

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 17 日現在

機関番号：53203

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23501127

研究課題名(和文) マルチメディア技術による中国語の発音教育システムの開発

研究課題名(英文) Development of training system of the Chinese pronunciation by multimedia technology

研究代表者

星野 朱美 (HOSHINO, AKEMI)

富山高等専門学校・一般教養科・准教授

研究者番号：90300566

交付決定額(研究期間全体)：(直接経費) 2,500,000円、(間接経費) 750,000円

研究成果の概要(和文)：日本人学生が中国語を学習する時「そり舌音」と「舌歯音」は調音の区別が最も難しい。そこで学生の自習のため、コンピューターを用いて、自分で発音のチェックができるように自動判別システム(CAI)を開発した。中国語話者40名の中国語の有気音の「そり舌音」と有気音の「舌歯音」の発話のスペクトルをMATLABで35チャンネルのフィルターバンクを形成して、VOT中のパワーと有声期間中のパワーの特徴を抽出し、自動判別システムを開発した。そのシステムを用いて50名の学生の発話を自動判別し、中国語話者の発話評価を基準にして、正確率は「そり舌音」と「舌歯音」がそれぞれ91%と95%以上の高い値を得た。

研究成果の概要(英文)：Chinese retroflex aspirates are generally difficult for Japanese students learning Chinese pronunciation. In particular, discrimination between utterances of aspirated dental and retroflex affricates is the most difficult for Japanese-speakers to learn. For the students' self-learning, the automatic judgment system has been developed using a computer, so as to check their pronunciations by themselves. We extracted the features of correctly pronounced aspirated dental affricates and aspirated retroflex affricates by observing the spectrum evolution of breathing power during both voice onset time and voiced period of sounds uttered by 40 Chinese native speakers. We developed a 35-channel computerized filter bank to analyze the evolution of the breathing power by using MATLAB and then automatically evaluated the utterances of the 50 Japanese students. The success rates of the system were higher than 91% and 95% for aspirated retroflex and dental affricates, respectively.

研究分野：総合領域

科研費の分科・細目：科学教育・教育工学・教育工学

キーワード：有気音のそり舌音 有気音の舌歯音 35CHのFBの作製 パワーの自動計算 VOTの自動測定 判定基準の作成 発話の自動判別

1. 研究開始当初の背景

(1) 地球上の3人中1人が中国語を話すと言われている。中国以外の国で中国語が話される国と地方は、香港、台湾、シンガポール、また、東南アジアには華僑も多い。それらの国との経済や文化の交流と共に、我が国においても英語の次に中国語を学習する学生が多くなっている。

(2) しかし、中国語の発音の殆どが日本語に無い音なので、多くの学生は中国語の発音を難しく感じている。その中でも「そり舌音」と「舌歯音」は調音の区別が困難で、特に有気音の「そり舌音」を発話するときには舌をそらせて息を吐きながら調音するが、日本語の中に「そり舌音」が無いから、日本人学生が「そり舌音」を発話するつもりが、舌をそる程度がたりないので、舌歯音になってしまう。

(3) それらの発音を正しく聞き分けることもできない。また、教師は授業時に一对一での発音矯正訓練は、多くの学生を対象にした場合には時間的にも不可能である。特に帰宅後の自習も自分の発音に対して正確な評価手段はなく、自分の発音が正確かどうか、チェックする手段もない。

(4) 現在のところ、それらの発音訓練用の教材(CAI)で市販されているものはない。

2. 研究の目的

(1) 本研究では、中国語話者20名と日本語話者40名の中国語の有気音の「そり舌音」 $cha[t\zeta'a]$ 、 $chi[t\zeta'i]$ 、 $che[t\zeta'ɤ]$ 、 $chu[t\zeta'u]$ と有気音の「舌歯音」 $ca[t\zeta'a]$ 、 $ci[t\zeta'i]$ と $ce[t\zeta'ɤ]$ 、 $cu[t\zeta'u]$ の発話スペクトルを詳細に解析することにより、正確な発話の特徴やパターンを抽出した。次にそれらのパターンを新たな発話の評価基準として用い、「そり舌音」と「舌歯音」の音節の自動判別システムを開発する。

