

平成 22 年 5 月 28 日現在

研究種目：基盤研究（C）

研究期間：2007～2009

課題番号：19500843

研究課題名（和文）中国語を学習する日本語話者向けの発音訓練のシステムの開発

研究課題名（英文）Pronunciation training system of Chinese Aspiration  
for Japanese Students

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

研究成果の概要（和文）：日本語話者は中国語の「有気音」と「無気音」を聴き取れなく、帰宅後の自習も自分の発音に対して正確な評価手段もない。本研究では、中国語話者9名と日本語話者20名の中国語の有気音二重母音唇音pai[p'ai], pao[p'ao], pei[p'ei]とpie[p'ie], 舌尖音 tai[t'ai], tao[t'ao], tie[t'ie]と tou[t'ou], 舌根音 kai[k'ai], kao[k'ao], kie[k'ie], kou[k'ou]の3種類の音節、全部で12音節の発話のVOTとVOT中に吐いた息のパワーを比較するにより、従来の有気音評価の結果と異なるパターンを発見した。すなわち、有気音の正確さはVOTの長さだけではなく、VOT中に吐いた息のパワーにも依存することが分かった。その結果より、新たな発話評価基準を確立した。更にその新しい発話評価基準を用いて、中国語の「有気音」と「無気音」の発話の訓練システムを開発し、「有気音」と「無気音」の発話の正確さを判定した。その結果「無気音」の判別率は約85%、「有気音」の判別率は76%であった。

研究成果の概要（英文）：A Chinese aspiration is generally considered to be very difficult for Japanese students to reproduce and perceive. We measured the voice on set time (VOT) and mean breathing power during VOT, as evaluation parameters, of diphthong sounds of the labial of pai[p'ai], pao[p'ao], pei[p'ei] and pie[p'ie] diphthong sounds of the alveolar of tai[t'ai], tao[t'ao], tie[t'ie], and tou[t'ou], diphthong sounds of the velar of kai[k'ai], kao[k'ao], kie[k'ie] and kou[k'ou] uttered by nine Japanese students and nine native Chinese speakers. Then, we found that the VOT was not the sole measure for evaluating the pronunciations; the mean breathing power during VOT is also a useful measure for evaluation. In order to offer an instruction tool for students for their private study, we designed the automatic measurement system of VOT and relative breathing power. And we apply it to the uttered sounds of above diphthong syllables and examine the performance of the system. The result shows that the averaged recognition rates are 85% and 76% for each syllable of unaspirated and aspirated, respectively.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2007年度	800,000	240,000	1,040,000
2008年度	500,000	150,000	650,000
2009年度	500,000	150,000	650,000
年度			
年度			
総計	1,800,000	540,000	2,340,000

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：中国語，有気音，無気音，VOTの長さ，VOT中パワー，VOTの自動測定，VOT中のパワーの自動計算，有気音と無気音の自動判別

### 1. 研究開始当初の背景

近年，中国の経済の発展と共に，わが国においても中国語を学習する学生がますます多くなって来ている。中国語の多数の発音の殆どが日本語に無い音なので，殆どの学生は中国語の発音を難しく感じている。それらの中でも困難なのは「有気音」と「無気音」の区別である。日本語話者の中国語教師の中には正しく発音し，正しく聞き分ける能力を持たない者もいるし，一対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もない。中国語の学習に際して，自分の発音は正確かどうか，判断する手段を持ち，自分の発音が正しいかどうか常にチェックしながら学習できることが望まれている。

