Internal Use Only (Do not copy)

TR-I-0010

Research Activities of Speech Processing Department

April 1986 through October 1987

ATR Interpreting Telephony Research Laboratories

1987.12.1

ATR Interpreting Telephony Research Laboratories

Research Activities of

Speech Processing Department

April 1986 through October 1987

Abstract:

The research areas of the Speech Processing Department now include:

(A) Phoneme Recognition and Segmentation.

(1) Feature-Based Approaches.

(2) Hidden Markov Models.

(3) Neural Network Approaches.

(B) Speaker Adaptation.

(1) Speaker Adaptation by Vector Quantization.

(C) Language Models.

(1) Word Trigram Models.

(2) Word Prediction by Neural Networks.

(D) Noise Reduction.

(1) Noise Reduction by Neural Networks.

(E) Speech Synthesis by Rule.

(1) High Quality Japanese Speech Synthesis by Rule.

(2) Modal Information and Prosodies.

(F) Voice Conversion.

(1) Voice Conversion by Parameter Mapping through Vector Quantization.

(2) Personal Characteristics Analysis by Analysis-by-Synthesis.

(G) Phonetics

(1) Analysis of Allophones and Coarticulation.

(H) Speech Database

(1) Japanese Large-Scale Speech Database with Phonetic Transcriptions.

(2) Speech Database Retrieval System.

(3) Speech Workbench.

For additional information, contact:

ATR Interpreting Telephony Research Laboratories

Twin 21, MID-Tower, 2-1-61 Shiromi, Higashi-ku, Osaka, 540, JAPAN

Telephone: +81−6−949−1830, Telefax: +81−6−949−1822,

Telex: 64660 ATR J,

E-mail: atr-la.atr.junet ! shikano @ uunet.UU.NET (from USA arpanet).

Contents

1. General

Necessity of automatic telephone interpretation research is described in brief. We introduce the Speech Processing Department at the ATR Interpreting Telephony Research Laboratories, and provide an overvview of its research activities from April, 1986 to October, 1987.

1.1. Automatic Telephone Interpretation Project

The Automatic Telephone Interpretation system is a facility which enables a person speaking in one language to communicate readily by telephone with someone speaking another language. It does so by automatically and simultaneously transforming the dialogue from the speaker's to the listener's language. At least three constituent technologies are necessary for such a system: speech recognition, machine translation and speech synthesis.

Since this is a new concept, a number of studies and evaluations on applicabilities must still be made. A high degree of performance from each of the constituent technologies and user friendliness of the system are also essential. According to a feasibility study report published by the Japanese Ministry of Posts and Telecommunications, realizing a system like this will require at least fifteen years. One more important requirement of this ambitious project is the international cooperation of research institutions in Japan and abroad.

1.2. Speech Processing Department

The ATR Interpreting Telephony Research Laboratories, whose president is Dr. Akira Kurematsu, have three departments. They are the Natural Language Understanding Department, the Knowledge and Database Department and the Speech Processing Department. These laboratories were founded in April of 1986 with the support of the Japan Key Technology Center, ATR International, NTT, KDD, NHK and other Japanese enterprises. The Speech Processing Department is one of these and is headed by Dr. Kiyohiro Shikano. The research areas of the Speech Processing Department are concerned with speech recognition , speech synthesis and speech databases. These research areas are aimed at advanced speech technologies capable of recognizing continuous speech from any speaker and synthesizing high quality speech with various kinds of speaker characteristics. These technologies, which are indispensable to realize an interpreting telephony system, have to be improved considerably by investigating these fields and exploring inovative techniques and break-through technologies.

The Speech Processing Department has two major groups. One is a speech recognition group and the other is a speech synthesis and phonetics group. The speech recognition group is mainly directed by Drs. K. Shikano, A. Waibel and T. Kawabata. The speech synthesis and phonetics group is mainly directed by Drs. H. Kuwabara and Y. Sagisaka.

As of October, 1987, efforts aimed at speaker-dependent phoneme recognition and speaker-independent phoneme segmentation have resulted in dramatically improved phoneme recognition performance. We are now pursuing three approaches. They are (1) *Feature-Based* approach especially for phoneme segmentation, (2) *Hidden Markov Model* approach, and (3) *Neural Network* approach. Word spotting efforts have also started. For speaker-independent speech recognition, a speaker adaptation approach has been undertaken using a concept of *Vector Quantization and Spectrum Mapping*. In the area of language models, a *Trigram* approach has been taken, and a word prediction approach utilizing Neural Networks has been initiated.

