone Interpretation Research** | A. Kurematsu | Proceedings of Pacific Rim International Conference on Artificial Intelligence, pp.11-16, 1990. | 229 |

List of Technical Publications from Speech Processing Department
(November 1989 through December 1990)
**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Kuwabara 90-8 | **Voice Quality Control through Vector Quantization** | H. Kuwabara<br>M. Abe | ISSPA(International Symposium on Sinal Processing and Its Applications), 1990 | 235 |
| Maruyama 89-11 | HMM音韻連結学習とNET gramを用いた英単語音声の認識<br><br>[Enlish Word Recognition using HMM Phone Concatenated Traning and Netgram] | 丸山　活輝<br>中村　雅己<br>川端　　豪<br>鹿野　清宏<br>[K. Maruyama<br>M. Nakamura<br>T. Kawabata<br>K, Shikano | ATR Technical Report<br>TR-I-0123 | - |
| Maruyama 89- | **HMM Based Word Recognition using Word Category Prediction Neural Network** | K. Maruyama<br>M. Nakamura<br>T. Kawabata<br>K. Shikano | The Journal of the Acoustical Society of America, S68, CC6, Supplement 1, Vol.86, Fall 1989 | 239 |
| Maruyama 89-12 | HMM音韻連結学習とNET gramを用いた英単語音声の認識<br><br>[English Word Recognition using HMM Phone Concatenated Training and Netgram] | 丸山　活輝<br>中村　雅己<br>川端　　豪<br>鹿野　清宏<br>[K. Maruyama<br>M. Nakamura<br>T. Kawabata<br>K. Shikano | 電子情報通信学会 技術研究報告<br>SP89-89, pp.25-30, 1989<br><br>[IECIE Technical Report, SP89-89, pp.25-30, 1989] | 240 |
| Maruyama 90-3a | NET gram を用いたHMM英単語音声認識の改善<br><br>[Improvement of HMM-based English Word Recognition Using NETgram] | 丸山　活輝<br>中村　雅己<br>川端　　豪<br>鹿野　清宏<br>[K. Maruyama<br>M. Nakamura<br>T. Kawabata<br>K. Shikano | 日本音響学会 平成2年度 春季研究発表会講演論文集, 3-3-7, pp.83-84, 1990.<br><br>[Proc. of ASJ Spring Meeting, 3-3-7, pp.83-84, 1990] | 246 |
| Maruyama 90-3b | NET gram を用いたHMM英単語音声認識の改善<br><br>[Improvement of HMM-based English Word Recognition Using NETgram] | 丸山　活輝<br>中村　雅己<br>川端　　豪<br>鹿野　清宏<br>[K. Maruyama<br>M. Nakamura<br>T. Kawabata<br>K. Shikano] | ATR Technical Report<br>TR-I-0133 | - |
| Matsunaga 90-11a | **Sentence Speech Recognition using Semantic Dependency Analysis** | S. Matsunaga<br>S. Sagayama | Proceedings of 1990 International Conference on Spoken Language Processing , 21.9, pp.929-932, Kobe, 1990. | 248 |
| Matsunaga 90-11b | **A Continuous Speech Recognition System Based on a Two-level Grammar Approach** | S. Matsunaga<br>S. Sagayama<br>S. Honma<br>S. Furui | Proceedings of 1990 International Conference on Spoken Language Processing, S11.7, pp.589-592, Kobe, 1990. | 252 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Minami 90-1 | TDNN音韻スポッティングと予測LRパーザを用いた大語彙単語音声認識<br>[Large Vocabulary Spoken Word Recognition Using Time-Delay Neural Network Phoneme Spotting and Predictive LR-Parsing] | 南 泰浩<br>沢井 秀文<br>宮武 正典<br>[Y. Minami<br>H. Sawai<br>M. Miyatake] | 電子情報通信学会 技術研究報告 SP89-99, pp.7-13, 1990.<br><br>[IECIE Technical Report, SP89-99, pp.7-13, 1990] | 256 |