本研究では，中国語話者なみに聞き分ける能力を持った，「有気音」と「無気音」発音訓練をする教材の開発を目指す。

### 2. 研究の目的

従来の有気音の研究では，VOT（破裂から声帯振動開始時点までの時間を音響学的には有声開始時間 VOT (voice onset time) と言う）が短ければ発話は悪く，無気音に聞こえ，VOT が長ければその発話は正確な有気音であると言われていたが，いままで，筆者は中国語の 21 有気音の単母音の音節に対して，学生の発話と中国語話者の発話の VOT を解析した。その結果は，有気音の発話の正確さは VOT の長さだけでなく，VOT 中に吐かれた息のパワーにも強く依存することが分かった。本研究では，中国語話者 9 名と日本語話者 20 名の「有気音」と「無気音」のいずれも二重母音の唇音  $\text{pai}[\text{p}'\text{ai}]$ ， $\text{pao}[\text{p}'\text{ao}]$ ， $\text{pei}[\text{p}'\text{ei}]$  と  $\text{pie}[\text{p}'\text{i}\epsilon]$ ，舌尖音  $\text{tai}[\text{t}'\text{ai}]$ ， $\text{tao}[\text{t}'\text{ao}]$ ， $\text{tie}[\text{t}'\text{i}\epsilon]$  と  $\text{tou}[\text{t}'\text{ou}]$ ，舌根音  $\text{kai}[\text{k}'\text{ai}]$ ， $\text{kao}[\text{k}'\text{ao}]$ ， $\text{kie}[\text{k}'\text{i}\epsilon]$  と  $\text{kou}[\text{k}'\text{ou}]$  の 3 種類の音節，全部で 12 音節の発話の VOT と VOT 中のパワーを比較することにより，学生の発話の問題点を分析した。その結果，従来の研究では有気音の発話の正確さは VOT の長さのみに依存すると言われていたが，必ずしもすべてがその判定基準に従うものではないことを見出し，VOT 中のパワーを解析することにより新たな評価基準の確立を試みた。本研究目的は新しい評価基準を用いて，「有気音」と「無気音」の発話の

正確さの判別可能になるシステムの開発を目指す。

### 3. 研究の方法

#### (1) 有気音と無気音の相違点

中国語の有気音は閉鎖子音であり，それを調音するとき，閉鎖段階と解放段階がある。音響学的には，閉鎖段階は声道が閉鎖され，解放段階においては，口の中に閉じ込められた空気が破裂後に氣息性の摩擦音を生成するか否かによって，「有気音」と「無気音」に分類される。有気音は空気が部分的に閉じた声帯を通過して咽頭を通過するときに生み出される氣息性雑音である。無気音では破裂後に雑音がほとんど現れない。

図 1 の右側の有気音  $\text{tai}[\text{t}'\text{ai}]$  は，解放段階で強く息を吐くと，声帯振動が始まる前に急激な強いパワーが放射される。破裂後から声帯振動開始時点までの有声開始時間 VOT は 91ms である。左側の無気音  $\text{tai}[\text{tai}]$  では閉鎖の段階が終わって殆どすぐ声帯振動によるフィルメントが現れて，VOT は殆どゼロである。

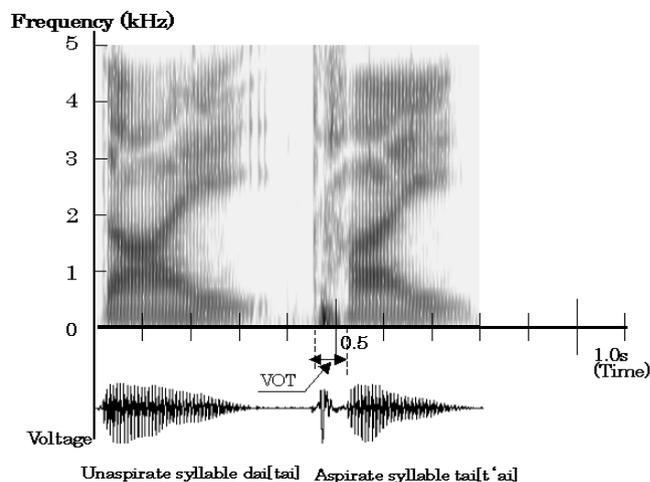


図. 1 中国語話者の舌尖音の有気音  $\text{tai}[\text{t}'\text{ai}]$  と無気音  $\text{dai}[\text{tai}]$  発話による無気音

#### (2) 調査対象と発音の評価方法

本研究において調査対象としたのは，中国語を週 2 コマ，約 1 年間学習した日本人学生 20 名，中国語話者 9 名である。発話評価の聴取試験は学生の各発話を中国語話者 8 人に聞