Analysis efforts for the exploration of *Allophones and Coarticulation* phenomena have been undertaken utilizing a large-scale speech database. In order to realize a high quality rule-based speech synthesis system, we are working in three areas, such as a speech synthesis by rule based on *Multiple Speech Units* from a phoneme-labelled speech database, a voice conversion algorithm from one speaker to another based on *Vector Quantization and Mapping*, and a relation analysis between *Modal information and Prosodics*. Moreover, non-parametric noise reduction algorithms using *Neural Networks* are now being investigated.

A large-scale speech database with phonetic transcriptions, mainly for research in speech recognition and speech synthesis, has been developed with much support of other speech research laboratories. A speech workbench and a database retrieval system have been implemented on a micro Vax station in order to provide researchers with a tool for easy access and manipulation of speech data.

2. Research Activities

Research activities in the Speech Processing Department from April, 1986 through October, 1987 are summalized below, and the related technical publications are quoted.

2.1. Phoneme Recognition and Segmentation

Reliable phoneme recognition and segmentation algorithms have been studied leading to considerable improvements beyond conventional approaches. We have been taking three approaches, a *Feature-Based* approach, a *Hidden Markov Model* approach, and a *Neural Network* approach.

2.1.1. Feature-Based Approach [Hatazaki, Tamura, Kawabata, Murayama]

A spectrogram reading seminar taught by Prof. Victor Zue was held in February of 1986. Since that time, spectrogram reading session meet on a weekly basis. Many researchers in our department can now read phonemes in Japanese utterances with relatively high accuracy.

In order to clarify the differences of spectrogram reading features between English and Japanese, statistical analysis has been carried out based on the large vocabulary speech database with phonetic transcriptions [Kawabata-87-10].

In particular, phoneme boundaries can be identified correctly in Japanese utterances. Now we are developing an expert system for phoneme segmentation using the expert tool, *ART* [Hatazaki-87-10]. A Symbolics lisp machine for running *ART* and a micro vax station for feature extraction of input speech are connected to each other and used to implement a phoneme segmentation expert system.

Technical Publications: [Hatazaki-87-10], [Kawabata-87-10]

2.1.2. Hidden Markov Models [Hanazawa, Kawabata, Shikano]

Hidden Markov models have been studied to know how to use models properly in the stage of training based on the forward-backward algorithm [Hanazawa-87-10]. The number of states, tied probabilities of transition and output, parameter smoothing techniques, and initial probability settings have been studied so far based on the task of phoneme recognition in our large vocabulary speech database described in section **2.8.1**. The Hidden Markov models based on phoneme units have now also been applied successfully to word spotting in continuous speech.

Technical Publications: [Hanazawa-87-10], [Results will be presented in ICASSP'88.]

2.1.3. Neural Network Approaches [Waibel]

Neural Networks have been investigated to recognize phonemes in continuous utterances since May of 1987. A time-delay neural network (TDNN) now achieves high phoneme recognition rates for the task of speaker-dependent discrimination among the voiced consonants, /b/,/d/, and /g/. All consonants were extracted from the phonetically labelled large vocabulary database , i.e., from 5240 common Japanese words spoken by three speakers. The TDNN attains 98.5% phoneme recognition rates, compared to a rate of 93.7% for the Hidden Markov model. Closer inspection revealed that the network invented well-known acoustic features such as F2-rise ,F2-fall and vowel-onset as useful abstraction. The back-propagation algorithm was also coded on an Alliant 9400 with four CPU units for a twenty fold speed-up compared to a Vax 8600. Further improvement are underway targetting a factor of 100 speed-up over aVax 8600.

Technical Publications: [The TDNN will be presented in ICASSP'88, and ATR technical report of the TDNN is available on request.]

2.2. Speaker Adaptation

Aiming at a general preprocessor for speaker normalization, speaker adaptation research has been studied using vector quantization techniques. The discrete spectrum space representation by vector quantization makes it possible to realize sophisticated speaker adaptation / normalization.