(2) これまでに確立した中国語の有気音の発話評価基準を用いて、これらを取り入れた音声解析機能を持ち、学生の発話の認識精度が高く、人に優しいマンマシンインターフェイスを持った教材の開発に取り組む。すなわち、マルチメディア技術を用いた中国語発音教育のコンピュータ援助指導システム(CAI)を開発する。これにより、パソコンの画面上で自分の発音が正しいかどうか正確に判断しながら、中国語の発音訓練のサポートが受けられるようになる。

3. 研究の方法

(1) 有気音の「そり舌音」と有気音の「舌歯音」の相違点

Fig.1は中国語の有気音の「そり舌音」 $cha[t\zeta'a]$ と有気音の「舌歯音」 $ca[t\zeta'a]$ のスペクトログラムを示す。図の左側は有気音の「そり舌音」 $cha[t\zeta'a]$ で、そり舌摩擦音とも言う。舌をそらせて後部歯茎にあてて閉鎖を作ったのち、開放すると同時に口を大きく開けて、破裂音を生成する。それと同時に同じ調音点

で舌と後部歯茎の間の隙間から生じる息の摩擦音も発生し、「そり舌音」 $cha[t\zeta'a]$ が生成される。ただし、舌をしっかりとそらさないと舌歯音になってしまう。日本語にはそり舌音がないので、日本人がそり舌音を発話するのはかなり難しい。この図のVOTの長さは150msである。VOTが0~70ms間で、周波数が2500~5000Hz(左上)に濃い黒い縦縞が見られる。これは破裂しながら舌と後部歯茎の間に生じる息の摩擦によって、口中に強いエネルギーを発生したもので、この部分のパワーが高くなっている。右側は「舌歯音」 $ca[t\zeta'a]$ で、VOTの長さは160ms、VOTが70ms~150ms間で、2500~5000Hzの周波数に薄く黒い縦縞が見られ、パワーがやや強いことが見て取れる。

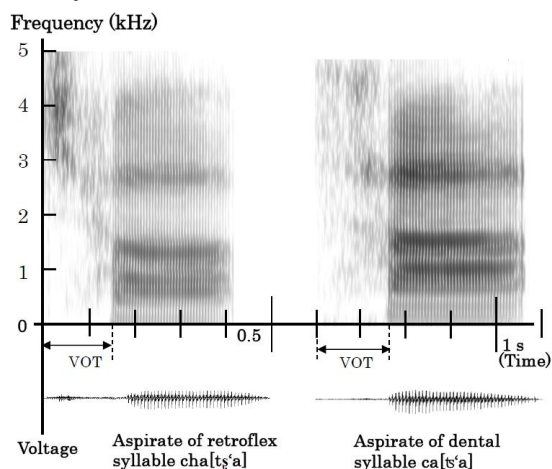


Fig. 1. Spectrograms of retroflex aspirate affricate $cha[t\zeta'a]$ (left) and dental aspirated affricate $ca[t\zeta'a]$ (right) pronounced by Chinese speaker

Fig.2は中国人の発話による有気音の「そり舌音」 $chi[t\zeta'i]$ と有気音の「舌歯音」 $ci[t\zeta'i]$ のスペクトログラムを示している。図の左側は有気音の「そり舌音」 $chi[t\zeta'i]$ で、VOT250ms中で、周波数区間2000Hz~5000Hzに黒い縦縞が見られる。図の右側は有気音の「舌歯音」 $ci[t\zeta'i]$ で、VOTの長さは225msである。パワーはVOT中ではあまり変化していないことが読み取れる。

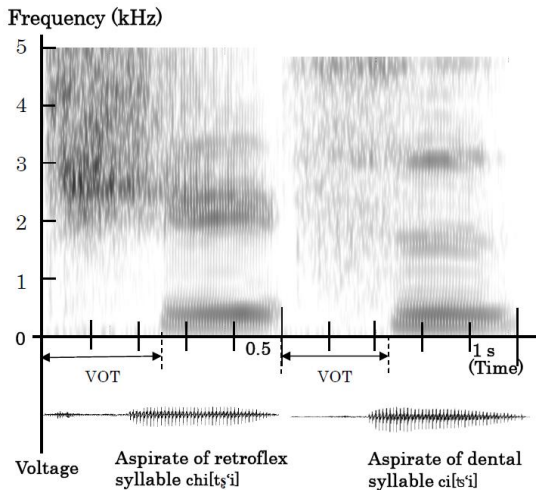


Fig. 2. Spectrograms of retroflex aspirate affricate $chi[t\zeta'i]$ (left) and dental aspirated affricate $ci[t\zeta'i]$ (right) pronounced by Chinese speaker