かせて、正確な有気音に聞こえる場合は「3」を、はっきり有気音に聞こえない場合（有気音と無気音の中間の音）は「2」を、はっきり無気音に聞こえる場合は「1」をつけた。正確なデータを得るため、8人の中国語話者の評価が大きく割れた発話と評価の標準偏差が0.64以上のもの、また録音したときにマイクが口付近すぎて、音が割れた発話とマイクが口から遠すぎてS/Nが小さい発話、それにVOTが充分長く、良い評価の発話の一部を除いた。なお、中国語話者については互いに発話評価し合った結果、いずれの音節も評価「3」であった。発話の合格基準としては、「1」の評価が無く、5人以上の中国語話者が「3」を付け、その他は「2」をつけた発話、すなわち平均評価「2.6」以上とした。

### (3) 中国語話者と日本人学生のVOTの比較

従来の研究では、VOTが短ければ発話は悪く、無気音に聞こえ、VOTが長ければその発話は正確な有気音であると言われていた。しかし、有気音の二重母音の音節に対して、調査の結果はこれと異なるパターンもあることが分かった。図2は中国語話者と日本人学生の舌尖音の有気音 tie[t' i ε] のスペクトログラムである。右側は中国語話者の発話で、VOTは57msである。左側の発話は日本人学生のもので、VOTは21msで中国語話者のVOTより約36msも短く、評価も悪く「1.0」である。

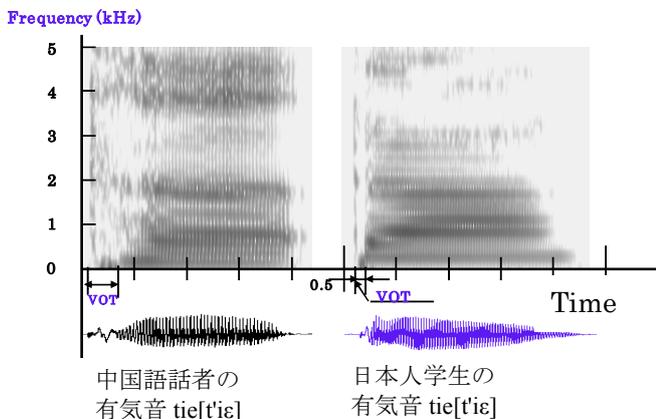


図.2 中国語話者と日本人学生の有気音 tie[t' i ε] のスペクトログラムの比較

今回 VOT の調査により、学生の発話の VOT は中国語話者の VOT より長いのに、発話評価が悪かったり、また、学生の発話の VOT は中国語話者のものとほとんど変わらないのに評価が悪いことが多い。他の種類の有気音も同様の例が散見されることから、中国語の有気音の正確さは VOT の長さだけでは評価できないという結論に達した。

### (4) 相対平均パワーの導出法

今回の調査で学生の有気音の発話は、VOT の長さは中国語話者の VOT と比較して、ほとんど差がないのに、発話評価が悪いことやまた、逆に VOT が長くても、発話評価が悪いことがある。この原因を追求するために、VOT 中に吐いた息のパワーを求め、発話評価との関係を調べた。

使用した音声分析ソフトでは、マイクロフォンで電圧に変換された音声の波形を、サンプリング周波数 11.025kHz で、16bit の AD 変換の後、種々の解析を行っている。パワーの導出は、その平均値  $V_A$  を基準として、連続するサンプル値  $V(n)$  と ( $n$  は正の整数) の差の絶対値を求め、5ms ごとの平均  $V_a(t)$

$$V_a(0.005m) = \frac{\sum_{n=55m+1}^{n=55m+55} |V(n) - V_A|}{55} \quad (1)$$

を求めて ( $m$  は 0 および正の整数)、デジタル値 300 を 0dB とし、そのパワー  $P(t)$  を次式により求めている。

$$P(0.005m) = 20 \log_{10} \left( \frac{V_a(0.005m)}{300} \right) \quad (2)$$

VOT 中の全エネルギーは VOT 中の  $P(t)$  の dB 値を 10 で除し、真数に戻して、5ms を掛けて加算して求めた。しかし、ここで求めたエネルギーは絶対的なものではなく、録音レベルの違いによる規格化を行っていない。これをここでは名目エネルギー  $W_{nom}$  と呼ぶことにする。すなわち VOT 中のパワーが強ければ有声化区間（母音の声帯振動の開始から発話終了までの時間）のパワーも強くなると考えられる。そこで、 $P_2$  を基準として摩擦性区間のエネルギーを正規化するために、まず VOT 中の最大パワー  $P_1$  と有声化区間中の最大パワー  $P_2$  との比  $R$  を次式により求めた。