2.2.1. Speaker Adaptation by Vector Quantization[S. Nakamura, Shikano]

As a general preprocessor to a phoneme recognizer, speaker normalization algorithms have been developed. The algorithms adopt vector quantization as a discrete representation of spectral space. The discrete representation makes it possible to carry out sophisticated spectral normalization or mapping from one speaker to another. In previous studies[Shikano-86-12-b], algorithms using a single codebook were developed, and a complex spectral distortion measure which is composed of a spectrum term, a differenced spectrum term and a differenced power term was adopted. The complex measure improved the word recognition results, but it resulted in significant degradation of the spectrogram. In order to reduce the degradation of the spectrogram, algorithms with multi-codebooks have been investigated[Nakamura-87-6][Nakamura-87-10]. These algorithms are called "speaker adaptation by separate vector quantization". So far, these algorithms attain good performance in experiments evaluating the spectrogram distortion and in word recognition experiments. They are successfully applied to voice conversion as described in section **2.6** [Abe,M-87-10]. Fuzzy vector quantization techniques are now being introduced to realize more accurate speaker adaptation.

Technical Publications: [Shikano-86-12-b], [Nakamura-87-6][Nakamura-87-10], [Abe,M-87-10], [Speaker adaptation and voice conversion will be presented in ICASSP'88]

2.3. Language Models

ATR Interpreting Telephony Laboratories have two other departments, which are studying language models more deeply than the Speech Processing Department. Nevertheless, the Speech Processing Department has been studying language models based on bottom-up word prediction and statistical word prediction. Now, we have been taking two approaches, a word trigram one and a neural network one , to predict a word set to be recognized next. Such bottom-up word prediction approaches should be combined with top-down approaches in language processing.

2.3.1. Word Trigram Models [Shikano]

Trigram language models based on word categories were introduced in order to improve word recognition results for English sentences uttered word by word. Probabilities of the trigrams of word categories were estimated using the Brown Corpus Text Database. Moreover, high probabilities of common word sequences such as frozen word sequences were extracted from the Brown Corpus Text Database. These probabilities were integrated using a dynamic programming algorithm, and improved the word recognition results from 80.9% to 89.1% [Shikano-87-3][Shikano-87-4][Shikano-87-6].

Technical Publications: [Shikano-87-3][Shikano-87-4][Shikano-87-6]

2.3.2. Word Prediction by Neural Networks [M. Nakamura, Shikano]

A neural network approach is being developed to predict words using the Brown Corpus Text Database. The results will be compared to results from trigram modeling. In order to realize this approach practically and quickly, further improvements of our learning network simulator are underway.

2.4. Noise Reduction

In order to bring speech recognition technologies into the field, reduction of various noises has to be performed. We are now looking for new noise reduction mechanisms using various kinds of neural networks.

2.4.1. Noise Reduction by Neural Networks [Tamura, Waibel]

Noise reduction algorithms using neural networks have been under investigation since August of 1987 using the back-propagation algorithm on multi-layered feed-forward and recurrent networks. During training, the networks develop their internal representation for a mapping from a set of noisy signals to the set of noise-free signals with noise-added speech signals as inputs and noise-free speech signals as the target outputs. The training is carried out using the digitized speech wave itself. The noise reduction algorithms are already working effectively for white noise. Effectiveness for non-stationary noises and speech noises are currently being studied.

Technical Publications: [The algorithms will be presented in ICASSP'88]

2.5. 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 is under development. Also, rules between speaking styles and prosodic control have been studied.

2.5.1. High Quality Japanese Speech Synthesis by Rule [Sagisaka, K. Abe]

In speech synthesis by rule, various kinds of synthesis units such as phonemes, dyads, CV-syllables and CVC units have been proposed, and their relative merits and drawbacks have been documented. Our proposed synthesis scheme is different from ordinary systems in two ways. One is the flexible use of non-uniform synthesis units and the other is a choice of optimal units of multi-templates using criteria such as feasibility of unit concatenation and fitness of units to target context [Sagisaka-87-10]. The strength of coarticulation for CV syllables are also investigated in depth. The place of articulation is much more dominant than the manner of articulation from the viewpoints of coarticulation phenomena [Abe,K-87-10].