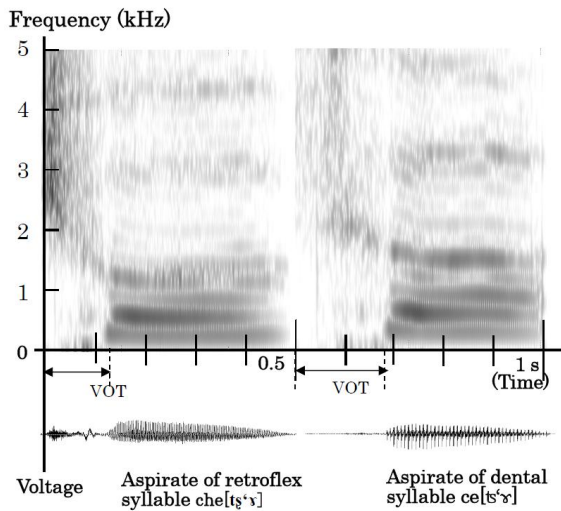


Fig. 3. Spectrograms of retroflex aspirate affricate che[tʂʰ] (left) and dental aspirated affricate ce[tʂʰ] (right) pronounced by Chinese speaker

Fig.3 は中国語の有気音のそり舌音 che[tʂʰ] と有気音の舌歯音 ce[tʂʰ] のスペクトログラムを示している。図の左側は有気音のそり舌音 che[tʂʰ] で、VOT は 125ms である。VOT の前半部分で、周波数 1300Hz~5000Hz に黒い縦縞が見られる。また、周波数 1200Hz 以下では縦縞は薄く、パワーが弱いことが分かる。図の右側は有気音の舌歯音 ce[tʂʰ] で、VOT の長さは 180ms、VOT の中間付近で、周波数約 2000Hz あたりの縦縞はすこし黒く、やや強いパワーが見られる。

Fig4 は中国語の有気音のそり舌音 chu[tʂʰu] と有気音の舌歯音 cu[tʂʰu] のスペクトログラムを示している。図の左側は有気音のそり舌音 chu[tʂʰu] で、VOT は 190ms である。VOT が 0~160ms 間で、周波数 3500Hz~5000Hz に黒い縦縞が見られる。図の右側は有気音の舌歯音 cu[tʂʰu] で、VOT は 185ms である。VOT 中 80ms~120ms で、周波数 0~300 Hz あたりの縦縞はすこし黒く、それ以外は、あまり変化していないことが読み取れる。

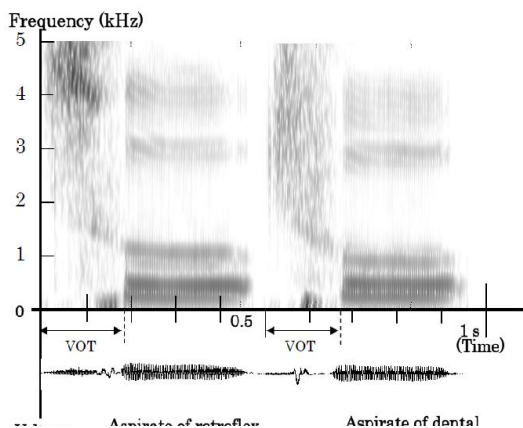


Fig. 4 Spectrograms of retroflex aspirate affricate chu[tʂʰu] (left) and dental aspirated affricate cu[tʂʰu] (right) pronounced by Chinese speaker