$$r \text{ (dB)} = P_1 - P_2 \text{ (dB)} \quad (3)$$

$$R = 10^{r/10}$$

$$(4)$$

次に名目エネルギー  $W_{nom}$  にこの比  $R$  を掛けて相対エネルギー  $W_{rel}$  を

$$W_{rel} = W_{nom} \times R \quad (5)$$

求め、これを VOT で除して相対平均パワー  $P_{rel}$  を

$$P_{rel} = W_{rel} / VOT \quad (6)$$

導出した。

## 4. 研究成果

### (1) VOT と VOT 中の相対平均パワーと発話評価の関係

正確な有気音の発話は VOT 中のパワーと関係があるかどうか、中国語における有気音の二重母音の音節を調査するため、日本人学生 20 名、中国語話者 10 名の各発話の VOT を測定し、前節で述べた相対パワーの計算方法を用いて、VOT 中のパワーを計算した。

表1は日本人学生と中国語話者の有気音の発話に対して、VOTが短すぎるとき、VOTの長さが一定の幅内のときと十分長いときの三つ区間に分けて、それぞれのVOTの平均値と平均パワーPavの比較を示す。

VOTが短すぎる場合は、pai[p'ai]の範囲0~10ms時には学生の平均VOTは4msで、相対平均パワーPavは弱く1.8で、その発話の評価の平均値は悪く、1.3しかない。また、pei[p'ei]のVOTの0~22ms範囲場合は、平均VOTは16msで、Pavは強く30.5があったのに、平均評価は悪く1.5しかない。VOTが短すぎる場合は、発話評価はVOTの長さだけに依存する。

VOTが十分長い時は、pei[p'ei]の30ms~の範囲に学生の平均VOTは43msで、パワーの平均値は強く146で、その平均評価は高く2.9である。しかし、tao[t'ao]の中国語話者のVOTの45ms~範囲の場合は、平均VOTは長く94msで、Pavは非常に弱く0.085しかないのに、発話の平均評価は高く、3.0である。VOTが十分長い時は、発話の正確さはパワーの強さに関係なく、VOTの長さだけに依存することが分かった。

VOTが一定の幅内のとき、tie[t'ie]の音節のVOTが35ms~50msでは学生の平均VOTは43ms、パワーは非常に弱く0.61しかない。発話の平均評価も悪く2.4で、不合格の評価となっている。pao[p'ao]の音節のVOTの範囲が35ms~40msでは学生のVOTは38msで、tie[t'ie]のVOTより5ms短いのに、平均パワーは非常に強く130で、平均発話の評価も高く2.7である。他のデータも同じ傾向を見られた。調査の結果により、発話の正確さはVOTの長さだけでなく、吐いた息のパワーにも依存することが分かった。