Technical Publications: [Sagisaka-87-10][Abe,K-87-10][An overview of the system will be presented in ICASSP'88]

2.5.2. Modal Information and Prosodics [Miyatake, Sagisaka]

Effects of speaking styles on prosodic parameters have been analyzed to clarify the interrelation among pitch patterns, power patterns and durations. The analysis shows that differences among speaking styles are strongly reflected in the pitch and power patterns. The results suggest the possibility of a new prosody control model for speech synthesis [Miyatake-87-9][Miyatake-87-10-a][Miyatake-87-10-b].

Technical Publications: [Miyatake-87-9][Miyatake-87-10-a][Miyatake-87-10-b]

2.6. Voice Conversion

Voice conversion from one speaker to another is an important aspect for realizing the automatic telephone interpreting system. Voice conversion algorithms based on vector quantization and spectrum mapping have been developed. We have also started to analyze personal voice characteristics using analysis-by-synthesis methods.

2.6.1. Voice Conversion by Parameter Mapping through Vector Quantization [M. Abe, S. Nakamura, Kuwabara, Shikano]

A new voice conversion technique through vector quantization has been studied [Abe,M-87-10]. The algorithms used in this voice conversion is basically the same as the spectrum mapping described in section **2.2.** The algorithms have three codebooks, which for spectrum, power and pitch. The voice conversion from male voice to female voice was tried and a hearing test shows that the converted voice was judged perfectly as female. In order to improve the converted voice, vector quantization of residual waves and use of fuzzy vector quantization techniques are now being explored.

Technical Publications: [Abe,M-87-10] [will be presented in ICASSP'88]

2.6.2. Personal Characteristics Analysis by Analysis-by-Synthesis [Segot, Kuwabara]

Aiming to improve the quality of speech, the modification of formant frequencies and bandwidths has been studied. These modification techniques will be applied to analyze personal voice characteristics. Attention is now focused on the dynamic aspects of spectral parameters that convey personal voice characteristics.

Technical Publications: [Kuwabara-86-11] [Kuwabara-86-12] [Kuwabara-87-3] [Kuwabara-87-4-a] [Kuwabara-87-4-b]

2.7. Phonetics

Analysis of speech phenomena has been started using our large vocabulary speech database in order to clarify allophonic and coarticulatory phenomena.

2.7.1. Analysis of Allophones and Coarticulation [Takeda, Kuwabara]

As a first step to analyze allophonic and coarticulatory phenomena, devocalization of vowels was studied using our large vocabulary speech database. The contexts of the devocalization are predicted by rules [Takeda-87-10]. Next, phoneme durations are now under investigation using our large speech database.

Technical Publications: [Takeda-87-10]

2.8. Speech Database

Speech database with phoneme labels are essential and necessary not only for speech recognition research but also for speech synthesis research.

2.8.1. Japanese Large-Scale Speech Database with Phonetic transcriptions [Takeda, Kuwabara, Katagiri, Sagisaka, M. Abe]

A large-scale Japanese speech database has been constructed at ATR with the help of several other laboratories. For multiple transcriptions, three types of categories are considered: phonetic labels, acoustic events and allophonic variations. To date, about 8500 words uttered by seven professional announcers have been collected and transcribed [Takeda-87-8] [Sagisaka-87-5] [Takeda-87-3] [Takeda-87-6]. The speech database has been effectively used to perform speech research at ATR.

Technical Publications: [Takeda-87-8] [Sagisaka-87-5] [Takeda-87-3] [Takeda-87-6]

2.8.2. Speech Database Retrieval System [Takeda, Kuwabara]

To access the speech database easily, a database management system has been developed [Takeda-87-8] [Takeda-87-9]. The system is implemented using the relational database *UNIFY*. The easy access language, *EAL*, is implemented using the host language interface on *UNIFY*. *EAL* is now widely used at ATR.

Technical Publications: [Takeda-87-8] [Takeda-87-9]

2.8.3. Speech Workbench [Maruyama, Murayama, Kawabata]

A speech workbench for speech researchers has been developed on a micro Vax station. The main features are high quality spectrograms, easy extension of functions and easy access to the speech database.

3. Research Staff

The research staff is mainly composed of members from the research institute and laboratories which support ATR and visiting scientists. The 17 members presently as of October, 1987, employed in the speech processing department are listed below:

Dr. Kiyohiro Shikano,	*Department Head*	*(NTT,*	*1986.6 —)*
Dr. Hisao Kuwabara,	*Supervisor*	*(NHK,*	*1986.10—)*
Dr. Yoshinori Sagisaka,	*Senior Researcher*	*(NTT,*	*1986.4 —)*
Dr. Takeshi Kawabata,	*Senior Researcher*	*(NTT,*	*1986.9 —)*
Dr. Alex Waibel,	*Invited Researcher*	*(Carnegie-Mellon Univ.,*	*1987.5 —)*
Shin'ichi Tamura,	*Researcher*	*(Sony,*	*1986.9 —)*
Kaichiro Hatazaki,	*Researcher*	*(NEC,*	*1986.12—)*
Masanori Miyatake,	*Researcher*	*(Sanyo,*	*1986.9 —)*
Masami Nakamura,	*Researcher*	*(Sumitomo Metal,*	*1987.9 —)*
Katsuo Abe,	*Researcher*	*(Toyocom,*	*1987.3 —)*
Satoshi Nakamura,	*Researcher*	*(Sharp,*	*1986.9 —)*
Masanobu Abe,	*Researcher*	*(NTT,*	*1987.4 —)*
Kazuya Takeda,	*Researcher*	*(KDD,*	*1986.8 —)*
Katsuteru Maruyama,	*Researcher*	*(Nitsuko,*	*1987.3 —)*
Toshiyuki Hanazawa,	*Researcher*	*(Mitsubishi Elec.,*	*1987.3 —)*
Kooichi Murayama,	*Engineer*	*(DEC Japan,*	*1987.9 —)*
Hubert Segot,	*intern student*	*(ENST, France,*	*1987.4 —)*

4. Research Facilities

The Speech Processing Department facilities include a number of computer systems which are connected through Ethernet. The computer systems include a Vax 8600, a Vax 8700, and an Alliant 9400 (four-CPU array processor). Every researcher also has a micro Vax station with an AD/DA device. Four Masscomps (MC5600) and one micro Vax II which are installed with 16 bit AD conversion devices are used to collect speech data. One S3670 (Symbolics) with *ART* is used for implementing an expert system.

Technical Publication List in Speech Processing Department
(April, 1986 through October, 1987)
(Bold letter titles show papers written in English or with an English abstract.)

Ref. ID	Title	Authors	Conference / Journal	Page
Shikano-86-9	カーネギー・メロン大学における音声認識・理解研究の現状 (Present State of Research Activities of Automatic Speech Recognition and Understanding at Carnegie-Mellon University)	Seiichi Nakagawa (Toyohashi Univ. of technology), Kiyohiro Shikano	Jonal of Acoustic Society of Japan, (1986-09)	15 ~ 19
Shikano-86-12-a	音声理解研究の動向 (Trends on Speech Understanding Researches)	Kiyohiro Shikano, Akira Kurematsu	Jonal of Acoustic Society of Japan, (1986-12)	20 ~ 24
Shikano-87-4	**Improvement of Word Recognition Results by Trigram Models**	Kiyohiro Shikano	Proceedings on ICASSP 86, 29.2, (1987-04)	25 ~ 28
Sagisaka-86-9	日本語テキストからの音声合成とアクセント (Speech Synthesis by Rule and Prosodies)	Yoshinori Sagisaka, Hirokazu Sato (NTT Electrical Communication Laboratories)	Phonetics Meeting in Kansai Area, (1986-09)	29 ~ 50
Shikano-86-12-b	ベクトル量子化による話者適応 **(Speaker Adaptation through Vector Quantization**)	Kiyohiro Shikano	IEICE Technical Report, SP86-65, (1986-12)	51 ~ 61
Kuwabara-86-11	音声の声質変換における信号処理 (Speech Processing in Speech Quality Change)	Hisao Kuwabara, Tohru Takagi (NHK Science and Technical Laboratories)	The First Digital Signal Processing Synposium, B2.3, (1986-11)	62 ~ 68

Technical Publication List in Speech Processing Department
(April, 1986 through October, 1987)

(Bold letter titles show papers written in English or with an English abstract.)