| Minami 90-2a | TDNN音韻スポッテイングと予測LRパーザを用いた大語彙単語音声認識<br><br>[Large Vocabulary Spoken Word Recognition Using Time-Delay Neural Network Phoneme Spotting and Predictive LR-Parsing] | 南 泰浩<br>沢井 秀文<br>宮武 正典<br>鹿野 清宏<br>[Y. Minami<br>H. Sawai<br>M. Miyatake<br>K. Shikano] | ATR Technical Report TR-I-0144 | - |
| Minami 90-2b | TDNNの構造の音韻認識率、シフトインバリアント性への影響<br>[The Effect of TDNN Structures on Phoneme Recognition Rates and Shift Invariance] | 南 泰浩<br>沢井 秀文<br><br>[Y. Minami<br>H. Sawai] | ATR Technical Report TR-I-0145 | - |
| Minami 90-3 | 入力層.中間層におけるベクトルの近傍の情報を利用したTDNN出力の平滑化<br><br>[Output Smoothing for TDNN Using Information on Input Vectors Neighborhood Input Layer] | 南 泰浩<br>田村 震一<br>沢井 秀文<br>鹿野 清宏<br>[Y. Minami<br>S. Tamura<br>H. Sawai<br>K. Shikano] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 1-3-18, pp.35-36, 1990.<br><br>[Proc. of ASJ Spring Meeting, 1-3-18, pp.35-36, 1990] | 263 |
| Minami 90-6 | 時間遅れ神経回路網(TDNN)による音韻スポッティング法と予測LRパーザを用いた大語彙単語音声認識<br>[Large Vocabulary Spoken Word Recognition Using Time-Delay Neural Network Phoneme Spotting and Predictive LR-Parsing] | 南 泰浩<br>沢井 秀文<br>宮武 正典<br><br>[Y. Minami<br>H. Sawai<br>M. Miyatake] | 電子情報通信学会論文誌D-II, Vol.J73-D-II, No.6, PP.788-795, 1990.<br><br>[Trans. of IEICE, Vol.J73-D-II, No.6, PP.788-795, 1990] | 265 |
| Minami 90-10 | 発話変動にロバストなTDNNの検討<br>[A Study on Robust TDNNs against the Variability of Speaking Styles] | 南 泰浩<br>沢井 秀文<br>[Y. Minami<br>H. Sawai] | ATR Technical Report TR-I-183 | - |
| Mimura 90-9 | 統計的手法を用いた文音声における振幅制御<br><br>[Power Control in Sentence Utterances Using Statistical Method] | 三村 克彦<br>海木 延佳<br>匂坂 芳典<br>[K. Mimura<br>Y. Kaiki<br>Y. Sagisaka] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 2-6-19, pp.257-258, 1990.<br>[Proc. of ASJ Fall Meeting, 2-6-19, pp.257-258, 1990] | 273 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Miyatake 90-4 | **Integrated Training for Spotting Japanese Phonemes Using Large Phonemic Time-Delay Neural Networks** | M. Miyatake<br>H. Sawai<br>Y. Minami<br>K. Shikano | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S8.10, pp.449-452, Albuquerque, 1990. | 275 |
| Miyatake 90-5 | 時間遅れ神経回路網(TDNN)による音韻スポッティングのための学習法とその効果<br>[Training Methods and Their Effects for Spotting Japanese Phonemes Using Time-Delay Neural Networks] | 宮武 正典<br>沢井 秀文<br>鹿野 清宏<br>[M. Miyatake<br>H. Sawai<br>K. Shikano] | 電子情報通信学会 論文誌, D-II, Vol.J73-D-II ,No,5 pp.699-706, 1990.<br>[Trans. of IEICE, Vol.J73-D-II , No.5, PP.699-706, 1990) | 279 |
| Miyatake 90-12 | 種々の発話様式に見られる韻律特徴とその制御<br>[Prosodic Characteristics and Their Control in Japanese Speech with Various Speaking Styles] | 宮武 正典<br>匂坂 芳典<br>[M. Miyatake<br>Y. Sagisaka] | 電子情報通信学会論文誌, D-II , Vol.J73-D-II No.12, pp.1929-1935, 1990.<br>[Trans. of IEICE, Vol.J73-D-II , No.12, pp.1929-1935, 1990] | 287 |
| Nagai 90-9 | HMM-LR 法における音素文脈依存型LRパーザの検討<br>[Study on Phonemic Context Dependent LR Parser in HMM-LR Method] | 永井 明人<br>北 研二<br>嵯峨山 茂樹<br>[A. Nagai<br>K. Kita<br>S. Sagayama] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 3-8-16, pp.125-126, 1990.<br>[Proc. of ASJ Fall Meeting, 3-8-16, pp.125-126, 1990] | 294 |
| Nakamura. M 89-10 | ニューラルネットにおけるバックプロパゲーション学習の効率化方法<br>[A New Method to Speed up the Back-Propagation Algorithm] | 中村 雅己<br>鹿野 清宏<br>[M. Nakamura<br>K. Shikano] | ATR Technical Report TR-I-0119 | - |
| Nakamura. M 90-3 | ニューラルネットによる音韻フィルター<br>[A Study of Phoneme Filter Using Neural Network] | 中村 雅己<br>田村 震一<br>[M. Nakamura<br>S. Tamura] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 2-P-24, pp.167-168, 1990.<br>[Proc. of ASJ Spring Meeting, 2-P-24, pp.167-168,1990] | 296 |
| Nakamura. M 90-6 | ニューラルネットによる音素フィルタを用いた母音認識<br>[Vowel Recognition by Phoneme Filter Newral Networks] | 中村 雅己<br>田村 震一<br>[M. Nakamura<br>S. Tamura] | 電子情報通信学会技術研究報告, SP90-11, pp.17-23, 1990<br>[IECIE Technical Report , SP90-11, 1990] | 298 |
| Nakamura. M 90-8 | **Neural Network Approach to Word Category Prediction for English Texts** | M. Nakamura<br>K. Maruyama<br>T. Kawabata<br>K. Shikano | COLING'90, Computational Linguistics, Vol.3, pp.213-218, Finland, 1990. | 305 |
| Nakamura. M 90-9a | ニューラルネットによる英単語品詞列予測モデル<br>[English Word Category Prediction -Based on Neural Network] | 中村 雅己<br>鹿野 清宏<br>[M. Nakamura<br>K. Shikano] | ATR Technical Report TR-I-176 | - |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Nakamura. M 90-9b | ニューラルネットによる音素フィルタを用いた母音認識<br>[Vowel Recognition by Phoneme Filter Neural Networks] | 中村　雅己<br>田村　震一<br>[M. Nakamura<br>S. Tamura] | ATR Technical Report<br>TR-I-177 | - |
| Nakamura. M 90-11 | **Vowel Recognition by Phoneme Filter Neural Networks** | M. Nakamura<br>S. Tamura | Proceedings of 1990 International Conference on Spoken Language Processing 16.3, pp.669-672, Kobe,1990. | 311 |
| Nakamura. M 90-9 | 時間遅れ神経回路網を用いた不特定話者の音韻認識<br>[Speaker -independent Phoneme Recognition Using Time-Delay Neural Networks] | 中村　悟<br>沢井　秀文<br>[S. Nakamura<br>H. Sawai] | ATR Technical Report<br>TR-I-178 | - |
| Nakamura. M 90-12 | 不特定話者音素認識のためのニューラルネットアーキテクチャの検討<br>[A Preliminary Study on Neural Network Architectures for Speaker-Independent Phoneme Recognition] | 中村　悟<br>沢井　秀文<br>[S. Nakamura<br>H. Sawai] | 電子情報通信学会技術研究報告<br>SP.90-61, pp.33-40, 1990.<br>[IECIE Technical Report SP.90-61, pp.33-40, 1990] | 315 |
| Nakamura. S 90-3 | ベクトル量子化話者適応化のTDNN音韻認識への適用<br>[VQ-based Speaker Adaptation Applied to TDNN Phoneme Recognition] | 中村　哲<br>鹿野　清宏<br>[S. Nakamura<br>K. Shikano] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 2-P-22, pp.163-164, 1990.<br>[Proc. of ASJ Spring Meeting, 2-P-22, pp.163-164, 1990] | 323 |
| Nakamura. S 90-4 | **A Comparative Study of Spectral Mapping for Speaker Adaptation** | S. Nakamura<br>K. Shikano | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S3.7, pp.157-160, Albuquerque, 1990. | 325 |
| Nakamura. S 90-12 | 話者重畳型HMMを用いた日本語音韻認識における話者適応化の改善<br>[Improved Speaker Adaptation Using Speaker Supplemented HMM for Japanese Phoneme Recognition] | 中村　哲<br>服部　浩明<br>鹿野　清宏<br>[S. Nakamura<br>H. Hattori<br>K. Shikano] | 電子情報通信学会論文誌, D-II, Vol.J73-D-II, No.12, pp.1919-1928, 1990.<br>[Trans. of the IEICE, D-II, Vol.J73-D-II, No.12, pp.1919-1928, 1990] | 329 |
| Ohkura 90-6 | コードブックマッピングを用いた雑音環境下での音声認識<br>[Speech Recognition of Under Noisy Environment by Codebook Mapping] | 大倉　計美<br>服部　浩明<br>杉山　雅英<br>鹿野　清宏<br>[K. Ohkura<br>H. Hattori<br>M. Sugiyam<br>K. Shikano] | 電子情報通信学会技術研究報告<br>SP90-12, pp 25-32, 1990.<br>[IECIE Technical Report, SP90-12, pp.25-32, 1990.] | 339 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Ohkura 90-9 | 波形入出力による雑音抑圧ニューラルネットの音声認識への応用<br>[The Apprication of Noise Reduction Nural Network to Speech Recognition] | 大倉　計美<br>杉山　雅英<br>[K. Ohkura<br>M. Sugiyama] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 1-8-3, pp.5-6, 1990.<br>[Proc. of ASJ Fall Meeting, 1-8-3, pp.5-6, 1990] | 347 |
| Rainton 90-9 | **Speech Analysis / Synthesis using a Partial outer Product Expansion of the Wigner distribution** | David Rainton | 日本音響学会 平成2年度秋季研究発表会講演論文集, 1-8-1, pp.1-2, 1990.<br>[Proc. of ASJ Fall Meeting, 1-8-1, pp.1-2, 1990] | 349 |
| Rainton 90-10 | **Speech Analysis and Enhancement using the Time-frequency Wigner distribution** | David Rainton | 電子情報通信学会技術研究報告, SP90-49, pp.39-46, 1990.<br>[IECIE Technical Report, SP90-49, pp.39-46, 1990] | 351 |
| Rainton 90-11 | **Time-Frequency Spectral Analysis of Speech** | David Rainton<br>S.J. Young | Proceedings of 1990 International Conference on Spoken Language Processing , 9.1, pp.349-352, 1990. | 359 |
| Sadanobu 90-4 | 合成的字音語のアクセントの字数<br>[Accent of Sino-Japanese Complex Word and the Number of Charcters] | 定延　利之<br>匂坂　芳典<br>[T. Sadanobu<br>Y. Sagisaka] | ATR Technical Report TR-I-158 | - |
| Sagayama 90-4 | 音素環境クラスタリングのHMM大語彙音声認識への適用<br>[Phoneme Environment Clustering Applied to Large Vocabulary Speech Recognition] | 嵯峨山　茂樹<br>[S. Sagayama] | 文部省科研費総合研究(A)‖マルコフモデル·ニューラルネットワークを包含する新しい音声認識手法の総合的研究‖第5回研究討論会 榊原温泉 H2.4.20<br>[5th meeting of the commitee of New Approaches for Combining Markov Models and Neural Networks, Sakakibara-onsen, April, 1990] | 363 |
| Sagayama 90-7 | 音声認識におけるいくつかの基本問題について<br>[Basic Problems in Speech Recognition ] | 嵯峨山　茂樹<br>[S. Sagayama] | 文部省科研費総合研究(A)‖マルコフモデル·ニューラルネットワークを包含する新しい音声認識手法の総合的研究‖シンポジウム(湯河原) H2.7.20<br>[Symposium on New Approaches for Combining Markov Models and Neural Networks, Yugawara-onsen, July, 1990] | 367 |
| Sagayama 90-11 | **Estimation of Unknown Context Using a Phoneme Environment Clustering Algorithm** | S. Sagayama<br>S. Honma | Proceedings of 1990 International Conference on Spoken Language Processing , 9.4, pp.361-364, 1990. | 371 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Sagayama 90-12 | 行列演算によるHMM音声認識アルゴリズムの表現について [A Matrix Representation of HMM-Based Speech Recognition Algorithms] | 嵯峨山 茂樹 [S. Sagayama] | 電子情報通信学会技術研究報告, SP90-65, pp.63-70, 1990. [IECIE Technical Report, SP90-65, pp.63-70, 1990 | 375 |
| Sagisaka 89-11a | **On the Unit Set Design for Speech Synthesis by Rule using Non-uniform Units** | Y. Sagisaka | The Journal of the Acoustical Society of America, Supplement, Vol.86 , S79, FF24, Fall, 1989. | 383 |
| Sagisaka 89-11b | 『音声研究』日本音響学会座談会 [Meeting on "Speech Research"] | 匂坂 芳典,他 [Y. Sagisaka] | 日本音響学会 学会誌, 45巻, 11月号, pp.861-871, 1989. [The Journal of the Acoustical Society of Japan, Vol.45, 11, pp.861-871, 1989] | 384 |
| Sagisaka 90-1 | $F_0$パタン概形制御の定量的検討 [On the Quantification of Global $F_0$ Pattern Control] | 匂坂 芳典 [Y. Sagisaka] | 電子情報通信学会 技術研究報告, SP89-111, pp.13-18, 1990. [IECIE Technical Report, SP89-111, pp.13-18, 1990] | 395 |
| Sagisaka 90-3 | **Speech Synthesis from Text** | Y. Sagisaka | IEEE Communication Magazine , Special Issue, pp.35-41, 1990 . | 401 |
| Sagisaka 90-4a | **On the Prediction of Global $F_0$ Shape for Japanese Text-to-Speech** | Y. Sagisaka | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S6a.9, pp.325-328, Albuquerque, 1990. | 409 |
| Sagisaka 90-4b | 音声合成の立場からみた音声処理単位 [Segment Unit for Speech Synthesis] | 匂坂 芳典 [Y. Sagisaka] | 文部省科研費総合研究(A)「マルコフモデル・ニューラルネットワークを包含する新しい音声認識手法の総合的研究」第5回研究討論会 榊原温泉 H2.4.20 [Symposium on New Approaches for Combining Markov Models and Neural Networks, Sakakibara-onsen, April, 1990] | 413 |
| Sagisaka 90- | Phonotactic Constraintsを用いた音声単位の構成法 [The design of Synthesis unit set Using Phonotactic Constraints] | 匂坂 芳典 [Y. Sagisaka] | 近畿音声言語研究会資料 [KINKI Society for Phonetics] | 415 |
| Sagisaka 90-9a | 統語構造に基づく$F_0$上昇現象の統計的分析 [Statistic Analysis on $F_0$ Boosting in Relation to Syntactic Structure] | 匂坂 芳典 [Y. Sagisaka] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 2-6-15, pp.249-250, 1990. [Proc. of ASJ Fall Meeting, 2-6-15, pp.249-250, 1990] | 422 |

List of Technical Publications from Speech Processing Department
(November 1989 through December 1990)
**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Sagisaka 90-9b | **Units and Rules for Speech Synthesis** | Y. Sagisaka | Proceedings of the Tutorial Day Workshop on Speech Synthesis, pp.1-10, Autrans, France, 1990. | 424 |
| Sagisaka 90-11 | **A Large-scale Japanese Speech Database** | Y. Sagisaka<br>K. Takeda<br>M. Abe<br>S. Katagishi<br>T. Umeda<br>H. Kuwabara | Proceedings of 1990 International Conference on Spoken Language Processing, 24.4, pp.1089-1092, Kobe, 1990. | 433 |
| Sawai 90-9a | TDNN-LR文節音声認識システムにおける追加学習の効果<br>[Effects of Incremental Training in the TDNN-LR Phrase Speech Recognition System] | 沢井　秀文<br>[H. Sawai] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 2-P-11, pp.151-152, 1990.<br>[Proc. of ASJ Fall Meeting, 2-P-11, pp.151-152, 1990] | 437 |
| Sawai 90-9b | 時間·周波数変動に強い時間遅れ神経回路網(TDNN)<br>[Time-Frequency Shift-Tolerant Time -Delay Neural Networks] | 沢井　秀文<br>[H. Sawai] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 2-P-12, pp.153-154, 1990.<br>[Proc. of ASJ Fall Meeting, 2-P-12, pp.153-154, 1990] | 439 |
| Sawai 90-10 | 時間遅れ神経回路網(TDNN)による音声認識<br>[Speech Recognition Using Time-Delay Neural Networks] | 沢井　秀文<br>[H. Sawai] | 電気関係学会関西支部連合大会予稿集, S14-2<br>[Proceedings of United Meeting of Kansai District in Electronics - related Societys, S14-2] | 441 |
| Sawai 90-11 | **The TDNN-LR Large-vocabulary and Continuous Speech Recognition System** | H. Sawai | Proceedings of 1990 International Conference on Spoken Language Processing, 31.4, pp.1349-1352, Kobe, 1990. | 442 |
| Shikano 89-10 | **Research Activities of the Speech Processing Department** | 鹿野　清宏<br>[K. Shikano] | ATR Technical Report TR-I-0115 | - |
| Shikano 89-12 | ニューラルネットワークによる音声認識<br>[Speech Recognition Using Neural Networks] | 鹿野　清宏<br>[K. Shikano] | コロナ社 コンピュートロール, No.29, pp.42-51, 1989.<br>[Computrol, No. 29, pp.42-51, 1989] | 446 |
| Shikano 90-5 | **Approaches to Continuous Speech Recognition Using Time-Delay Neural Networks and Learning Vector Quantization** | K. Shikano | 日独情報技術フォーラム通産省、西ドイツ, 1990.5<br>[Deutsch-Japanisches Forum Informationstechnologie, Mai 1990] | 456 |
| Shimodaira 90-12 | ピッチパタン連続整合による連続音声のセグメンテーション<br>[Phrase Segmentation of Continuous Speech] | 下平　博<br>嵯峨山　茂樹<br>木村　正行<br>[H. Shimodaira<br>S. Sagayama<br>M. Kimura] | 電子情報通信学会技術研究報告, SP90-72, pp.33-40, 1990.<br>[IECIE Technical Report, SP90-72, pp.33-40, 1990] | 461 |

List of Technical Publications from Speech Processing Department
(November 1989 through December 1990)
**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Sugiyama 90-5 | 歪み尺度測地線を用いた音声スペクトルの補間<br><br>[Spectral Interpolation using distortion Geodesic Lines] | 杉山　雅英<br><br>[M. Sugiyama] | 電子情報通信学会技術研究報告, SP90-3, pp.17-24, 1990.<br><br>[IECIE Technical Report, SP-90-3, pp.17-24, 1990] | 469 |
| Sugiyama 90-8 | **Automatic Language Recognition using Acoustic Features** | M. Sugiyama | ATR Technical Report TR-I-167 | - |
| Sugiyama 90-9 | ニューラルネットによる集合間写像の教師なし学習<br><br>[Unsupervised Training Methods for Set Mappings Using Neural Networks] | 杉山　雅英<br>福沢　圭二<br>沢井　秀文<br>嵯峨山　茂樹<br>[M. Sugiyama<br>K. Fukuzawa<br>H. Sawai<br>S. Sagayama] | 日本音響学会 平成2年度秋季研究発表会講演論文集, 2-P-10, pp.149-150, 1990.<br><br>[Proc. of ASJ Fall Meeting, 2-P-10, pp.149-150, 1990] | 477 |
| Sugiyama 90-10a | ATRにおけるNeural Network を用いた音声情報処理<br>[Neural Networks Applied to Speech Processing in ATR] | 杉山　雅英<br><br>[M. Sugiyama] | ATR Technical Report TR-I-173 | - |
| Sugiyama 90-10b | ニューラル ネットワーク を用いた音声情報処理<br><br>[Speech Information Processing Using Neural Networks] | 杉山　雅英<br><br>[M. Sugiyama] | 神経情報科学特選講座 第9回「ニューロモデルの新しい流れ」pp.1-16, 1990.<br>[Draft for Seminar of Neural Network and Information Science , New Trends of Neural Network Models, pp.1-16, 1990] | 479 |
| Sugiyama 90-11 | **Spectral Interpolation Using Distortion Geodesic Lines** | M. Sugiyama | Proceedings of 1990 International Conference on Spoken Language Processing , 11.15, pp.477-480, Kobe, 1990. | 503 |
| Takahashi 90-11 | **Isolated Word Recognition Using Pitch Pattern Information** | S. Takahashi<br>S. Matsunaga<br>S. Sagayama | Proceedings of 1990 International Conference on Spoken Language Processing , 13.9, pp.553-556. Kobe, 1990. | 507 |
| Takeda 90-1 | エキスパートシステムを用いた単位選択の検討<br><br>[A Rule-based Approach for Synthesis Unit Selection] | 武田　一哉<br>匂坂　芳典<br>安部　勝雄<br>[K. Takeda<br>Y. Sagisaka<br>M. Abe] | 電子情報通信学会 技術研究報告 SP89-113, pp.27-32, 1990.<br><br>[IECIE Technical Report, SP89-113, pp.27-32, 1990] | 511 |

# List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Takeda 90-2a | 複合音声単位を用いる規則合成実験システム<br>[Speech Synthesis System Using Non-Uniform Units] | 武田 一哉<br>安部 勝雄<br>匂坂 芳典<br>海木 延佳<br>[K. Takeda<br>M. Abe<br>Y. Sagisaka<br>N. Kaiki] | ATR Technical Report<br>TR-I-0140 | - |
| Takeda 90-2b | 種々の音韻連接単位を用いる規則合成方式の診断的な評価<br>[Evaluation and Diagnosis of Selective Use of Non-uniform Units for Speech Synthesis-by-rule] | 武田 一哉<br>阿部 匡伸<br>匂坂 芳典<br>[K. Takeda<br>K. Abe<br>Y. Sagisaka] | ATR Technical Report<br>TR-I-0142 | - |
| Takeda 90-2c | 大規模音声データベースに基づく音声合成<br>[Automatic Rule Derivation for Speech Synthesis from Large-scaled Speech Database] | 武田 一哉<br>[K. Takeda] | ATR Technical Report<br>TR-I-0143 | - |
| Takeda 90-3 | 合成単位の音響的属性と合成音との関係<br>[On the Acoustic-phonetic Properties of Synthesis Units and the Resulting Synthesized Speech] | 武田 一哉<br>安部 勝雄<br>匂坂 芳典<br>[K. Takeda<br>K. Abe<br>Y. Sagisaka] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 1-4-14, pp.205-206, 1990.<br>[Proc. of ASJ Spring Meeting, 1-4-14, pp.205-206, 1990] | 517 |
| Takeda 90-9 | **On Unit Selection Algorithms and Their Evaluation in Non-uniform Speech Synthesis** | K. Takeda<br>K. Abe<br>Y. Sagisaka | Proceedings of the Tutorial and Research Workshop on Speech Synthesis, pp.35-38, France, 1990. | 519 |
| Takeda 90-11 | **On the Unit Search Criteria and Algorithms for Speech Synthesis using Non-uniform Units** | K. Takeda<br>K. Abe<br>Y. Sagisaka | Proceedings of 1990 International Conference on Spoken Language Processing , 8.8, pp.341-344. Kobe, 1990. | 523 |
| Takeda 90-12 | 選択的に合成単位を用いる規則音声合成<br>[Speech Synthesis by Rule Based on Adaptive Unit Selection] | 武田 一哉<br>安部 勝雄<br>匂坂 芳典<br>[K. Takeda<br>K. Abe<br>Y. Sagisaka] | 電子情報通信学会論文誌, Vol. J73-D-II , No.12 pp.1945-1951, 1990.<br>[Trans. of the IEICE , Vol. J73-D-II , No.12 pp.1945-1951, 1990] | 527 |
| Takami 90-6 | 対判定型TDNNによる音素認識<br>[Phoneme Recognition by Pairwise Discriminant TDNN] | 鷹見 淳一<br>嵯峨山 茂樹<br>[J. Takami<br>S. Sagayama] | 電子情報通信学会技術研究報告 SP90-10, pp.9-16, 1990.<br>[IECIE Technical Report, SP90-10, pp.9-16, 1990] | 534 |

List of Technical Publications from Speech Processing Department
(November 1989 through December 1990)
**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Takami 90-9 | 対判定型TDNNにおける中間値学習の効果<br>[Effects of the Training with Middle Value on Pairwise Discriminant TDNN] | 鷹見 淳一<br>嵯峨山 茂樹<br>[J. Takami<br>S. Sagayama] | 日本音響学会 平成2年秋季研究発表会講演論文集, 2-P-15, pp.159-160, 1990.<br>[Proc. of ASJ Fall Meeting, 1990] | 542 |
| Takami 90-11 | **Phoneme Recognition by Pairwise Discriminant TDNNs** | J. Takami<br>S. Sagayama | Proceedings of 1990 International Conference on Spoken Language Processing , 16.5, pp.677-680, Kobe, 1990. | 544 |
| Tamura 89-10 | フィードフォワードニューラルネットワークの解釈<br>[On Interpretations of a Feedforward Neural Network] | 田村 震一<br>[S. Tamura] | ATR Technical Report TR-I-0116 | - |
| Tamura 90-2 | 波形入出力による雑音抑圧ニューラルネットワークの最適化<br>[Improvements to the Noise Reduction Neural Network] | 田村 震一<br>中村 雅己<br>[S. Tamura<br>M. Nakamura] | 電子情報通信学会技術研究報告, SP89-120, pp.9-14, 1990.<br>[IECIE Technical Report SP89-120, pp.9-14, 1990] | 548 |
| Tamura 90-3 | 波形入出力による雑音抑圧ニューラルネットワークの改良<br>[Improvements to the Noise Reduction Neural Network] | 田村 震一<br>中村 雅己<br>[S. Tamura<br>M. Nakamura] | 日本音響学会 平成2年度 春季研究発表会講演論文集, 3-4-18, pp.303-304, 1990.<br>[Proc. of ASJ Spring Meeting, 3-4-18, pp.303-304, 1990] | 554 |
| Tamura 90-4 | **Improvements to the Noise Reduction Neural Network** | S. Tamura<br>M. Nakamura | Proceedings of 1990 International Conference on Acoustics, Speech, and Signal Processing, S15b.5, pp.825-828, Albuquerque, 1990. | 556 |
| Umeda 90-8 | 声質変換技術と高品質ピッチ変換法<br>[Voice Conversion Techniques and a High-quality Pitch Conversion Algorithm] | 梅田 哲夫<br>[T. Umeda] | ATR Technical Report TR-I-175 | - |
| Yamashita 90-11 | **A Support Environment Based on Rule Interpreter for Synthesis by Rule** | Y. Yamashita<br>H. Fujiwara<br>Y. Nomura<br>N. Kaiki<br>R. Mizoguchi | Proceedings of 1990 International Conference on Spoken Language Processing, 19.2, pp.769-772, Kobe, 1990. | 560 |
| Yoshida 90-5 | 母音無声化の要因分析<br>[Factor Analysis of Vowel devoicing in Japanese] | 吉田 夏也<br>匂坂 芳典<br>[N. Yoshida<br>Y. Sagisaka] | ATR Technical Report TR-I-159 | - |
| Waibel 89-11 | **Connectionist Large Vocabulary Word Recognition** | Alex Waibel | ATR Technical Report TR-I-0120 | - |

List of Technical Publications from Speech Processing Department

(November 1989 through December 1990)

**(Bold titles indicate papers written in English)**

| Ref.ID | Title(題名) | Authors(著者) | Journal(掲載誌) | Page |
|---|---|---|---|---|
| Waibel 89-12 | **Modularity and Scaling in Large Phonemic Neural Network** | Alex Waibel<br>H. Sawai<br>K. Shikano | IEEE Transactions on Acoustics,Speech,and Signal Processing, Vol.37, No.12, pp.1888-1898, 1989. | 564 |