(2) VOT とパワーの自動測定アルゴリズム

① VOT の自動測定アルゴリズム

有気音の「そり舌音」と「舌歯音」の発話の分析により発話の正確さは VOT 中の音声周波数の高さ とパワーの強さに密接な関係があることが分かった。

本自動判別システムでは中心周波数 50Hz~6850Hz、帯域幅 200Hz の 35 チャンネル (CH) のフィルターバンク (FB) を形成する。入力された音声信号を 11.025kHz で標本化し、FB に通し、CH 毎に音声パワーの時間変化を求め、それぞれの CH でフレーム間のパワーの差分を計算し、その差が正の値となる CH 数が最も多い時点をパワーの変化のピークとし、破裂点 t1 とする。一方、母音の声帯振動が始まる時、フォルマントが現れる。この時点を VOT の終了時点 t2 とすると、VOT は

$$VOT = t_2 - t_1 \quad (1)$$

となる。

② パワーの自動測定アルゴリズム

これまでに筆者らは VOT および VOT 中の平均パワーの自動測定システムを開発したが [2,3]、今回は「そり舌音」と「舌歯音」の特徴抽出のため、35 CH の FB により、VOT 中および有声期間中の周波数スペクトルの自動測定を行った。

$$P_i = \sum_{j=1}^J P_{i,j}(t_j) \quad (2)$$

この値に 5ms を掛けて、i-CH のエネルギー W_i ($i=0\sim34$) が式(3)のように得られる。

$$W_{i,VOT} = P_i \times 5ms \quad (3)$$

VOT 期間中の CH 毎の平均パワー $P_{i,av}$ は

$$P_{i,av} = W_{i,VOT} / VOT \quad (4)$$

となる。有声期間中の平均パワーも同様に定義できる。

(3) 「そり舌音」と「舌歯音」の発話の正確さと VOT 中の周波数とパワーの関係

本研究では、有気音の舌歯音 ca[tʂʰa], ci[tʂʰi], ce[tʂʰe], chu[tʂʰu] と有気音のそり舌音 cha[tʂʰa], chi[tʂʰi], che[tʂʰe], cu[tʂʰu] の発話の正確さを調査するため、まず、20 名の学生の発話を中国語話者 8 人に聞かせて、正確な有気音の舌歯音とそり舌音に聞こえる場合は「3」を、どちらともはっきりと区別できない場合は「2」を、有気音の「そり舌音」なのに有気音の「舌歯音」に聞こえた場合は「1」をつけ、また、有気音の「舌歯音」なのに有気音の「そり舌音」に聞こえた場合にも「1」をつけた。

有気音の「そり舌音」と「舌歯音」の発話の分析により発話の正確さは VOT 中の音声周波数の高さ とパワーの強さに密接な関係があることが分かった。そこで、VOT とパワーの自動計算システムを用い、これにより学生の発話の VOT とその VOT 中のそれぞれの周波数チャンネルのパワーの平均値を求め、

パワー及び周波数と発話の正確さの関係を詳しく分析した。

Fig.5 に学生の「舌歯音」ca[ts'a]と「そり舌音」cha[tɕ'a]の発話評価を VOT(横軸)とパワー(縦軸)平面図上に示す。この図の各発話のパワーは VOT の開始時間から VOT/2 の時間までの周波数 2450Hz (13CH) から 6850Hz (35CH) までの間のパワーを自動計算したものである。「そり舌音」cha[tɕ'a]のデータは図の中央より右上に集まっている。「舌歯音」ca[ts'a]のデータは図の中央によりやや下に集まっている。有気音の「そり舌音」cha[tɕ'a]の発話はパワーが高く、パワーが17以上の発話は2.6以上の合格の評価が得られている。舌をうまくそらせなかった、すなわち、舌をそる力が足りなかったものは、パワーも弱く、評価も悪い。「舌歯音」ca[ts'a]の発話で、8~12までのパワーの発話は2.8以上の高い評価が得られた。ca[ts'a]の上から二つの発話は発話するとき舌をそる必要がないのに、舌を少しそらしてしまったので、パワーが高くなって、発話評価が悪くなったものと思われる。

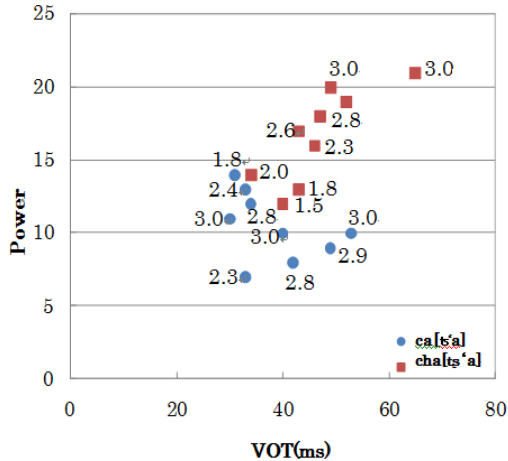


Fig5 Data distribution and scores for retroflex aspirated syllable cha[tɕ'a], and dental aspirated syllable ca[ts'a] with VOT on the abscissa and P_{av} on the ordinate

ci[ts'i] - chi[tɕ'i], ce[ts'ɤ] - che[tɕ'ɤ], cu[ts'u] - chu[tɕ'u]の対についても同様の手法で評価を行い、同様の傾向が得られた。

4. 研究成果

(1) 自動判別用のパラメータ

本研究では、中国語話者40名の有気音の「そり舌音」cha[tɕ'a], chi[tɕ'i], che[tɕ'ɤ] chu[tɕ'u]と有気音の「舌歯音」ca[ts'a], ci[ts'i], ce[ts'ɤ], cu[ts'u]の発話を対象にして、今回開発した発話のパワースペクトルの自動測定システムを用いて、発話のスペクトルを詳細に解析し、35CH毎にそれぞれVOT中のパワーと有声期間中のパワーを自動的に計算することにより、VOT中と有声期間中にそれぞれの「そり舌音」と「舌歯音」の特徴的なパターンを抽出した。

Table I: Evaluation criteria on utterance of retroflex aspirated affricates during VOT

Syllable	Channels (CH)	Frequency domain (Hz)	VOT range	Ave. Power in VOT
cha[tɕ'a]	CH15~CH29	2750~5750	0~VOT/2	17 or more
chi[tɕ'i]	CH10~CH32	1750~6350	Whole VOT	25 or more
che[tɕ'ɤ]	CH07~CH30	1150~5950	0~VOT*2/3	34 or more
chu[tɕ'u]	CH17~CH31	3250~6150	0~VOT*2/3	11 or more

Table II: Evaluation criteria for formants of retroflex affricates during voiced period after VOT

Syllable	F1 (Hz)/(CH)	F2 (Hz)/(CH)	F3 (Hz)/(CH)
cha[tɕ'a]	550~750/(CH4)	950~1150/(CH6)	1750~1950/(CH10)
chi[tɕ'i]	150~350/(CH2)	1950~2150/(CH11)	2750~2950/(CH15)
che[tɕ'ɤ]	350~550/(CH3)	950~1150/(CH6)	2350~2550/(CH14)
chu[tɕ'u]	150~350/(CH2)	750~950/(CH5)	2750~2950/(CH15)

Table III: Evaluation criteria for formants of dental affricates during voiced period after VOT

Syllable	F1 (Hz)/(CH)	F2 (Hz)/(CH)	F3 (Hz)/(CH)
ca[ts'a]	750~950/(CH5)	1150~1350/(CH7)	2150~2350/(CH12)
ci[ts'i]	150~350/(CH2)	1350~1550/(CH8)	2550~2750/(CH14)
ce[ts'ɤ]	350~550/(CH3)	1150~1350/(CH7)	2350~2550/(CH13)
cu[ts'u]	150~350/(CH2)	750~950/(CH5)	2550~2750/(CH14)

Table I と II はそれぞれ「そり舌音」の発話の評価基準と有声期間の発話のフォルマント周波数値を示している。

「そり舌音」の cha[tɕ'a]の音節は、VOT 開始から VOT/2 の間の周波数 2750Hz(CH15)~5750Hz(CH29)までパワーが17以上で、F1, F2 と F3 はそれぞれ 550~950Hz, 950~1150Hz と 1750Hz~1950Hz である。

chi[tɕ'i]の音節は VOT 中の周波数 1750(CH10)~6350Hz(CH32)までパワーは25以上である。フォルマント周波数値 F1, F2 と F3 はそれぞれ 150~350Hz, 1950~2150Hz と 2750~2950Hz である。

che[tɕ'ɤ]の音節は、0~VOT/2/3の間の周波数 1750(CH10)~6350Hz(CH32)までパワーが34以上である。フォルマント周波数値 F1, F2 と F3 はそれぞれ 350~550Hz, 1150~5950Hz と 2350~2550Hz である。

chu[tɕ'u]の音節は、0~VOT/2/3の間の周波数 3250(CH17)~6150Hz(CH31)までパワーが11以上である。フォルマント周波数値 F1, F2 と F3 はそれぞれ 150~350Hz, 750~950Hz と 2550~2750Hz である。VOT中の各チャンネルのパワーの自動計算により、「そり舌音」のパワー変化の特徴から、「そり舌音」の音節の自動判別が可能となる。

Table IIIは「舌歯音」ci[ts'i], ce[ts'ɤ]と cu[ts'u]の音節のフォルマントを示す。「舌歯音」の有声期間中の各チャンネルのパワーを自動計算し、それによるフォルマント周波数 F1 からフォルマント周波数 F3 までの値を測定した。ca[ts'a]の音節は、チャンネル5, 7 と 12 に

それぞれ強いパワーが現れ、フォルマント周波数 F1~F3 の値はそれぞれ 750~950Hz, 1150~1350 Hz と 2150~2350 Hz であることが分かる。「舌歯音」の音節は有声区間中の 35 チャンネル (CH) 毎のパワーを自動的に計算することにより、F1~F3 の値が得られ、自動的に判別することが可能となる。

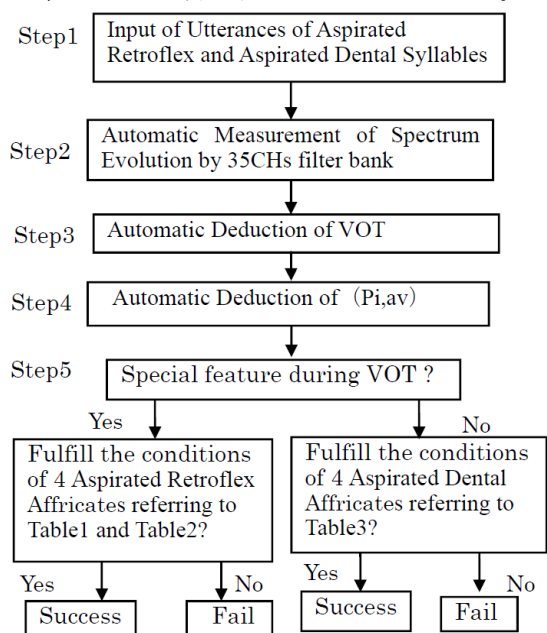


Fig.6 : Discrimination diagram of aspirated retroflex and dental affricate

(2) 自動評価の手順

本研究では、「そり舌音」cha[tʂ'a], chi[tʂ'i], che[tʂ'ɤ]と「舌歯音」ca[tʂ'a], ci[tʂ'i]と ce[tʂ'ɤ], cu[tʂ'u]の音節に対して、日本人学生 50名(中国語話者 8名の聴取試験に合格したもの)、中国語話者 9名の発話を用いて自動判別のための判別基準を作成し、その基準に基づき自動判別システムを用いて、発話の評価を行った。

Fig.6 は中国語の有気音の「そり舌音」と「舌歯音」の音節の自動判別の流れ図である。Step1 では「そり舌音」と「舌歯音」の発話をマイクによりコンピュータに入力する。Step2 では各発話を 35 チャンネル (CH) のフィルターバンク (FB) により自動解析する。Step3 では各発話の VOT の長さを自動的に測定する。

Step4 では、今回開発した発話のパワースペクトルの自動測定システムを用いて、35 の CH 毎にそれぞれ VOT 中の平均パワーと有声期間中のパワーを自動的に計算する。Step5 では、VOT 中の各 CH のパワーに特徴的なパターンがあれば、Table I と Table II に示している基準に従って「そり舌音」の発話の自動判別を行った。VOT 中の各 CH のパワーに特徴的なパターンがなければ、TABLE III により「舌歯音」の発話の自動判別を行った。

(3) 実験結果

Table IV に示しているように本実験は日本人学生 50 名の「そり舌音」cha[tʂ'a], chi[tʂ'i], che[tʂ'ɤ]と chu[tʂ'u]に対して、「そり舌音」と「舌歯音」の自動判別システムの実験を行った。その結果、正確な「そり舌音」の発話と判別されたのはそれぞれ 36,40,30,32 であったが、中国語話者 9 名によれば、それぞれ 38, 41, 32, 35 個の発話が正確な「そり舌音」と評価され、中国語話者が正確と判定した発話のうち、それぞれ 94%, 97%, 93%,91%を正確な発話と判定しており、いずれも高い判別率と言える。同じく日本人学生 50 名の「舌歯音」ca[tʂ'a], ci[tʂ'i]と ce[tʂ'ɤ], cu[tʂ'u]の発話は、自動判別システムにより 44, 40, 37, 37 個が正確な「舌歯音」と判別され、中国語話者 9 名には 44, 42, 38, 39 個の発話が正確な「舌歯音」と評価された。「舌歯音」では 100%, 95%, 97%, 95%とさらに高い判別率が得られた。

本研究は「そり舌音」と「舌歯音」の自動判別を目的として、VOT 中と有声期間中の発話の周波数スペクトルを自動計算し、それぞれの発話の特徴を抽出した。求めた判別基準により自動判別システムを開発し、それぞれの発話の音節の VOT 中と有声期間中の平均パワーを自動測定することにより「そり舌音」と「舌歯音」の音節の自動判別の実験を行った。その結果、いずれも高い判別率が得られた。

Table IV : Number of correctly judged student's pronunciations among 50 ones by native Chinese speaker and our judgment system

	Aspirated retroflex syllables				Aspirated dental syllables			
	cha[tʂ'a]	chi[tʂ'i]	che[tʂ'ɤ]	chu[tʂ'u]	ca[tʂ'a]	ci[tʂ'i]	ce[tʂ'ɤ]	cu[tʂ'u]
Number of correctly judged ones by native Chinese speakers (1)	38	41	32	35	44	42	38	39
Number of correctly judged ones by our judgment system (2)	36	40	30	32	44	40	37	37
Correct judgment ratio (2)/(1)	94%	97%	93%	91%	100%	95%	97%	95%

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

- ① 星野朱美, 安田明生
Automatic Judgment System for Chinese Retroflex and Dental Affricates Pronounced by Japanese Students, International Journal of Information and Education Technology, 査読有, Vol.4, No.2, April 2014, pp.167-171
- ② 星野朱美, 安田明生
Automatic Discrimination of Pronunciations of Chinese Retroflex and Dental Affricates Springer, 査読有, LNAL 8202, 2013, pp. 303-314

[学会発表] (計 5 件)

- ① 星野朱美, 安田明生
Automatic Judgment System for Chinese Retroflex and Dental Affricates Pronounced by Japanese Students, 2013 6th International Conference on Computer Science and Information Technology, 査読有, December, 21, 2013, Paris, France
- ② 星野朱美, 安田明生
Automatic Discrimination of Pronunciations of Chinese Retroflex and Dental Affricates, CCL2013 and 1st International Symposium NLP-NABD 2013, 査読有, October, 11, 2013, Suzhou, China.
- ③ 星野朱美, 安田明生
Discrimination of Chinese pronunciations of aspirated dental and retroflex syllables according to breathing power and its frequency dependency during VOT, Acoustics, 2012, 査読有, April, 23-27, 2012, Nantes, France.
- ④ 星野朱美, 安田明生
「中国語のそり舌音と舌歯音の発話の自動判別」日本音響学会 2013 年春季研究発表会, 3, 14, 2013, 東京工業大学
- ⑤ 星野朱美, 安田明生
「周波数の高さとも VOT 中のパワーによりそり舌音と舌歯音の発話の識別」, 日本音響学会 2012 年春季研究発表会, 日本音響学会, 3, 13, 2012, 神奈川大学

6. 研究組織

(1) 研究代表者

星野 朱美 (HOSHINO AKEMI)
富山高等専門学校・一般教養科・准教授
研究者番号 : 90300566