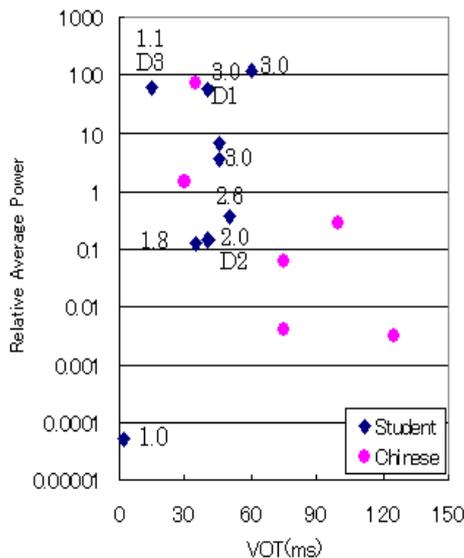


図3. 舌尖音 tao[t'ao]のVOTとVOT中の相対平均パワー分布図

Syllable	VOT Range	Japanese Student				Native Chinese speaker		
		Average VOT(ms)	Average Pav	Average Grade	Data	Average VOT	Average pav	Data
pai[p'ai]	0-10ms	4	1.8	1.3	3	---	---	0
	20-30ms	25	6	2.3	4	---	---	0
	35ms~	38	310	3.0	8	72	106.5	6
	Average	28	242.5	2.5		72	106.5	
pao[p'ao]	0-20ms	8.3	0.81	1.1	3	---	---	0
	35-40ms	38	1300	2.7	3	33	130.7	3
	41ms~	60	570	2.9	9	91	13.2	4
	Average	36	62.7	2.7		72	52.4	
pei[p'ei]	0-22ms	16	30.5	1.5	3	---	---	0
	25-29ms	25	4.3	2.5	2	---	---	0
	30ms~	43	146.3	2.9	8	65	220.7	5
	Average	30	76.3	2.6		65	221.0	
pie[p'ie]	0-25ms	18	13.3	1.3	4	---	---	0
	26-40ms	40	37.6	2.6	3	---	---	0
	60ms~	80	600	3.0	6	106.3	4.7	4
	Average	47	43.6	2.5		106	4.7	
tai[t'ai]	0-10ms	6	0.88	1.1	2	---	---	0
	30-44ms	38	1.91	2.6	3	---	---	0
	45ms~	49	38.5	3.0	8	74	14.2	5
	Average	39	13.8	2.7		74	14.2	
tao[t'ao]	0-15ms	8.5	31.5	1.0	3	---	---	0
	20-40ms	38	18.8	2.4	3	33	36.8	2
	45ms~	50	31.2	2.9	8	94	0.085	4
	Average	36.9	27.1	2.6		73.3	12.3	
tie[t'ie]	0-30ms	25	0.71	1.7	3	---	---	0
	35-50ms	43	0.61	2.4	4	---	---	0
	55ms~	74	2.4	3.0	7	90	7.5	7
	Average	56.3	1.5	2.5		90	7.5	
tou[t'ou]	0-25ms	14	0.43	1.5	2	---	---	0
	35-45ms	42	3.2	2.4	3	---	---	0
	50ms~	63	35.2	3.0	7	70	24.4	6
	Average	48	16.8	2.5		70	24.4	
kai[k'ai]	0-18ms	12	1.6	1.1	4	---	---	0
	37-44ms	39	3.7	2.7	5	---	---	0
	45ms~	51	21.4	3.0	9	74	14.6	9
	Average	35.8	18	2.8		74	14.6	
kao[k'ao]	0-14ms	9	0.9	1.3	3	---	---	0
	35-46ms	40	2.8	2.6	4	45	36.8	1
	51ms~	62	37.7	2.9	8	77	13.6	7
	Average	39	22	2.7		61	14.2	
kei[k'ei]	0-21ms	13	2.7	1.2	3	---	---	0
	34-45ms	39	23.3	2.5	5	41	4.3	2
	54ms~	69	40.1	3.0	8	76	44.9	7
	Average	48	32.5	2.6		72	43.6	
kou[k'ou]	0-23ms	15	0.51	1.5	3	---	---	0
	35-48ms	42	22.6	2.4	4	---	---	0
	53ms~	67	45.2	3.0	8	73	37.3	8
	Average	48	40.8	2.5		73	37.3	

表1 日本人学生と中国語話者の有気音の発話における平均VOTと平均パワーPavの比較

図3は舌尖音 tao[t'ao]のVOTとVOT中の相対平均パワー分布図でVOTが40ms付近に縦に並ぶD1, D2で、パワーはそれぞれ87, 0.2で発話評価は「3.0」と「2.0」と差がある。VOTはほとんど同じでも、パワーが強い方が評価も高いことがわかる。中国語話者の発話(相互聴取評価、いずれも「3.0」)はVOTが短いデータはパワーが強い。殆どの発話のVOTは充分長い、パワーは弱いことが分かった。

従来の研究では、VOTが短ければ発話は悪く、無気音に聞こえ、VOTが長ければその発話は正確な有気音であると言われていたが、本研究の結果により、有気音のVOTの長さが一定の幅内であれば、発話の正確さはパワーに依存すること、VOTが短すぎる場合と十分長い場合は発話の正確さはパワーに依存しないことも分かった。