Ref. ID	Title	Authors	Conference / Journal	Page
Kuwabara-86-12	分析合成による声質変換と嗄声音改善への応用 **(Speech Quality Control by the Analysis-Synthesis Method and Application to the Enhancement of Abnormal Speech)**	Hisao Kuwabara, Tohru Takagi (NHK Science and Technical Laboratories)	IEICE Technical Report, SP86-57 / H-86-56, (1986-12)	69 ~ 77
Kawabata-87-4	**Word Spotting Method Based on Top-Down Phoneme Verification**	Takeshi Kawabata, Masaki Kohda (NTT Electrical Communication Laboratories)	Proceedings on ICASSP 86, 34.7, (1987-04)	78 ~ 81
Takeda-87-8	**Acoustic-Phonetic Labels in a Japanese Speech Database**	Kazuya Takeda, Yoshinori Sagisaka, Shigeru Katagiri	European Conference on Speech technology, pp13-16, (1987-08)	82 ~ 85
Sagisaka-87-5	**Phonetic Labeling and Acoustic Correlates for Building Japanese Speech Data Base**	Yoshinori Sagisaka, Shigeru Katagiri ,Kazuya Takeda	ASA-Meeting '87, Spring , (1987-05) including its handout.	86 ~ 91
Shikano-87-3	**Improvement of Word Recognition Results by Trigram Models**	Kiyohiro Shikano	ASJ-Meeting '87 Spring, 3-5-1 (1987-03)	92 ~ 93
Takeda-87-3	音声データベース構築のための音韻ラベリング (Acoustiic-Phonetic Labeling in a Japanese Speech Database)	Kazuya Takeda, Yoshinori Sagisaka	ASJ-Meeting '87 Spring, 2-5-10 (1987-03)	94 ~ 95

Technical Publication List in Speech Processing Department
(April, 1986 through October, 1987)
(Bold letter titles show papers written in English or with an English abstract.)

Ref. ID	Title	Authors	Conference / Journal	Page
Kuwabara-87-3	声の明瞭性を支配する物理的特徴量の抽出 (Extraction of Acoustic Features Specifying the Clearness of Natural Speech)	Tohru Takagi (NHK Science and Technical Laboratories), Hisao Kuwabara	ASJ-Meeting '87 Spring, 2-6-15 (1987-03)	96 ~ 97
Kuwabara-87-4-a	声の個人性に関する諸問題 (Some Problems on the Personal Charactteristics of Speech)	Hisao Kuwabara	Journal of IECEJ, Vol. 70, No. 4, (1987-04)	98 ~ 105
Kuwabara-87-8	**Quality Control of Speech by Modifying Formant Frequencies and Bandwidths**	Hisao Kuwabara, Tohru Takagi (NHK Science and Technical Laboratories)	the 11th International Congress of Phonetic Sciences, Tallin Estonia, (1987-08)	106 ~ 109
Kurematsu-87-4	自動翻訳電話のための研究課題ー音声を主体としてー (Research for Realization of Interpreting Telephone System)	Akira Kurematsu	IEICE Technical Report, Invited Talk, (1987-04)	110 ~ 117
Kuwabara-87-4-b	日本語音声データベースとそのラベリング (Japanese Speech Database and Its Phoneme Labels)	Hisao Kuwabara, Yoshinori Sagisaka, Kazuya Takeda, Shigeru Katagiri	Phonetics Meeting in Kansai Area, (1987-04)	118 ~ 123
Shikano-87-6	**Improvement of Word Recognition Results by Trigram Models**	Kiyohiro Shikano	IEICE Technical Report, SP87-23, (1987-06)	124 ~ 132

Technical Publication List in Speech Processing Department
(April, 1986 through October, 1987)
(Bold letter titles show papers written in English or with an English abstract.)

Ref. ID	Title	Authors	Conference / Journal	Page
Takeda-87-6	音韻ラベルを持つ日本語音声データベースの構築 **(Construction of an Acoustically-Phonetically Transcribed Japanese Speech Database)**	Kazuya Takeda, Hisao Kuwabara, Yoshinori Sagisaka, Shigeru Katagiri	IEICE Technical Report, SP87-19, (1987-06)	133 ~ 141
Nakamura-87-6	ベクトル量子化を用いたスペクトログラムの正規化 **(Spectrogram Normalization Based on Vector Quantization)**	Satoshi Nakamura, Kiyohiro Shikano	IEICE Technical Report, SP87-17, (1987-06)	142 ~ 150
Hatazaki-87-9	連続音声中の音韻認識エキスパートシステムの検討 (An Expert System for Phoneme Recognition in Continuous Speech)	Kaichiro Hatazaki, Shin'ichi Tamura, Takeshi Kawabata, Kiyohiro Shikano	IPS Fall Meeting, 1L-4, (1987-09)	151 ~ 152
Takeda-87-9	音声データベース管理システムの構築 (Construction of a Japanese Speech Database Management System)	Kazuya Takeda, Hisao Kuwabara, Shogo Morikawa (TIS)	IPS Fall Meeting, 2H-4, (1987-09)	153 ~ 154
Miyatake-87-9	会話文音声合成のための音声合成 (Speech Synthesis by Rule for Conversation)	Masanori Miyatake, Yoshinori Sagisaka	IPS Fall Meeting, 2H-3, (1987-09)	155 ~ 156
Takeda-87-10	母音無声化の要因分析と予測手法の検討 (Analysis and Prediction of Devocalized Phenomena)	Kazuya Takeda, Hisao Kuwabara	ASJ Fall Meeting, 3-3-8, (1987-10)	157 ~ 158

Technical Publication List in Speech Processing Department (April, 1986 through October, 1987)

(Bold letter titles show papers written in English or with an English abstract.)

Ref. ID	Title	Authors	Conference / Journal	Page
Nakamura-87-10	ベクトル量子化を用いたスペクトログラムの正規化 (Spectrogram Normalization Using Vector Quantization)	Satoshi Nakamura, Kiyohiro Shikano	ASJ Fall Meeting, 3-3-8, (1987-10)	159 ~ 160
Sagisaka-87-10	種々の複合音声単位からの音声合成 (Speech Synthesis by Rule Using Non-Uniform Phonemic Clusters)	Yoshinori Sagisaka	ASJ Fall Meeting, 3-3-8, (1987-10)	161 ~ 162
Kawabata-87-10	日本語スペクトログラム特徴の英語との比較 (Sound Spectrogram Features in Japanese)	Takeshi Kawabata, Shin'ichi Tamura, Kaichiro Hatazaki	ASJ Fall Meeting, 3-3-8, (1987-10)	163 ~ 164
Miyatake-87-10-a	発声様式の違いが韻律パラメータに与える影響の分析 (Acoustic Manifestation of Prosody Control)	Masanori Miyatake, Yoshinori Sagisaka	ASJ Fall Meeting, 3-3-8, (1987-10)	165 ~ 166
Abe,M-87-10	ベクトル量子化による音質変換 (Voice Conversion through Vector Quantization)	Masanobu Abe, Satoshi Nakamura, Kiyohiro Shikano, Hisao Kuwabara	ASJ Fall Meeting, 3-3-8, (1987-10)	167 ~ 168
Hanazawa-87-10	Hidden Markovモデルを用いた日本語有声破裂音の識別 (Discrimination of Japanese Voiced Stops Using Hidden Markov Models)	Toshiyuki hanazawa, Takeshi Kawabata, Kiyohiro Shikano	ASJ Fall Meeting, 3-3-8, (1987-10)	169 ~ 170
Abe,K-87-10	抽出環境が音節スペクトルに与える影響の分析 (Spectrum Analysis by Syllable Environments)	Katsuo Abe, Yoshinori Sagisaka	ASJ Fall Meeting, 3-3-8, (1987-10)	171 ~ 172

Technical Publication List in Speech Processing Department
(April, 1986 through October, 1987)
(Bold letter titles show papers written in English or with an English abstract.)

Ref. ID	Title	Authors	Conference / Journal	Page
Miyatake-87-10-b	種々の発声様式における韻律パラメータの性質について **(On the Prosodic Characteristics in Various Utterances with Different Speaking Styles)**	Masanori Miyatake, Yoshinori Sagisaka	IEICE Technical Report, SP87-62, (1987-10)	173 ~ 181