(2) VOTの自動測定とパワーの計算

図4の下のトレースは中国語の有気音 pai[p 'ai] の発話パワーの包絡線である。有気音を発話する時に口の中に閉じ込められた空気を吐き出すことにより急に息の強さが増し、発話パワーが急に大きくなる。この時点を破裂点(t1)と言い、VOTの開始点でもある。一方、母音の声帯振動が開始する時点がVOTの終了時点(t2)である。

本システムはその特徴を用いて、VOTの開始点(破裂点 t1)をまず検出する、中心周波数 50Hz~6850Hz、幅 200Hz で、35 チャンネル (CH) のフィルタバンク (FB) を設計する。音声信号を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、そのパワー差の正の値の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。母音の声帯振動が始まる時、フォルマントが現れ、この時点で f0 を検出し、VOTの終了時点(t2)とする。すなわち

$$VOT = t2 - t1 \quad (1)$$

VOT中の相対平均パワーの自動計算については、まず発話全体の各値に5msずつのパワーを計算し、それからVOT中の全値を加算し、VOTで除することによりVOT中の平均パワー  $P_{nom}$  を求める。録音レベルの正規化を行うため、それぞれVOT中と有声区間でパワーが最大となる時点を検出し、その時点のパワーを計算しP1とP2とし、相対平均パワーを規格化している。

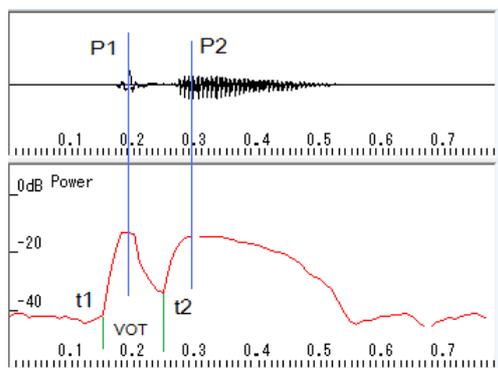


図4. pai[p 'ai]のパワーの包絡線(下)

(3) 有気音と無気音の判別

図5は中国語の発音訓練システム構成図である。日本語話者による中国語の発話を収録し、コンピュータにより発話の正確さの重要なパラメータであるVOTを計算する。VOTが短すぎる時は無気音と判断する。VOTが十分長い時は有気音と判断する。VOTが一定の幅内のときにはVOT中のパワーを計算し、著者らが確立した有気音の発話評価基準により有気音か無気音かの判別を行う。その時

はパワーが大きければ、発話の評価も高い。



図5. 発音の学習訓練のシステム構成図

(4) 結果

本システムを用いて、発話の判別の実験を行った。その結果は唇音の無気音の判別率は87%、有気音の判別率は79%、舌尖音の無気音の判別率は85%で、有気音の判別率77%、舌根音の無気音の判別率は84%で、有気音の判別率は74%であった。

本研究では学生の中国語有気音の発話の問題を分析することにより、新たな評価基準を確立し、それらを用いて「有気音」か「無気音」かの自動判別ができる中国語の発話訓練システムを開発した。無気音の判別率は良く84%以上であるが、有気音の判別率まだ高くない。更にデータを解析し、誤差が大きい原因を追究し、精度を上げるように努力する。

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1件)

星野朱美, Evaluation of Aspiration Sounds of Chinese Labial Diphthong Uttered by Japanese Students Using VOT and Breathing Power. 富山商船高等専門学校研究収録 2009. 巻42 pp151-156, 査読無

〔学会発表〕(計2件)

① 星野朱美, 安田明生, 中国語を学習する日本語話者向けの発音訓練システム, 日本語学 2010年春季研究発表会, 2010.03.8, 電気通信大学

② 星野朱美, 安田明生, 中国語有気音の二重母音の発話評価におけるVOTとパワー依存性について, 電子情報通信学会2010年総合大会, 2010.03.17, 東北大学

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

(2) 研究分担者 なし  
( )

研究者番号:

(3) 連携研究者 なし  
( )

研究者番号: