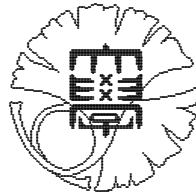


修士論文

英語話者を対象とした
日本語CALLシステムの構築と評価



平成19年2月2日提出

指導教員 峯松 信明 助教授

東京大学大学院情報理工学系研究科
電子情報学専攻

56442 羅 徳安

概要

1980年代から日本語学習熱が世界的に高まり、現在では世界127カ国で235万人以上の人々が日本語を学習している。これに伴い、高等教育機関の学習者、上級学習者の数も年々増え続けている。特に、オーストラリアでは、日本語は最も人気のある外国語であり、日本における英語のような存在であると言われている。一方、日本語教師が足りていないことも懸念されている。その解決策の一つとして、CALL (Computer Aided Language Learning) システムが注目されている。

近年、計算機技術の飛躍的向上に支えられ、音声・言語情報処理技術は著しく進歩し、各種技術をツールとして利用することも比較的容易になってきた。これらの技術に基づいて、外国語学習の更なる高度化・効率化を目的とし、学習支援環境構築を目的とした研究が広く行なわれ、種々の有効な技術、ツールの提供が行なわれてきた。しかし、従来のCALLシステム開発では、学習者群全体の誤り傾向及び教師群全体の評価戦略を統計分析し、およそ平均的な教師による平均的な学習者の発音指導を模擬する形態となっているものが多い。

本研究で検討するCALLシステムは、日本語教師の使い勝手を考え、1) その日本語教師による発音評価を模擬すること、2) 学習者の習熟度に応じて柔軟に発音誤り検出ができること、を念頭に置いた。それを実現するために、a) 当該日本語教師自身が定義する発音エラーカテゴリ、及び、b) 学習者毎の評定項目を準備し、更には、c) 当該教師自身の学習者音声に対する評価戦略を事前に聴取実験などを通して取得し、それを模擬する形で、豪英語の母語話者を対象とした、日本語発音とコミュニケーション能力の育成に焦点をあてた自律学習支援型システムを開発し、評価実験を行なった。

目次

| | | |
|----------|-----------------------------|-----------|
| 概要 | i | |
| 1 | 序論 | 1 |
| 1.1 | 日本語教育とその技術支援の現状 | 2 |
| 1.2 | 本研究の目的 | 2 |
| 1.3 | 本論文の構成 | 3 |
| 2 | 研究の背景 | 4 |
| 2.1 | 日本語と英語の音声学的差異 | 5 |
| 2.1.1 | 母音 | 5 |
| 2.1.2 | 子音 | 5 |
| 2.1.3 | 音節 | 5 |
| 2.1.4 | アクセント | 5 |
| 2.1.5 | リズム | 7 |
| 2.2 | 外国語学習における発音教育の位置づけ | 7 |
| 2.3 | 発音教育システムの概要 | 9 |
| 2.3.1 | マルチメディア技術に基づく英語 CAI システム | 9 |
| 2.3.2 | 音声情報処理技術に基づく英語 CAI システム | 9 |
| 2.4 | 発音教育における本研究の位置づけ | 12 |
| 3 | 本研究における先行研究 | 13 |
| 3.1 | 日本語発音誤りの自動検出 | 14 |
| 3.1.1 | 学習者音声データベース | 14 |
| 3.1.2 | 学習者音声誤り分析 | 14 |
| 3.1.3 | 発音誤り検出法 | 14 |
| 3.1.4 | 検出精度 | 15 |
| 3.1.5 | まとめ | 15 |
| 3.2 | 日本語学習者の文発話の発音評価法 | 17 |
| 3.2.1 | 日本語教師による発音評価 | 17 |
| 3.2.2 | 音響モデル | 17 |
| 3.2.3 | 各音響特徴量と日本語教師の評価との相関 | 17 |
| 3.2.4 | 統計的発音評価法 | 18 |
| 3.2.5 | 考察 | 19 |
| 3.3 | CALL 構築を目的とした学習者音声データベースの構築 | 19 |
| 3.3.1 | DB の特性 | 20 |
| 3.3.2 | DB 構成 | 20 |

| | | |
|----------|---------------------------|-----------|
| 3.3.3 | 考察 | 20 |
| 3.4 | まとめ | 21 |
| 4 | 本研究で提案する CALL システム | 22 |
| 4.1 | 提案する CALL システムの特徴 | 23 |
| 4.2 | システムの概要 | 23 |
| 4.3 | 発音エラーカテゴリーの定義 | 26 |
| 4.3.1 | 予測される豪英語話者による発音誤り | 26 |
| 4.3.2 | 書起こしによる誤り傾向分析 | 28 |
| 4.3.3 | エラーパタンリスト | 28 |
| 4.4 | 学習者依存型誤り検出ネットワーク文法 | 30 |
| 4.5 | 音響モデルと後処理手法 | 33 |
| 4.6 | まとめ | 34 |
| 5 | 評価実験 | 35 |
| 5.1 | 発音誤り検出評価実験 | 36 |
| 5.1.1 | 使用した音声データと正解ラベル | 36 |
| 5.1.2 | 使用したネットワーク文法 | 36 |
| 5.1.3 | システム精度 | 36 |
| 5.1.4 | 異なる正解ラベルによる精度の変化 | 38 |
| 5.1.5 | 結果と考察 | 38 |
| 5.2 | 教師による評定との相関分析 | 38 |
| 5.2.1 | 日本語教師による文節評定データの収集 | 39 |
| 5.2.2 | 各教師が異なる音声提示条件での評定比較 | 40 |
| 5.2.3 | 日本語教師間の文節採点相関 | 40 |
| 5.2.4 | CART による文節評定推定 | 40 |
| 5.2.5 | システム評定と教師による評定分析 | 41 |
| 5.2.6 | 結果と考察 | 42 |
| 5.3 | 学習者による主観的評価 | 42 |
| 5.4 | まとめ | 43 |
| 6 | 結論 | 44 |
| 6.1 | 終わりに | 45 |
| 6.2 | 今後の課題 | 46 |
| | 参考文献 | 47 |
| | 謝辞 | 50 |
| | 発表文献 | 51 |

| | |
|----------------|----|
| 付録 | 52 |
| A 読上げ音声収録用テキスト | 52 |

目 次

| | | |
|-----|--|----|
| 2.1 | 日本語における母音図 | 6 |
| 2.2 | 英語における母音図 | 6 |
| 2.3 | 日本語・英語における子音の分類 | 6 |
| 2.4 | Bachman's Model | 8 |
| 2.5 | マルチメディア技術に基づく英語 CAI システム:Microsoft ENCARTA . . . | 10 |
| 2.6 | 音声情報処理技術に基づく英語 CAI システム:English Now! | 11 |
| 3.1 | 発音誤り検出システム | 16 |
| 4.1 | 音声評価システム | 24 |
| 4.2 | 学習者用録音画面 | 24 |
| 4.3 | 音声評価画面 | 25 |
| 4.4 | 管理画面 | 25 |
| 4.5 | 学習者レベル別誤りの総数 | 30 |
| 4.6 | ネットワーク文法 | 32 |
| 4.7 | ネットワーク文法生成 GUI | 32 |
| 5.1 | システム精度 (後処理を行わない) | 37 |
| 5.2 | システム精度 (後処理を行った後) | 37 |

表 目 次

| | | |
|-----|--|----|
| 2.1 | 日本語のモーラと英語のシラブルの構造的差異 | 8 |
| 3.1 | 韓国語母語話者の発音誤り傾向 | 15 |
| 3.2 | 等誤り率 (%) の比較 | 16 |
| 3.3 | 文ごとの推定スコアと教師の評価の相関 | 19 |
| 4.1 | エラータイプとその頻度 (高1 低19) | 29 |
| 4.2 | 音響分析条件 | 33 |
| 5.1 | 各文法のカバー率 | 36 |
| 5.2 | 正解書起こし基準による認識率変化 | 38 |
| 5.3 | 採点者背景 | 39 |
| 5.4 | 各教師が文節のみを提示する条件での文節採点と文全体を提示する条件での 文節採点との相関 | 40 |
| 5.5 | システム評定と各教師評定との相関関係 | 41 |
| 5.6 | 模擬対象教師による評定との相関関係 | 42 |
| 5.7 | アンケート結果 | 43 |

第1章

序論

1.1 日本語教育とその技術支援の現状

1980年代から日本語学習熱が世界的に高まり、現在では世界127カ国で235万人以上の人々が日本語を学習している。例えば、オーストラリアでは、日本語は他の外国語と比べて人気が高い。日本語学習者が38万人にも達している。オーストラリアのほとんどの大学では、日本語を履修、また専攻することができる。大学で日本語を専攻する場合は、語学だけでなく社会、文化、経済などの日本学も取るのが普通であり、Bachelor of Asian StudiesまたはBachelor of Artsとして位置付けられる。また、日本語の学位だけでなく、例えば日本語と法学、日本語と商学などDouble Degreeとして二つの学位を専攻する場合もある[1]。

この日本語学習熱に伴い、高等教育機関の学習者、上級学習者の数も年々増え続けている一方、日本語教師が足りていないことが懸念されている[2]。周知のように、日本語と英語は音声学的（リズム、イントネーションを含む）にも、言語学的にも非常に大きな差異を持つ言語対である[3, 4]。特に漢字学習は、英語のような表音文字を第一言語としている日本語学習者にとって、困難極める問題である[5]。これらに加え、各言語を母語とする話者が持つ音声知覚プロセス間における差異[6]、物事を考える発想、論理プロセス間における差異[7]、更にはコミュニケーションにおける文化的差異[8, 9]もが存在し、英語話者にとって、日本語の習得・運用は非常に高いハードルとなっている。発音教育に絞った場合も、教員が個々の学習者に、絶えずface to faceで発音指導を実施するには、多大な時間とエネルギーが必要である。

一方、近年の計算機技術・音声情報処理技術の進展に伴い、語学学習に対する技術的サポートも活発に議論され、発音学習支援システムの研究開発も盛んに行われ、日本語の特殊拍に着目した学習システム、および、音素レベルでの発音誤りの検出と指摘ができる学習システムすでに開発されている[10, 11]。しかし現在の音声・言語情報処理技術は、その多くが数理統計的な枠組みに基づいて構築されており、その結果、大規模データベースは技術高度化の必須条件となっている[12]。このような大規模データベースに頼る発音学習支援システムは、平均的な学習者像を前提とし、教師の平均的な評価を求めるもので、学習者レベルや学習状況の多様性に柔軟に対応することができず、教育現場のニーズに合わない場合も考えられる。実際、最近になって発音学習支援システムに懐疑的な報告を耳にするのも事実である[13]。教育現場で十分利用できる信頼性を持つ新しいCALLシステム開発技術を求められている。

1.2 本研究の目的

本研究では従来とは異なり、日本語教師の使い勝手を考え、当該日本語教師自身が定義する発音エラーカテゴリ、及び、学習者毎の評定項目を準備し、更には、学習者音声に対する評価戦略を事前に聴取実験などを通して取得し、それを模擬する形で、豪英語の母語話者を対象とした、新しいCALLシステムの開発を行なった。このような開発は、日本語教育、教育工学、音声情報工学を専門とする研究者が相互協力しながら、教育現場の二-

ズに沿った開発を試みる必要があるが、本研究は各々の分野を専門とする研究者によって行なわれたものである。

1.3 本論文の構成

本論文では、まず2章で発音教育に関する背景について述べ、3章では、CALLシステムに関する先行研究を紹介する。それらを踏まえて、4章では本研究で提案した新しいCALLシステム開発手法を紹介する。5章では、システムに対する種々の評価実験について述べる。最後に、6章で本研究のまとめと今後の課題を述べる。

第2章

研究の背景

2.1 日本語と英語の音声学的差異

2.1.1 母音

日本語の母音は/a,i,u,e,o/の5母音である。図.2.1は、日本語母音を発音するとき、舌の位置を模式的に表しており、縦軸は舌の高さ（口の開き具合）、横軸は舌の部位（前舌か後舌か）を表している [14]。一方英語における母音図を図.2.2に示す（ただし、二重母音に関しては省略してある）。短母音5つ（/ɪ,ʊ,ɛ,ʌ,æ/）と長母音4つ（/i,u,ɔ,ɑ/）の合計9つ、二重母音は/eɪ,ɔɪ, aɪ, aʊ, oʊ/の4つ、さらに弱母音（/ə/）を入れると全部で21種類になる。

日本語における長母音は音質的に見ると短母音が並んだものとしてとらえることが出来るが、英語の長母音は短母音を長く発音したものではなく、例えば beat の [i] と bit の [ɪ] は発音の仕方も音質も異なる別種の音である。二重母音も同様に、[aɪ] は [a]+[ɪ] ではなく、[aɪ] で1つの音である。そのため、長母音と二重母音は短母音とは別の音として取り扱われる。また、英語には弱母音（schwa）と呼ばれる、弱く発音されるときのみ現れる母音が存在する。日本語では基本的にどの母音も同じ強さで発音されるため、英語を母語話者とする日本語学習者は特に弱母音化のミスが多いと言われている。

2.1.2 子音

子音に関しては、調音位置と調音方法によって図.2.3のように分類される [15]。無声/有声の対で示しており、発音記号により表記してある。日本語におけるラ行の子音である [r] は英語の /l/ と /r/ とは異なるため、英語話者学習者は、この子音を /l/ と /r/ に置き換えられることが多い。

2.1.3 音節

日本語と英語にはリズム構造に関する基本的韻律単位が異なる。英語は音節、あるいはシラブル (syllable) と呼ばれるものを基本的単位とする言語であり、日本語は音節より小さな単位であるモーラ (mora) を基本的単位とする言語である [16]。表 2.1 にモーラとシラブルの構造的差異を示す。日本語の母音数は5種類であるが、英語の母音の種類は約20種類となっており、シラブル/モーラの構造的差異から、モーラの種類数は約100であるが、シラブルは約10,000と言われる。

2.1.4 アクセント

日本語において、音の高さの変化を指し、高さアクセントと呼ばれ、音声情報処理分野では、基本周波数のみを用いて記述される。一方の英語においては、強さアクセントと呼ばれる。音声情報処理の分野ではパワーのみでは十分に記述できず、基本周波数、短時間

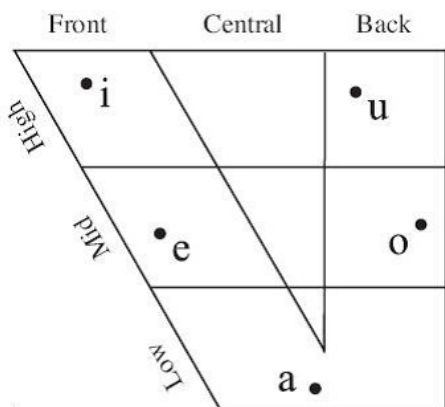


図. 2.1: 日本語における母音図

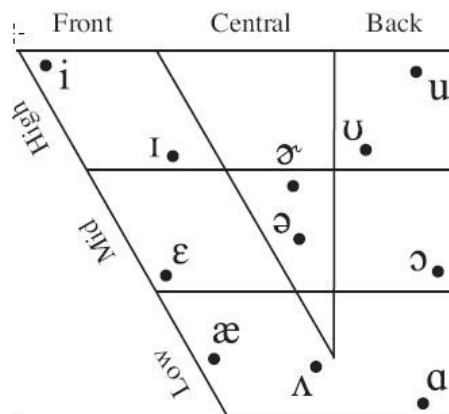


図. 2.2: 英語における母音図

| 調音位置 調音方法 | | 唇音 | | 歯 | 歯茎 | 後部 歯茎 | そり 舌 | 硬口蓋 | 軟口蓋 | 口蓋垂 | 声門 |
|--------------|---|-------|-------|-------|---------|----------|---------|-----|-------|-----|----|
| | | 唇 | 唇歯 | | | | | | | | |
| 閉鎖音 | 日 | p / b | | | t / d | | | | k / g | | ʔ |
| | 英 | p / b | | | t / d | | | | k / g | | ʔ |
| 摩擦音 | 日 | ɸ | | | s / z | | | ç | | | h |
| | 英 | | f / v | θ / ð | s / z | ʃ / ʒ | | | | | h |
| 破擦音 | 日 | | | | ts / dz | tʃ / dʒ | | | | | |
| | 英 | | | | | tʃ / dʒ | | | | | |
| 鼻音 | 日 | m | | | n | | | ɲ | ŋ | ɴ | |
| | 英 | m | | | n | | | | ŋ | | |
| 弾き音 | 流 | | | | r | | | | | | |
| | 音 | | | | l ɹ | | ɻ | | | | |
| 接近音 | 渡 | | | | | | | j | w | | |
| | り | | | | | | | j | w | | |
| 音 | 日 | w | | | | | | j | w | | |
| | 英 | w | | | | | | j | w | | |

図. 2.3: 日本語・英語における子音の分類

パワー，短時間パワーの時間累積値，持続時間，母音の音質などがアクセントの記述に用いられている．英語のアクセントを以降，特に「強勢」という言葉を用いる．

強勢 (stress) ある音節を発音するに当たって音源である呼気が強くなったりその量が多くなると喉頭や調音器官が緊張して調音のエネルギーが強くなり，聞き手が感じる音の大きさ (loudness) が増大する現象をいう．強勢を受けた音節はピッチが高まり，音が長めになる．強勢は強強勢 (strong stress) と弱強勢 (weak stress) とに2分され，強勢アクセントでは全ての音節はいずれかを受ける．

単語強勢 (word stress) 語中にある音節に置かれた強い強勢

文強勢 (sentence stress) 文中の特定の音節が持つ強勢およびその強弱の差．語義を持つ内容語 (content word) は強い文強勢を受け，機能語 (function word) と呼ばれる語義が希薄で主として内容語同士の文法的関係を示す働きをする語の文強勢は弱い．

2.1.5 リズム

リズムにおける等時性に関しても日本語と英語は異なる．日本語は，フランス語やイタリア語と同様に，一つ一つの音節が等間隔で発音され，音節拍のリズム (syllable-timed rhythm¹⁾) を持つといわれる．一方の英語の方は，ドイツ語，ロシア語と同様に強い強勢が等間隔に繰り返され，強勢拍のリズム (stress-timed rhythm) を持つといわれる．特に，英語のリズムを形成するものとして，ある一つの強勢音節から次の強勢音節までを脚 (foot) という．

2.2 外国語学習における発音教育の位置づけ

外国語学習における目標として，読む能力 (Reading)・書く能力 (Writing)・聞く能力 (Listening)・話す能力 (Speaking) の4技能の習得が挙げられるが，近年ではこれらの言語知識を習得するだけでなく，自分の意図を相手に伝えることや状況に合わせて言い方を変える等といった現実世界において実際に言語を運用する能力，即ちコミュニケーション能力の育成に主眼が置かれている．このようなコミュニケーション能力においては，特に話す能力・聞く能力を重視した教育が行われる．そして，語学学習に対して音声情報処理技術による支援を行う場合，主にこれら2つの能力の学習が支援の対象となる [17]．外国語によるコミュニケーション能力を評価する場合には，上記のような言語能力以外にも様々な観点から評価がなされるべきであり，例えば Bachman によって図 2.3 のように外国語コミュニケーションにおける様々な要素の分類がなされている [18]．そして，これらの各要素の中で，発音 (Pronunciation) はコミュニケーション能力を評価する上で非常に重要な要素となっており，[19] によると下記のような式で記述できるとしている．

¹日本語の場合，音節構造がモーラを単位としているため，mora-timed rhythm と呼ぶべきであろう．

表 2.1: 日本語のモーラと英語のシラブルの構造的差異

| | |
|------|--|
| モーラ | 母音 (V), 子音+母音 (CV), 撥音 (/N/), 促音 (/Q/) |
| シラブル | 母音を中心にその前後に0個以上の子音が連結した形をとる. 最長シラブルは CCCVCCCC . |

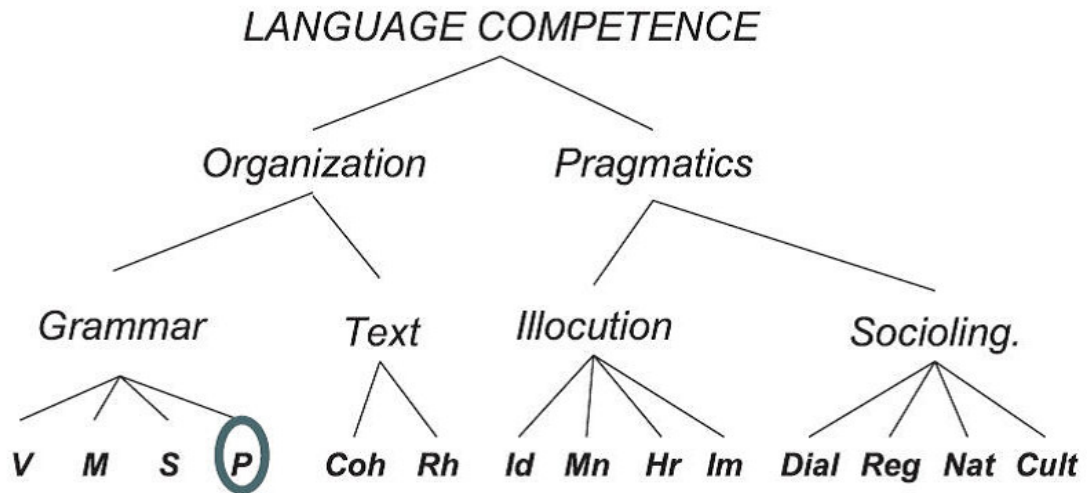


図. 2.4: Bachman's Model

$$comm. \simeq pron. \bullet lex. \bullet (1 + syn. + rhet. + illoc. + soc.) \quad (2.1)$$

発音能力 (pron.) と語彙力 (lex.) の二つがコミュニケーション能力に大きく寄与していることが示されている。即ち、英語学習を外国語によるコミュニケーション能力の習得として考えた場合、発音教育は非常に大きな位置を占めており、その技術支援には重要な意味を持つといえる。

2.3 発音教育システムの概要

本章では、発音教育システムにおける概要について述べる。

CAI(Computer Assisted/Aided Instruction) システム

指導者が学習者を指導する時の主要な機能をコンピュータに代行させ、あらかじめ用意された教材(コースウェア)に従って学習者個々に応じた最適な学習指導を行うシステム²

CALL(Computer Aided Language Learning) システム

コンピュータが生身の先生を取って代わるのではなく、コンピュータを使って言語学習の促進を図ることが主な目的³

現在、多くの英語 CAI システムが商用のソフトとして市場に流通している。以下にその英語 CAI ソフトの例を示す。

2.3.1 マルチメディア技術に基づく英語 CAI システム

一般的に商用の英語 CAI ソフトには音声情報処理技術を用いたものよりも、ビデオ映像などのマルチメディア技術を伴ったものが多い。その一例である Microsoft ENCARTA[20] を図 2.5 に示す。

これらの多くが、音声処理技術よりもビデオ映像を見た後にそれについての設問に答える形式になっている。また、千葉大学の協力でメディア開発センターが英語の CALL 教材 [21] を発表しているが、これもリスニング中心でリスニング後に設問に答える教材である。

2.3.2 音声情報処理技術に基づく英語 CAI システム

音声情報処理技術を用いた英語 CAI ソフトの一例である English Now![22] を図 2.6 に示す。この例のように音声波形を出力して比較をするものも多い。

この教材には「リズム」に関する教師音声との比較が導入されているが、この教材における「リズム」というのは「音の長さ」のみを考慮した結果である。また、分析結果によるフィードバックの意味するところが不明確で分かりにくく効果は薄いように考えられる。また音声認識技術を用いた英語 CALL システムとして [23] では、インターネット上で公開実験を行なっているが、韻律的特徴よりも音韻学習が中心の内容になっている。

²<http://sociolab.tamacc.chuo-u.ac.jp/~saizeHP/2000zemi/komine-p/txt1.html>

³<http://www15.freeweb.ne.jp/diary/yujii/english/qa/tesol.htm#CALL>



図. 2.5: マルチメディア技術に基づく英語 CAI システム:Microsoft ENCARTA

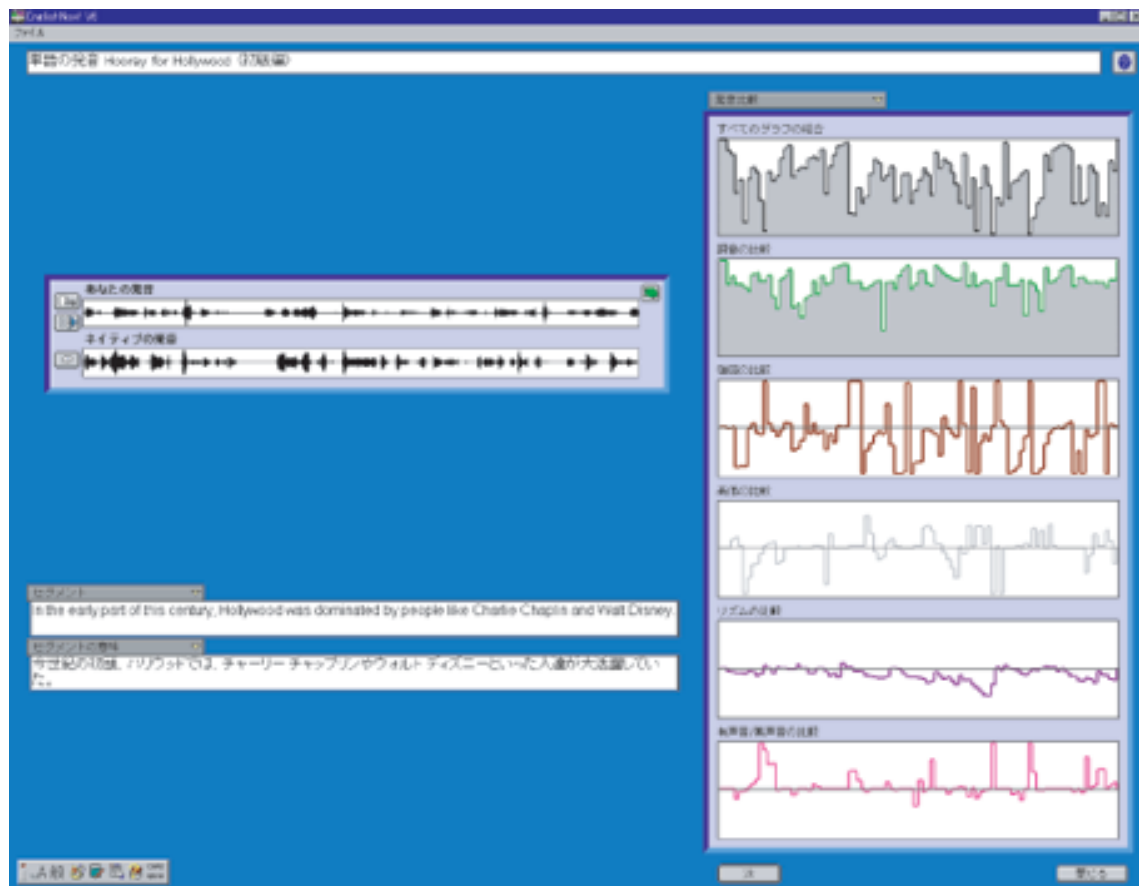


図. 2.6: 音声情報処理技術に基づく英語 CAI システム:English Now!

2.4 発音教育における本研究の位置づけ

近年の語学教育では、発音に誤りがあっても、通じればよいとするコミュニケーション重視の傾向がある。しかし、発音は音声コミュニケーションの基礎となる要素であり、間違った発音を直されずに進むと外国語の癖が残り、コミュニケーションに支障をきたしたり、相手に不愉快な思いをいだかせて円滑なコミュニケーションが進みにくくなることも多い。かといって、教員が個々の学習者に、絶えず face to face で発音指導を実施するには、多大な時間とエネルギーが必要である。また、学習者は自分の発音上の誤りを指摘されても、何が間違っていたのか、どう直せばよいのかわからないことが多く、教師も音声学を専門にしていない場合には、具体的な指導に窮することが多い [24]。

学習者の発音を波形やスペクトルで表示することは、現在の音響音声学 + PC というツールを用いれば可能であり、これらのツールを基にした発音学習アプリケーションは前節で紹介したように広く開発が行われている。しかし、既存の CALL システムで使っている現在の音声情報処理技術では、数理統計的な枠組みの上に構築されている。これは、学習者モデルが多数の学習者の平均分布的なモデルとして構築されることを意味する。また、教示方法（スコアリング）も教師戦略の平均的な方法論がモデル化されることを意味する。したがって、大規模データベースに頼り、音声技術者だけで構築してきた従来の CALL は、学習者の多様性及び教師の多様性に対応できない可能性がある。

本研究では、このような CALL の不安定性を考慮し、一般性を求めてきた従来の CALL 開発手法と異なって、特定の教師が自分の経験を生かし、学習者個人の学習状況に応じて、システムの特長・性能を調整し、常に「該当教師に最も似ている」柔軟なシステムを開発することを目指している。

第3章

本研究における先行研究

近年の音声情報処理技術の進展に伴い、日本語 CALL システムに関する研究は盛んに行われている。音声情報処理を用いた CALL システムは、機能的に、主に2種類の技術に支えられている：

1. 発音誤り検出技術 [25]
2. 学習者の発音を評価技術 [26, 27]

以下は、CALL システムがどのようにこれらの機能を実現するのかを、いくつかの関連研究例を通して紹介する。

3.1 日本語発音誤りの自動検出

小俣らは、音声認識技術を用いた、日本語自律学習システムのための音素発音誤りの検出手法を提案した [28]。この手法は、一般的の音声認識で利用されている音響特徴量として声道の形状特徴を反映したケプストラム情報だけではなく、声帯の振動情報を併用した発音誤り検出手法である。

3.1.1 学習者音声データベース

文部科学省科学技術研究費特定領域研究「メディア教育利用」で作成された日本語音声 DB [4] 中の韓国人母語話者 (32 名) 分を使用した。各話者は、日本語の難音音素ミニマルペア 115 単語を発声している。

3.1.2 学習者音声誤り分析

子音の置換誤りのみを対象とした。日本語母語とする 5 人に書起こしを依頼し、分析した結果は表.3.1 になる。

3.1.3 発音誤り検出法

発音誤り検出の流れは図.3.1 に示す。誤りやすい音声を含む発声を入力とし、正しく発声された場合と誤って発声された場合を仮定し、それぞれの尤度を算出する。尤度スコアは、声道の特徴を反映するケプストラム情報に基づくスコア (ケプストラムスコア: S_c) と声帯の振動情報を反映する有声・無声の度合いを表すスコア (有声・無声スコア: S_p) を用意し、それぞれを融合して最終的なスコア (S_f) とする。その計算式を式.3.3 に示す。

表 3.1: 韓国語母語話者の発音誤り傾向

| 正 | 誤 |
|------|------|
| /b/ | /p/ |
| /d/ | /t/ |
| /g/ | /k/ |
| /bj/ | /pj/ |
| /ɕ/ | /tʃ/ |
| /z/ | /ɕ/ |
| /ts/ | /tʃ/ |

$$S_c = (\text{正しい音素 } HMM \text{ に対する対数尤度}) - (\text{誤った音素 } HMM \text{ に対する対数尤度}) \quad (3.1)$$

$$S_p = (\text{正しい音素の } acpeak \text{ に対する対数尤度}) - (\text{誤った音素の } acpeak \text{ に対する対数尤度}) \quad (3.2)$$

$$S_f = \lambda_c S_c + \lambda_p S_p \quad (\lambda_c + \lambda_p = 1) \quad (3.3)$$

λ_c λ_p はケプストラムスコア, 有声・無声スコアに対する重み係数である. 得られた融合スコアを, 閾値と比較することによって, 対象の音素が正しく発音されているか, 誤って発音されているかを判断する.

3.1.4 検出精度

検出実験における強制切り出しで使用する音素 HMM には, 連続音声認識コンソーシアム 2000 年度版ソフトウェアの高精度音響モデル (性別依存の不特定話者 triphone モデル) を, MLLR によって話者適応したものをを用いる. 正解ラベルは, 日本語母語話者の評価において, 5 人中 4 人以上が正解としたものを「正解発音」, 5 人中 4 人以上が対立する誤りやすい音素で発音されたと判断したものを「誤り発音」として実験に使用する.

融合スコアの発音誤り検出性能を等誤り率 (EER) によって評価した. 表.3.2 に実験結果を示す. /b/ , /bj/ , /d/ , /g/ , /ɕ/ の同一調音構造の有声・無声子音の対立では, 有声・無声スコアを併用することにより, 等誤り率が平均で 3.8% 削減された.

3.1.5 まとめ

この研究で提案した音素誤り検出手法は, 音響特徴量として, ケプストラム情報だけでなく, 声帯の振動情報を併用し, ある程度の精度向上を果たした. しかし, その改良はわ

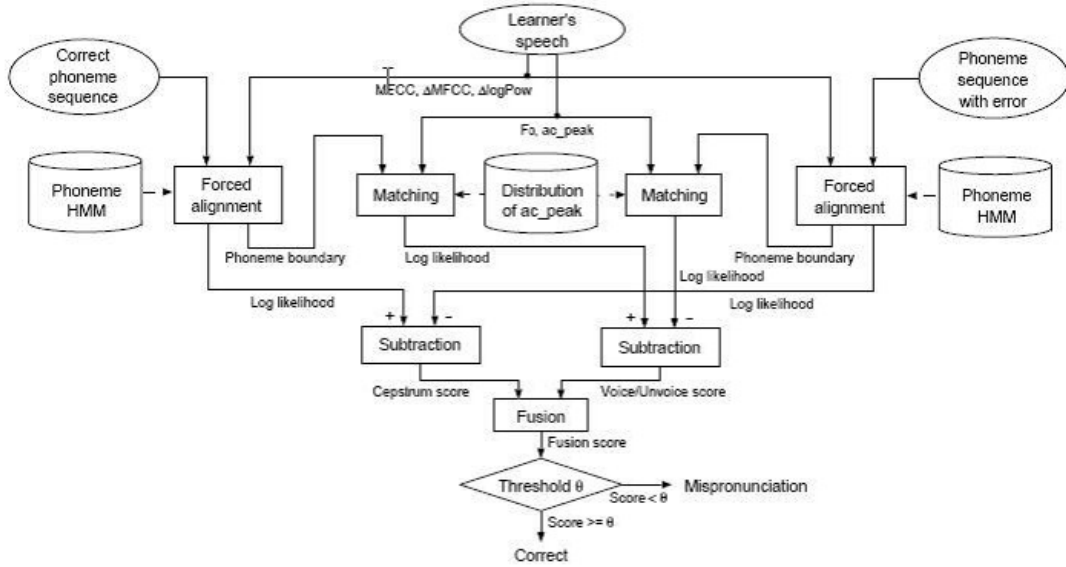


図. 3.1: 発音誤り検出システム

表 3.2: 等誤り率 (%) の比較

| | $S_{c\text{only}}$ | $S_{p\text{only}}$ | $S_f(\lambda_c)$ |
|------|--------------------|--------------------|------------------|
| /b/ | 17.5 | 22.8 | 15.1 (0.5) |
| /d/ | 18.5 | 17.2 | 15.0 (0.2) |
| /g/ | 10.8 | 27.8 | 10.0 (0.7) |
| /bj/ | 37.2 | 20.0 | 20.0 (0.0) |
| /dʒ/ | 24.0 | 30.0 | 24.0 (1.0) |
| /z/ | 12.5 | 44.8 | 11.0 (0.7) |
| /ts/ | 8.0 | 44.4 | 8.0 (1.0) |

ずかであり、音素の誤りパターンを一通りの置換誤りに限定しても、誤り率がまだ高いことが表.3.2 から明らかである。また、評価実験における強制音素切出しで使用した標準日本語音響モデルは、学習者に話者適応を行ったものの、100%正確に切出すことができないと思われる。

CALLシステム構築の際、データベースの精度に頼る音声認識技術のみでは、一部の誤りを検出できるものの、全ての学習者の発音エラーを検出できる学習支援システムの構築は困難であり、学習者個別の習熟度に応じて、検出すべき誤りパターンを絞る必要がある。

3.2 日本語学習者の文発話の発音評価法

多田らは留学生による日本語音声を対象とし、種々の音響的特徴量を教師の発音評価と相関が高くなるように組み合わせる方法を統計的に決める発音評価法の研究を行った [29]。

3.2.1 日本語教師による発音評価

「留学生による読み上げ日本語音声データベース」から男性話者71人、各話者5文を使用し、そのデータベースに付与されている日本語教師4人(A,B,C,D)による留学生の発声した日本語文音声データの発音を5段階(1~5, 5が最も良い)で評価したスコアの平均をターゲット値として利用する。

3.2.2 音響モデル

2種類の音響モデルを使用した

1. 日本語音節 HMM： ASJ データベースの音声試料(30名 4514文)と新聞記事読み上げ音声 JNAS(125名 13039文)を用いて学習した116種類の HMM
2. 留学生適応化音節 HMM： 日本語音節 HMM を留学生の日本語音声データ(45名 1021文)を用いて MAP 適応により追加学習したもの

3.2.3 各音響特徴量と日本語教師の評価との相関

日本語文の発音の評価は、尤度、話速などの組合せと、日本語教師によるスコアとの1文ごとの相関を求め、相関が高い方法を採用する。

対数尤度

発声された日本語文に対応する日本語音節モデルを連結して求めた対数尤度をフレーム数で正規化したもの (LL_{native}) と日本語教師の発音評価との相関, 留学生適応化音節モデルを連結して求めた対数尤度をフレーム数で正規化したもの ($LL_{non-native}$) と日本語教師の発音評価との相関を求めた。日本語音節モデルを使用した場合の相関は0.165, 留学生適応化モデルを使用した場合は0.077 となった。

事後確率

日本語音節モデルを使用した正規化した対数尤度と最適な音節列とのマッチング結果 (LL_{best}) を正規化した尤度を用い, 事後確率 L' (2つの尤度の差) を次式で求めた。

$$L' = LL_{native} - LL_{best} \quad (3.4)$$

日本語教師の発音評価との相関は0.462 となった。

音声認識結果

最適な音節列とのマッチング結果を用いて置換率 (Sub), 脱落率 (Del), 挿入率 (Ins), 正解率 (Cor), 精度 (Acc) を求めた。それぞれ日本語教師の評価との相関は-0.128, 0.085, 0.651, 0.097, 0.601 となり, 挿入率と精度で高い相関となった。

話速

入力された音声のフレーム数を N_{frame} , 発音辞書による音節数を $N_{syllable}$ とし, 話速 S を次の式で定義する。

$$S = \frac{N_{syllable}}{N_{frame}} \quad (3.5)$$

この式を用いて求めた話速と日本語教師の発音評価との相関を計算した結果, 相関は0.663 となった。

3.2.4 統計的発音評価法

様々な音響的特徴量を組み合わせる方法として重回帰分析を用いた統計的な発音評価では, 表.4.3の結果となる。重回帰分析の一例として, M 種類の音響特徴量の組合せを利用する場合, その特徴量を $x^{(k)}$ と表し, 目的値 (教師による採点スコア) を y とすると, 以下のようにモデル化する。

$$y = a^{(0)} + \sum_{k=1}^M a^{(k)} x^{(k)} \quad (3.6)$$

表 3.3: 文ごとの推定スコアと教師の評価の相関

| 組合せ | CLOSE | OPEN |
|-------------------------------|-------|-------|
| LL_{native} , パワー, モーラ長, 話速 | 0.719 | 0.699 |
| LL_{native} , 精度, 話速 | 0.715 | 0.701 |
| LL_{native} , 精度, パワー, 話速 | 0.719 | 0.703 |

n 個のサンプル値を用いて最小2乗法により、次式を最小とするように係数 $a^{(k)}$ を決める。

$$D(a, b, c) = \sum_{i=1}^n \left\{ y_i - \left(a^{(0)} + \sum_{k=1}^M a^{(k)} x^{(k)} \right) \right\}^2 \quad (3.7)$$

対数尤度 (LL_{native}), 精度, パワーの偏差値 (パワー), 話速を用いる場合は, 係数推定した結果, 以下の式になる。

$$y = 1.51 + 0.0192332 \times LL_{native} + 0.27134 \times \text{精度} \\ - 0.06494461 \times \text{パワー} + 0.39442 \times \text{話速} \quad (3.8)$$

3.2.5 考察

この研究は, 様々な音響と特徴量と日本語教師平均採点との相関を調べ, その最適な組合せを求めた。この手法はある程度, 日本語教師の採点「癖」を解析することができる。しかし, 4名の日本語教師による評価の平均をターゲット値とすることは, 教師の個人性をなくしてしまう。各教師の評定戦略が異なるため, それぞれ個別に対処が必要があると思われる。

3.3 CALL構築を目的とした学習者音声データベースの構築

これまで紹介した先行研究において, CALLシステムに使用した音響モデルは, いずれも, 母語話者による大規模データベースから学習した音響モデルをベースに, 学習者ごとに話者適応を行ってきた。また, 3.1節で紹介した研究では, 学習者誤りを分析するために, 学習者音声DBを使用した。したがって, 現在のCALLシステム構築手法では, 高度な学習者音声データベースがシステム精度向上に重要である。

峯松らは日本人学生による読上げ英語音声データベースの構築を試みた[30]。このDBが外国語発音学習支援環境構築における各種側面からの要請を考慮していることは, 新規であると言える。

3.3.1 DBの特性

音声処理技術的要請

1. 学習対象言語（方言）は米語とする
2. 学習者の年齢としては大学生，大学院生を対象とする
3. 音響的歪みのみを対象とし，言語的歪み（英作文時の誤用など）は対象外とする
4. 特定の学習者に見られる「特異的な」音響的歪み，及び各学習者に見られる「一時的な」音響的歪みはなるべく排除する

教育的要請

分節的側面と韻律的側面に分かれ，各側面に焦点をあてた発声リスト構築を行なう．

英語教師による発音評定ラベリング

発音評定者としては，英語音声学の知識を有し，対象言語を母語とし，日本人学生に対しての教育経験を持つ，対象言語の語学教師であることを必要条件となっている．

3.3.2 DB構成

1. 話者： 男性 100 人，女性 101 人
2. テキスト： 全文から 8 サブセット（各々約 120 文），全単語から 5 サブセット（各々約 225 単語）を構成し，1 サブセットずつを一人当たりの発声タスク量とする
3. テスクが大きいいため，一部を選択して，評定ラベリングを行った．

3.3.3 考察

このデータベースは，語学教育者と音声情報工学者と協力しながら，語学教育と音声工学の各種側面を念頭において，幅広い習熟度分布を網羅する日本人英語発声を収録した．学習者音声の特性を分析するに非常に役に立つことは明らかである．また，経験を持つ英語教師による評定データは，音声工学者が CALL を構築するときの参考になり，ある程度教育現場の知識をシステムに導入することができる．しかし，本 DB のすべての学習者の特性を一般化し，また全ての教師による評定を標準として CALL システム構築を行う研究者もいる [31]．このようなアプローチは，DB に含まれている全ての学習者の全体で見ると，性能向上につながる可能性あるが，学習者個人の特性をなくし，異なる評価尺度を持つ教師

の特性を無視してしまう可能性がある。したがって、従来のCALL開発手法では、学習者個人、ないし教師個人の特性に応じた、語学教育独特なニーズに沿ったCALLシステムの構築は不可能である。

3.4 まとめ

本章では、CALLシステムのための発音誤り検出法、発音評価法、及びCALLを目的としたデータベース構築法に関する研究を紹介し、それぞれの特長と欠点を述べた。次章以降は、これら先行研究を踏まえ、本研究で開発した新規なCALLシステムの構築手法とその評価について述べる。

第4章

本研究で提案するCALLシステム

4.1 提案する CALL システムの特徴

すでに述べたように、従来の CALL 開発手法は、学習者レベルの多様性を無視してしまい、学習者個人の特性に応じた適切な診断・評価を行うことが不可能である。本研究では、20年日本語教授歴を持つ豪グリフィス大の日本語教師と共同研究を行い、その教師の知識と経験に基づき、その教師の日本語クラスに所属している日本語学習者の特性を分析し、その教師を模擬するという形で、学習者レベルや特性に応じた CALL を開発することを目的として、本研究を発足した。言い換えれば、ある特定の日本語教師が様々なレベルの学習者に対して行う発音指導をシミュレートする CALL システムの開発である。

4.2 システムの概要

本研究で開発した CALL システムはクライアント・サーバ方式であり、学習者がクライアント PC で例文を読み上げて録音し、サーバに音声を転送する。サーバは音声分析サーバ、Web サーバ、データサーバとして機能し、読み上げ音声、分析結果、学習者 ID、フィードバック文を管理し、学習者が Web ページで発音評価結果とフィードバックを確認できるようになっている。また、日本語教師は管理モードにより、ユーザ管理、オリジナル教材作り、学生のレベルに応じた発音誤り検出モジュールの構築などの教材開発ができる。学習者音声評価システムを図.4.1、録音ソフトの画面を図.4.2、学生の音声評価画面を図.4.3、日本語教師のための管理画面を図.4.4 で示す。

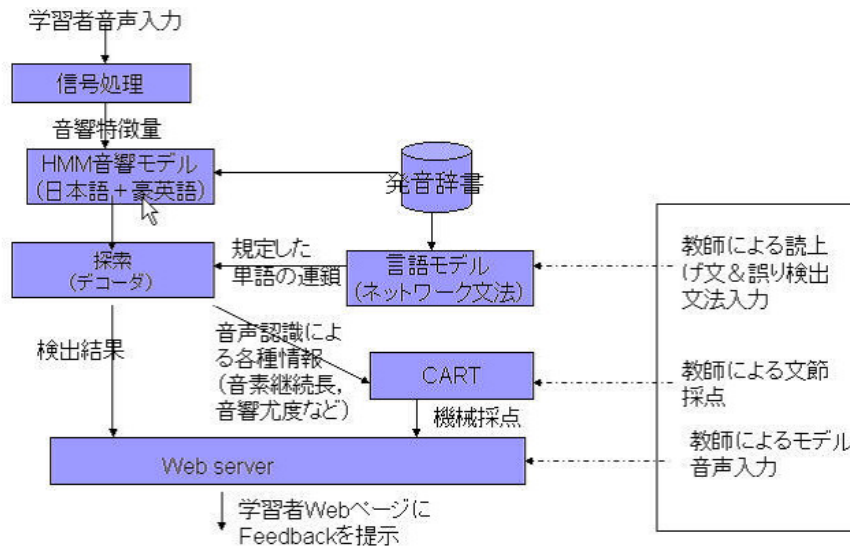


図. 4.1: 音声評価システム

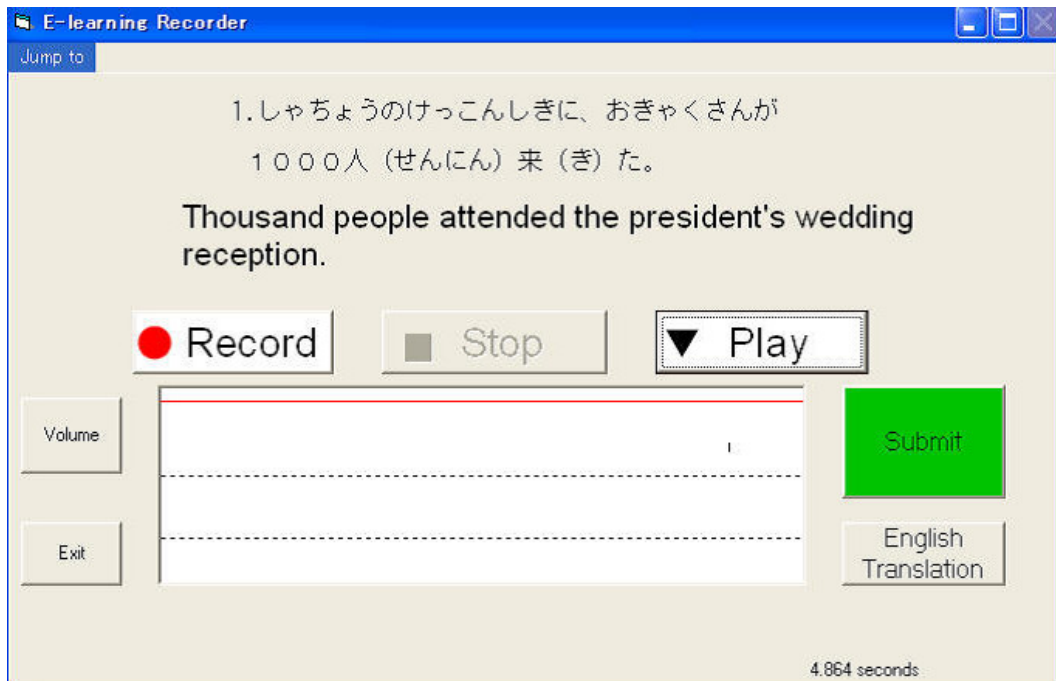


図. 4.2: 学習者用録音画面

15. まだ ちょっとわかりません

m a d a c h o q t o w a k a r i m a s e n
 m a : d a c h o q t o w a k a r i m a s e n

69% 69% 100%

Total 80%

m a d a

- The length of vowel is long.
- Shorten it considerably. Listen to the model pronunciation and try again.
- You lengthened the vowel between the two syllables.
 You might be assigning the pitch pattern 'High -Low' on the two syllables.
 Japanese words do not have a stress accent as English words do.
 Try to use the flat pitch pattern.
- おばさん(LHHH)、しゃちょう(LHH)、ごまい(LHH)

(※H=high, L=low)

図. 4.3: 音声評価画面

Scenario Management

Management tool Menu

- Account Management
 - >> [User-Account Management](#)
- Scenario Management
 - >> Topics ([List/Create](#))
 - >> Q & A ([List/Create](#))
 - >> Sentence ([List/Create](#))

List of the Sentences

[Create New Sentence](#)

| EDIT | ID | Sentence | Delete |
|----------------------|----|---------------------|--------------------------|
| Edit | 1 | ぎょうの てんきわ はれ でした | <input type="checkbox"/> |
| Edit | 3 | ぎょうわ あめが ふっています | <input type="checkbox"/> |
| Edit | 4 | ぎょうわ くもり です | <input type="checkbox"/> |
| Edit | 5 | かさ を もっています | <input type="checkbox"/> |
| Edit | 6 | どうきょう に りゅうがく したいです | <input type="checkbox"/> |
| Edit | 7 | ぎょう と に りゅうがく したいです | <input type="checkbox"/> |
| Edit | 8 | はい すぎ です | <input type="checkbox"/> |
| Edit | 9 | いいえ すぎ じゃ ありません | <input type="checkbox"/> |
| Edit | 10 | はい つくり ます | <input type="checkbox"/> |
| Edit | 11 | いいえ つくり ませ ン | <input type="checkbox"/> |
| Edit | 12 | はい | <input type="checkbox"/> |
| Edit | 13 | いいえ | <input type="checkbox"/> |

図. 4.4: 管理画面

4.3 発音エラーカテゴリーの定義

自動誤り検出するために、エラーカテゴリーを定義する必要がある。本研究では、まず日本語教師により、本システムが対象とする発音誤りを定義してもらった。まず自らの教育経験を振り返って定義させ、次に学習者コーパスを聞取して、そのリストの修正を行ってもらった。以下、その作業の様子を記述する。

4.3.1 予測される豪英語話者による発音誤り

分節音誤り

単純に英語と日本語の音素を比較することにより、考えられる誤りは以下の通りである。

<子音>

1. ϕ (両唇摩擦音 ふ) f (歯摩擦音) : さいふ
2. r (ら行 はじき音) r (流音, 接近音) / l (流音, 側音) : いろいろ らいねん
3. ts (歯茎破擦音) s (歯茎摩擦音) : つぎ すぎまたは、その混同で、[s] を [ts] と取り違えてしまうミスもある。

4. 有声子音の無声化

英語の閉鎖音は語頭で強い呼気を伴うため、声帯の振動開始時間にずれがおこり、語頭の有声閉鎖音が無声子音に聞こえてしまうことがある。その他、誤って、強い呼気が与えられた場合も、同じ結果となる。

- b (有声両唇閉鎖音) p (無声両唇閉鎖音) : ぼく ぼく
g (有声軟口蓋閉鎖音) k (無声軟口蓋閉鎖音) : つぎ つき* ([つ] を発音する呼気につられるため)

<母音>

1. 母音の混同

日本語の5母音の境界線が英語の各母音と重なるため、次のような母音の混同が起こることがある。

o ɔ e ε a æ u o
o u e a a ɔ u ə

2. つづりから来るミス

i ai (ローマ字表記を思い浮かべ、[i] を [ai] と発音する . e.g. [kai] in “kite”)

3. 連続母音の二重母音化

o.i oi “おい”という語が glide oi に発音される。
 a.i ai * [.] はモーラ境界。なお ui, ei は存在しない。

4. 母音脱落

ここでは、モーラ一つ分の長さが保たれず、次の音に融合してしまう場合を指す。
 しょう しょう (shiyou shoo)
 ですよ でしょう (desuyo deshoo)

< 拗音 >

拗音“CjV”は、music [mju:zik], beautiful[bju:tɪfl], cute [kju:t] にみられるように英語にも存在する音節だが、母音は、長母音の [u] のみと言う制約がある。したがって、日本語の中で頻度が高いが英語にない組み合わせであり困難とされる [rjV, kjV, gjV] でも、[rju, kju, gju] は、[a],[o] に比べて易しいと予測できる。

普通、[r] の音素自体の難しさもあって、/rjV/ が最も困難な拗音になる。拗音の誤りのパターンとしては、子音連続の誤りとして、通常、どちらかの音素を落として簡素化する次の二つが挙げられる。

CjV CV / jV [32]

それに加えて、モーラ感覚を習得していない学習者の場合、2モーラの長さにまで延びて、CiJV とする場合も見られる。その他、単語内の位置によって、促音挿入 CjVQ や長母音化 CjVV も起こりうる。

韻律的誤り

日本語のピッチアクセントに比べ、英語では、ストレスアクセントを置かれた音節が、重音節であることを原則とし、ピッチだけでなく長さ、大きさも増大する。従って、英語の韻律構造の転移から予測される長さの誤りは、ストレスアクセントパターンからの次の点が考えられる [33]。

1. ストレスが置かれた日本語の音節は、長さ、ピッチが変化し、軽音節の場合、促音挿入、長母音化が起こり、重音節となる。
2. ストレスが置かれなかった音節は母音が弱音化したり、消えたりする。また、無アクセントは英語に存在しないため、ピッチパターンも影響を受け、

3. 無アクセント型の語は、語のどこかにアクセントを置かれる。アクセントの位置としては、英語のアクセントルールの最も一般的なパターンから推測し、語末から2番目の音節にアクセントが置かれやすい。

などの点も考慮する必要がある。これらによって引き起こされる誤りは、時間長の制約があまりない言語の話者にとって、モーラ時間の制約を無視した母音、子音の長さに関する誤りとなる。

4.3.2 書き起こしによる誤り傾向分析

以上の考察に対して、実際に豪英語話者の日本語音声データを解析することで、本研究で開発する CALL システムが対象とする発音誤りを定義する。豪人日本語学習者初級 27 名、中級 19 名に、上記で考察した誤りの生起を考慮して作成された日本語例文 30 文を発音してもらった。それを、発音誤りの検出能力が比較的高い日本人母語話者を選んで、その全文の書き起こしを依頼した。30 文は、学習者のレベルに合わせ、仮名書きされたもので、学習者には、自分が考える正しい発音をするよう、指示した。書き起こし者は、音声工学や言語学を専攻とする学生から募集し、書き起こしの精度のレベルが音声研究者とほぼ一致する者を選んだ。各文を三人の母語話者が仮名表記によって書き起こし、二人以上が同時に指摘した誤りを「発音誤り」として定義した。書き起こしの結果は、音声研究者がランダムにピックアップしてチェックし、精度の低いもの（30 文中明らかなミスが 2 つ以上あるもの）は（精度の比較的高かった）他の書き起こし者に再度書き起こさせ、結果の精度向上に務めた。

4.3.3 エラーパターンリスト

日本語教師の経験及び書き起こしを元に、最終的に定義されたエラーパターンを、発生頻度の高いと思われるものから並べたものが表 4.1 である。ただし、ピッチ自体の誤りは、今回作成中のシステムでは対象外としたため、ここではリストに含めていない。

表 4.1: エラータイプとその頻度 (高1 低19)

| | |
|----|---|
| 1 | 促音挿入 e.g. hito hitto |
| 2 | 長母音の単音化 e.g. a aa, u uu |
| 3 | 促音の脱落 e.g. itte ite |
| 4 | 短母音の長音化 e.g. aa a |
| 5 | 拗音の simplification CyV (拗音) CV りゅ る/ゆ |
| 6 | 拗音における母音挿入 CyV (拗音) CiyV りゅ りゆ |
| 7 | 子音置換 R (flap in ら行) l (lateral approximant) |
| 8 | 子音置換 ts (alveolar affricate) s (alveolar fricative) |
| 9 | 母音の置換 e.g. [o] in "moss" i ai |
| 10 | 鼻音の挿入, 脱落 |
| 11 | 子音置換 F (bilabial fricative in っ) f (dental fricative) |
| 12 | 語頭での有声子音の無声化, 摩擦音の破擦音化 |
| 13 | 連続母音の二重母音化 o.i oi |
| 14 | 無声子音間, 文末の母音の無声化 |
| 15 | しょう しょう ですよ でしょう |
| 16 | ください ku:dsai |
| 17 | d rの置換ミス |
| 18 | 促音「っ」と長音の置換 |
| 19 | 長音と促音「っ」の置換 |

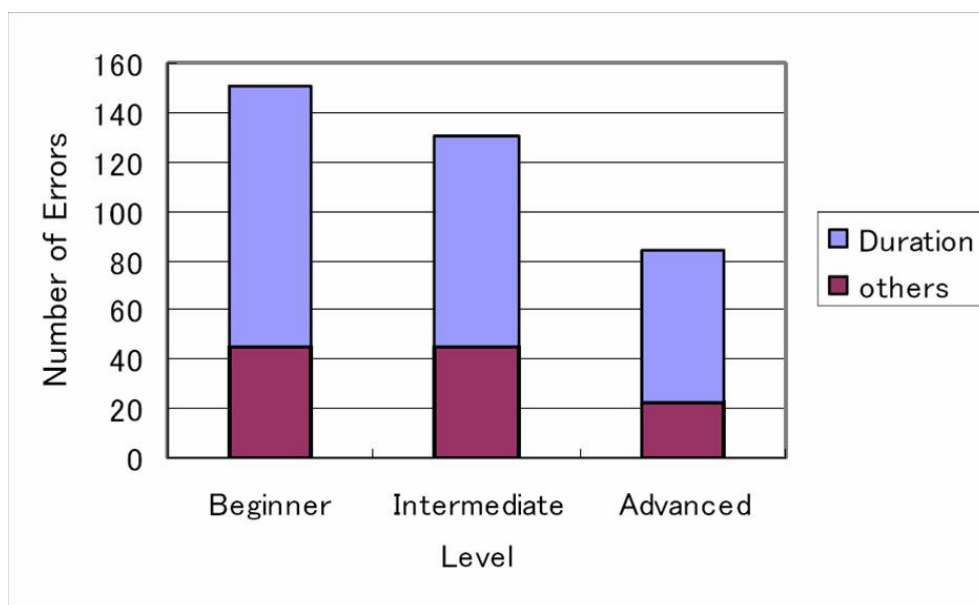


図. 4.5: 学習者レベル別誤りの総数

これとは別に、共同研究者である日本教師が、読上げテキストから10文を使って、初級、中級、上級の各6人の学習者に発音してもらい、エラー分析を行った。

初級、中級、上級の学習者は、以下のように定義する：

1. 初級： 日本語学習暦1年未満
2. 中級： 日本語学習暦2年未満，大学以外（高校，短期留学）での学習暦1～2年
3. 上級： 日本語学習暦4年以上，日本滞在1～2年

集計した結果は、図.4.5 に示すように、モーラ継続長に起因するエラーが各級において7割以上あった[34]。しかし、現在の音声認識技術では、継続長の情報は十分モデル化されていないため、通常の音声認識処理のみでは十分に誤り検出できない可能性がある。

4.4 学習者依存型誤り検出ネットワーク文法

発音誤りの検出は、従来より広く利用されているネットワーク文法を用いて行なった。例えば、前節で定義された発音誤りパターンより、「きいろいかさを」に対しては、図.4.6 に示すような発音のバリエーションが可能であり、各々に対してそれを受理するためのネットワークを構成する。そして音声認識技術によって、どのパスが実際の発声であったのかを同定することになる。当然、パスに含まれない発声であった場合にはその誤りは検出できない。任意の誤りを同定するためには、入力音声を「任意の平仮名列」としてネットワーク

文法を構成すればよいが，網羅する事象数が急増するため，自ずと検出率は低下する．そのため，適度に幅をもった文法が要求される．

このようにして，誤り予測ネットワーク文法が作成されるが，更に，日本語教育者がネットワーク文法の各項目に制約を加えた．上記の誤りパターンが起こりうる箇所（例えば，文節頭のみにかかる，など）を日本語例文の中で指定することにより，ネットワークの肥大による誤検出を回避した．以下のように指定は文節末に置かれたが，1と2に関しては，語末から2番目の音節に起こると言う制約があるため，文節末尾の助詞の手前に置いた（10～19は，記号操作のため，A～Iと書き換えられた）．

例： わたし(1,2)の こーこーで(4JH) いっしょに(3I,5,6) しゃしん(2)を(5,6 A) とりましょー(4,7)

予測される誤りパターンというのは，習熟度依存，更には学習者依存の傾向が高い．その意味において，現場の教師が学習者の習熟度に応じて適切な誤りパターンを埋め込む形でネットワーク文法を web 上で構成できるよう，GUI も整備した（図.4.7）．こうすることで，誤り検出精度の向上も期待できる．従来の e-learning 教材の場合，技術的問題の解決は開発者のタスクであり，現場の教師は構築されたシステムのユーザとして位置づけられることが多いが，本システムでは現場の教師の積極的な働きかけによっていくつかの技術的課題の解決を図ることにした．このような環境を整備することによって，教師の持つ，システムの内部動作についての知識も自ずと向上するものと期待される．

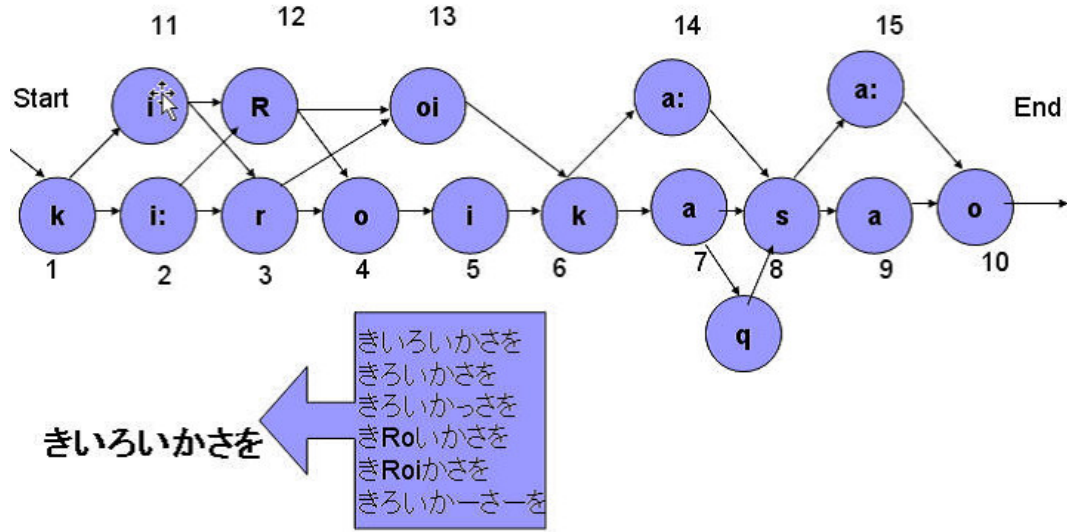


図. 4.6: ネットワーク文法



図. 4.7: ネットワーク文法生成 GUI

4.5 音響モデルと後処理手法

本研究で同定した発音誤りは、本来あるべき音が無くなる（脱落）、不必要な音が入る（挿入）、異なる音で代用している（置換）の3種類に分類される。前者2種類の誤りに関しては、ATR 日本語多数話者音声データベース（話者 4,130 名、性別非依存）で学習した HMM モデルを用いる。一方、置換誤りは、例えば「ふ」という平仮名を豪英語の /fu/ として発声するなどの誤りがある。この場合、日本語の /f/ ではなく、豪語の /f/ になっていることを指摘する必要があるため、豪語の音響モデルが必要となる。そこで、豪語音声データベース ANDOSL (Australia National Database Of Spoken Language) を用いて音響モデル構築を行なった。音響分析時のパラメータは表 4.2 に示す。

音響モデルは 43 音素、豪語音素 47 音素（置換誤りを検出ために使う音素は /r/, /f/, /oi/, /ai/ の 4 つのみ）からなる 5 状態 3 分布の left to right 型の HMM である。モノフォンモデルは、各状態につき 16 個の混合ガウス分布を持つ。トライフォンモデルは、状態共有（2000 状態）がなされており、各状態の混合ガウス分布は 16 である。なお、前節に述べたように、音声認識で使われる音素の音響モデルは、例えば、短母音と長母音の区別の精度が低くなる傾向がある。その理由は、音響モデルは音声のスペクトル特性を主に用い、その音素が本来どの程度の時間長を有するのかについての特性は十分にモデル化されないためである。その結果、長短の違いが音素の違いを生む場合に、その弁別精度が落ちてくる。予備実験の結果、学習者音声における非常に短い母音を長母音として認識する（或はその逆の）場合が比較的多く散見されたため、ここでは、以下のヒューリスティックな手法を後処理的に用いて、音素継続長に関する検出誤りの低減を試みた：

1. 長さ 120[msec] 以上の短母音を長母音とみなす
2. 長さが 100[msec] 以下の長母音を短母音とみなす
3. 長さが 60[msec] 以下の促音挿入と判断したエラーを無視する
4. 子音の長さに関しては、/p/, /t/, /k/ であれば、100[msec] 以上、/s/, /sh/ は 150[msec]、それ以外の子音は 120[msec] 以上であれば、促音が入っていると判断する。

表 4.2: 音響分析条件

| | |
|---------|-------------------------|
| サンプリング | 16[kHz], 16bit |
| 分析窓長 | 25[msec], ハミング窓 |
| 分析間隔 | 10[msec] |
| 特徴パラメータ | MFCC+ MFCC+ pow (25 次元) |

4.6 まとめ

本章では、学習者に焦点を当てた CALL システムの開発手法を提案した。本システムでは現場の教師の積極的な働きかけによって、その教師の知識と経験を生かした検出機能を実装した。また本システムは柔軟なインターフェースを持ち、特定した任意の教師の知識と経験を簡単に導入し、学習者一人一人の習熟度に応じて、システム機能をチューニングすることが可能にした。次章は、本システムにたいする種々の評価実験について述べる。

第5章

評価実験

5.1 発音誤り検出評価実験

5.1.1 使用した音声データと正解ラベル

豪日本語学習者46名による30文の読み上げ音声を評価実験用データとして利用した。また、発音誤りラベリングとしては、第4章で述べた書起こし作業で、書起こし者3人のうち2人以上が一致した書起こしをラベルとして使用した。

5.1.2 使用したネットワーク文法

下記3つの文法を使用して、モノフォンとトライフォンで評価実験を行った。日本語教師によるマニュアル操作による文法構築が行なわれているはG3である。

1. ラベル1: G1: エラー対象となる音素が、単語のどこに位置するのかに関する制限を加えずに、19種類のエラー全てを文全体に適応する。
2. ラベル3: G2: 対象音素の単語内位置に関する制限を加えて19種類のエラー全てを文全体で考慮する。
3. ラベル3: G3: 対象音素の単語内位置に関する制約を入れ、文節毎に、日本語教師が必要十分なエラーのみ指定する。

5.1.3 システム精度

まず、3つの文法の評価データの正解ラベルに対する発音エラーカバー率を求めた。結果を表5.1に示す。G1, G2, G3となるに従い、カバー率は低下しているものの、低下の割合は非常に小さい。ネットワーク化によって複数の音素群からの同定作業が行なわれることになり、例えば図4.6のように1つの文節に対して、複数の発音形式が得られる。G1の場合、文節あたりの発音形式数は平均182、G3の場合は平均7である。発音形式数を大幅に低減していることが分かる。

次に、G1~G3の各文法を用いて、後処理を行う前と行った後の検出精度をそれぞれ図5.1と図5.2に示す。

表 5.1: 各文法のカバー率

| | G1 | G2 | G3 |
|------|-------|-------|-------|
| カバー率 | 92.3% | 90.5% | 89.6% |

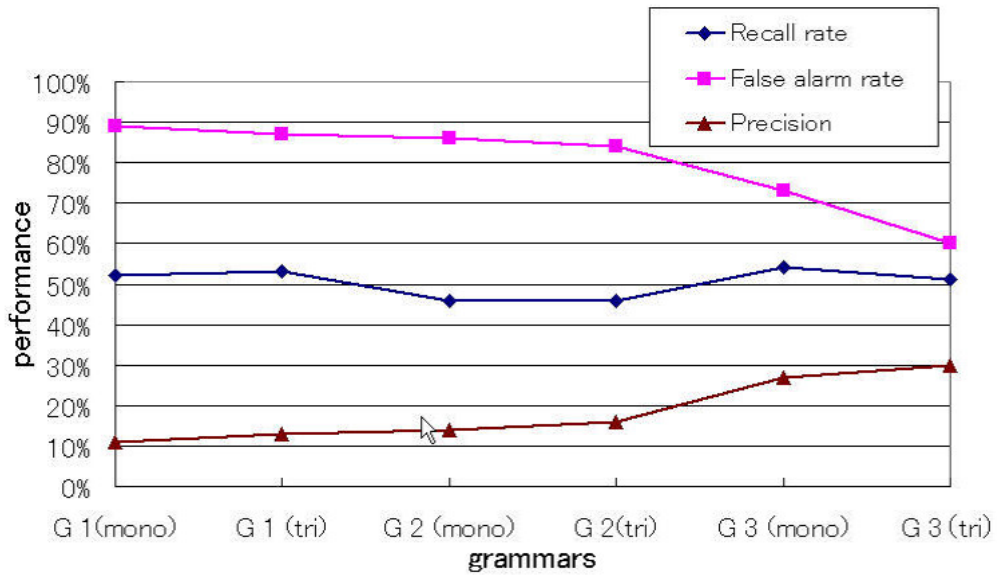


図. 5.1: システム精度 (後処理を行わない)

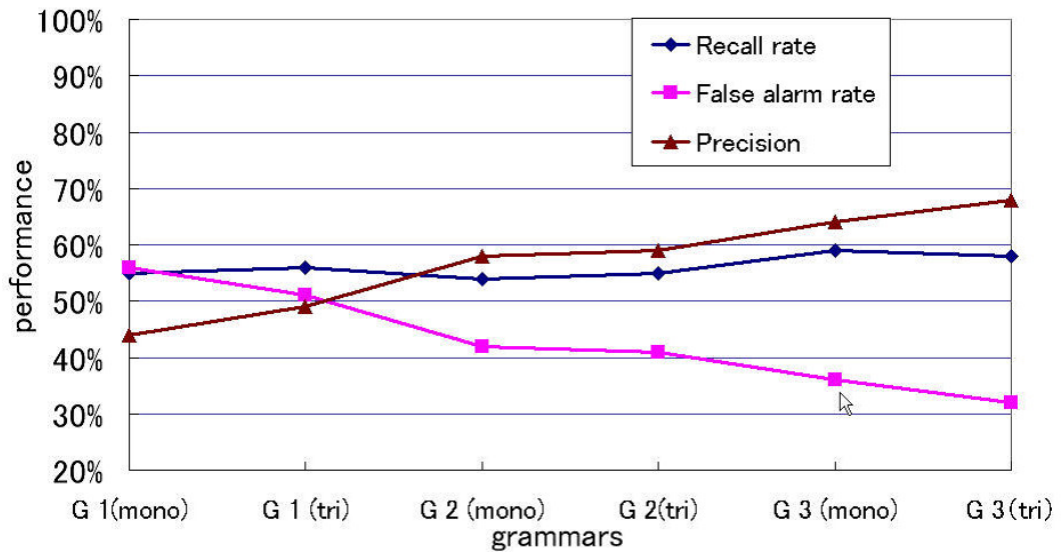


図. 5.2: システム精度 (後処理を行った後)

表 5.2: 正解書起こし基準による認識率変化

| | ラベル1 | ラベル2 | ラベル3 |
|-----------|------|------|------|
| 再現率 | 52% | 65% | 75% |
| 精度 | 73% | 69% | 56% |
| F-measure | 61% | 67% | 64% |

5.1.4 異なる正解ラベルによる精度の変化

前節における評価は、書起こし協力者が3人のうち2人以上が一致した書起こしを正解ラベルとしていた。しかし、2人以上が同様な発音誤りを見逃することなどによる正解書起こしの誤差も否定できないため、以下のような3つの正解ラベルに基づいて、検出精度の変化を求めた。

1. ラベル1: 3人のうち1人以上が判定したエラーを実際のエラーとする。
2. ラベル2: 3人のうち2人以上が判定したエラーを実際のエラーとする。
3. ラベル3: 3人とも判定したエラーを実際のエラーとする。

5.1.5 結果と考察

誤り検出評価実験で得られた最大の検出率は、G3と後処理を導入した場合は69.0%であった。指定した音素誤りパターンの全てを検出できるCALLシステムにおいては、十分な検出精度を持っている。しかし、False alarm率はG3と後処理を導入した後も34.5%となり、また改良の余地がある。図.5.1と図.5.2で表示しているように、いずれの場合、G3を導入することによって、精度を向上していることが分かる。また、後処理によって、検出精度、特にfalse alarm率が大幅に低減することは、第4章で検討したように、音響モデルの中に十分反映されない継続長の情報に関する発音エラーは、通常の音声認識処理のみでは検出することが困難であるといえる。

表.5.2で示すように、3種類の正解ラベルによる精度の変動が見られ、Precisionが高いラベルが逆にRecall率が低くなり、両者を一つの尺度にまとめるパラメータF-measureを求めたところ、システム精度評価で使っていたラベル2が最もいい評価となる。F-measureの観点から、正解ラベルをラベル2にすることは妥当であったと言える。

5.2 教師による評定との相関分析

3.2節で述べた先行研究では、日本語教師間の評価の相関は1文ごとで平均0.603であった。(3人の平均スコアと他の1人のスコアとの相関は、1文ごとで平均0.701であった。

また、機械と4人の教師の評価の平均との相関は0.703であった。) 教師間の評価の相関が低いことは、それぞれ独自の評価尺度を持っていることを示している。その研究は、4人の教師の評価の平均をモデル化したが、本システムでは、教師間の評価の差異を解明するために、特定の教師をモデル化し、当該教師を含む日本語教師間の相関や機械と各教師との相関を求めた。

5.2.1 日本語教師による文節評定データの収集

採点用音声データ

収録したすべての読上げ音声を文節ごとに採点することは、現実的に不可能なため、以下のようにデータ選択を行った。

1. 学習者選択：初級，中級学習者の中から，さらにそれぞれレベルをA，B，C三段階に分け，各段階の学習者を男女1名ずつ計6名を選ぶ。
2. テキスト選択：30文から，それぞれのレベルに応じて，最も誤りが起きやすい10文を選ぶ。

文節採点用の音声には，合計636個文節が含まれている..

採点者

モデル化対象とした日本語教師以外に，オーストリアにいる日本語教師4名（全員日本人）に評定を依頼した。各教師の日本語教育背景は表.5.3で示す。

音声提示手法

音声認識で行われている文節に関する評価（音響尤度・音素継続長など）は，基本的に前後の文節に影響されずに，限定されているセグメントのみを対象としている。その一方，

表 5.3: 採点者背景

| 教師 | 滞豪年数 | 日本語教育年数 |
|-------|------|---------|
| モデル教師 | 17年 | 20年 |
| 教師1 | 13年 | 13年 |
| 教師2 | 13年 | 14年 |
| 教師3 | 8年 | 15年 |
| 教師4 | 6年 | 4年 |

表 5.4: 各教師が文節のみを提示する条件での文節採点と文全体を提示する条件での文節採点との相関

| 模擬対象 | 教師 1 | 教師 2 | 教師 3 | 教師 4 |
|------|------|------|------|------|
| 0.80 | 0.79 | 0.71 | 0.60 | 0.57 |

日本語教師が学習者の音声を評価する際、文全体の情報を把握した上で、文の中における文節音や音素を評価することは普通である。したがって、本実験では、文節音声を採点する際、文コンテキストが文節評価への影響を考慮し、以下の2通りの方法で音声を提示し、各教師の文節評定データを収集した。

1. 文全体の読上げ音声を日本語教師に提示し、文節ごとに採点を行う
2. 音声を文節ごとに切出し、孤立した文節音声を採点者に提示し、それについて採点を行う

5.2.2 各教師が異なる音声提示条件での評定比較

前節で述べた条件で収集した評価データを、教師ごとに、同じ教師が文節のみを提示する場合の採点と文を提示する場合の文節採点との相関を求めた。その結果は表.5.4 に示す。同じ教師でも、異なる条件で文節評価の差が見られ、文節前後の情報が文節評価に影響されていることが示されている。また、相関が教師ごとに異なり、各教師の評価尺度において、文全体の情報による影響の個人差があることを示している。

5.2.3 日本語教師間の文節採点相関

当該教師を含む5名の日本教師間の文節採点相関を調べた。文節のみ提示する場合は、教師間の文節評価の平均相関は0.60となり、文全体を提示し文節採点を行う場合は、0.66であった。この5名の教師は、各自の採点戦略を持っていることが分かる。

5.2.4 CARTによる文節評定推定

本システムはCART (Classification and Regression Trees) を用いて、学習者音声に対して、文節の評定を推測する。

表 5.5: システム評定と各教師評定との相関関係

| 模擬対象 | 教師 1 | 教師 2 | 教師 3 | 教師 4 |
|------|------|------|------|------|
| 0.72 | 0.68 | 0.55 | 0.63 | 0.45 |

CART の説明関数

1. 文節に関する情報：文節で起きる発音誤り数，平均音響尤度，母音継続長の平均と分散 [35]，単語信頼度 [36]
2. 文全体に関する情報：文全体に起きる発音誤り数，平均音響尤度，文全体における母音の継続長の平均と分散，平均単語信頼度
3. その他：学習者レベル

モデル化用試料

5.2.1 節で述べた文全体を提示する条件で，学習者音声に対して，模擬対象となる日本語教師による文節音声評定データ。

実験条件

話者・テキスト Open（学習者 12 名のうち，8 名の音声を学習用試料，4 名の音声をテスト試料とする。また，読上げテキスト 10 文のうち，9 文を学習用試料，1 文をテスト用試料とする組合せで，文節音声の評価を推定する。）

5.2.5 システム評定と教師による評定分析

本システムは特定した教師の評価を模擬し，その教師を支援することを前提に開発したため，本システムの評価が他の教師よりもその教師に似ていることを証明できれば，システムの有効性を示すことができる。

システム評定と教師の評定との相関

システムによる文節評定と教師の評定を調べ，表 5.5 に示す。システムによる評定と模擬対象教師による評定との相関は他の教師の評定との相関より高いことを確認した。

表 5.6: 模擬対象教師による評定との相関関係

| システム | 教師 1 | 教師 2 | 教師 3 | 教師 4 |
|------|------|------|------|------|
| 0.72 | 0.70 | 0.63 | 0.66 | 0.55 |

模擬対象教師による評定との相関関係

本システムの模擬対象となる教師の評定とシステムによる評定，また模擬対象となる教師の評定と他の各教師による評定との相関を求めた．結果は表.5.6 に示すように，システムによる評定が他の教師よりも模擬対象となる教師に似ていることを判明し，模擬対象となる教師にとっては，他の教師（教師 1～教師 4）よりも類似度が高いことが分かる．

5.2.6 結果と考察

模擬対象となる教師の評定を他の教師の評定およびシステム評定と比較し，その相関を求めた結果，システム評定がある程度高い信頼度を持つことを確認した．しかし，表.5.6 に示すように，相関値そのものに関しては，教師 1 とシステムがそれほど差が大きい．本システムより，模擬対象教師に似ている日本語教師の存在を完全に否定することはできない．相関のみでは，各教師の評定戦略を解明することができない．CART の各説明関数を調整し，それぞれのパラメーターがどのように各教師の評定に影響するのかを詳しく調べる必要がある．また，システムの拡張性を考慮し，全ての教師のモデルを作り，それぞれの特性を考察する必要がある．

5.3 学習者による主観的評価

本システムの有効性を調べるため，初級 4 人，中級 5 人，上級 5 人の豪人日本語学習者に実際に，使用してもらった．中上級者には 2 つのレベルを初級には初級のレベルのみを試してもらい，アンケート用紙で以下の項目についてシステムの評価を行った．

- 1) プログラムの使いやすさ
 - 1-1) 録音部分
 - 1-2) 評価部分
- 2) プログラムは面白かったか
- 3) 自分の発音の弱点が分かったか
- 4) フィードバックは分かり易かったか．
- 5) プログラムは自分の発音学習に役に立ったか．

1～5 で示してもらった回答を集計，平均し，表.5.7 のような結果が得られた．ただし，点数が低いほうが高評価に相当する．プログラムそのものの有益度，使いやすさに関して

表 5.7: アンケート結果

| | 平均値 | 標準偏差 |
|------------------|------|------|
| 1-1) Recording | 1.36 | .50 |
| 1-2) Results | 1.86 | .77 |
| 2) Fun | 1.29 | .469 |
| 3) Weak points | 1.57 | .76 |
| 4)How to improve | 1.86 | .66 |
| 5) Usefulness | 1.46 | .66 |

はどれも高い支持が得られた。しかし、1-2)と4)が若干低めなのは、前節分析したように、検出制度、特に false alarm 率の低減をまた改良する余地があることを示している。

5.4 まとめ

本章は開発した CALL システムに対して、3つの観点からシステム評価を行った。本システムが学習者個々のレベルに応じて、より適切な評価をしていることを証明した。特定した教師の評価を模擬するモデルは、他の教師よりもその教師の評価に似ていることを明らかにした。また、本システムは学習者から比較的に高い評価を得ている。したがって、本研究で開発した CALL は、教育現場に適応していると言える。ただし、各教師の採点「癖」そのものがまだ解明すべきところが多い。今回の評価実験において、一人の教師のみ模擬しモデルを用いた。他の教師を模擬した場合は、同様な性能を示せるかを検証すべきこととして思われる。このような検証を行うためには、モデル化とする教師がシステムチューニングに参加してもらう必要があるため、今後の課題とする。

第6章

結論

6.1 終わりに

平均的な学習者像を前提とし、教師の平均的な評価戦略を模擬する形で開発が行われてきた従来のCALL開発と異なり、本研究では、特定した教師の教科戦略を模擬し、学習者それぞれの習得度に応じて、柔軟な発音検出できる新規なCALLの開発と評価を行った。

第1章では、現在の発音教育の現状とその技術支援における問題点を述べ、それとともに本研究の目的を述べた。

第2章では、英語発音教育に関する研究の背景として、日本語と英語における差異、英語教育における発音学習の位置づけとその技術支援としての発音教育システムの概要について紹介し、発音教育の技術支援において求められているものと、発音教育における本研究の位置づけについて述べた。

第3章では、本章では、CALLシステムのための発音誤り検出法、発音評価法、及びCALLを目的としたデータベース構築法に関する研究を紹介し、それぞれの特長と欠点を述べた。

第4章では、学習者に焦点を当てたCALLシステムの開発手法を提案した。従来のCALLの場合、技術的問題の解決は技術者の仕事であり、現場の教師は構築されたシステムのユーザとして位置づけられることが多いが、本システムでは現場の教師の積極的な働きかけによっていくつかの技術的問題の解決を図ることにした。本システム柔軟なインターフェースを持ち、特定した教師の知識と経験を導入することが可能にした。

第5章では、開発したCALLシステムに対して、3つの観点からシステム評価を行った。本システムが学習者個々のレベルに応じて、より適切な評価をしていることを証明した。特定した教師の評価を模擬するモデルは、他の教師よりもその教師の評価に似ていることを実験的に示した。

6.2 今後の課題

今後の課題としては、まず第4章におけるヒューリスティックな手法を後处理的に用いて、音素継続長に関する検出誤りの低減の試みについては、厳密な理論にも基づいた手法ではないため、継続長以外でも、ピッチなどの韻律情報も考察し、検出精度向上につながると考えられる。

また、本研究で開発したシステムは、現在音素誤りの検出のみ対応しているため、日本語学習者が頻繁に発生するアクセントの不具合への対応も必要となる。第5章における CART による発音評価予測では、説明関数にアクセント型や F0 に関するパラメータを追加することによって、予測評定精度の改良が考えられる。また、各説明関数を調整し、教師の採点戦略にどのように影響を与えるのかを今後行うべきこととして考えられる。

システムの拡張性を考慮し、特定した教師以外の日本語教師の評価もモデル化し、比較実験を行うことも必要である。

参考文献

- [1] 国際交流基金, “ 海外の日本語教育の現状 - 日本語教育機関調査・2003年 - (概要版)”, 2003
- [2] 岡本佐智子, “ 日本語教師養成の現状と課題 ”, 北海道文教大学論集第六号, 2003
- [3] 小川直樹, “ 理屈で分かる英語の発音”, ノヴァ・エンタープライズ, 1999
- [4] 深澤俊昭, “ 英語の発音パーフェクト辞典”, アルク, 2000
- [5] 佐々木正人, 渡辺章, “ 「空書」行動の文化的起源 漢字圏・非漢字圏との比較 ”, 教育心理学研究, vol.32, pp.182-190, 1984
- [6] 田嶋圭一, 山田玲子, 山田恒夫, “ Perceptual learning of English syllable rhythm by elderly Japanese listeners,” 第16回日本音声学会全国大会予稿集, pp.103-108, 2002
- [7] 松本亨, “英語で考える本”, 英友社 (1968)
- [8] 浅間正通, “英語教育における「異文化コミュニケーション」の普遍的視点をめぐって”, 静岡大学情報学研究, vol.3, pp.1-10, 1997
- [9] R. B. Kaplan, “ Cultural thought patterns in inter-cultural education, ” Language Learning, vol.16, pp.1-20 (1996)
- [10] 文部科学省科学研究費補助金特定領域研究 (1) 「高等教育改革に資するマルチメディアの高度利用に関する研究」平成14年度研究成果報告書 (2003)
- [11] 壇辻正剛, “ 道しるべ: IT化時代の語学環境としてのCALL”, 情報処理, vol.42, no.10, pp.1001-1005, 2001
- [12] 峯松 信明, 仁科 喜久子, 中川 聖一, “ 外国語学習用読み上げ音声データベース”, 日本音響学会誌, Vol.59, No.6, pp.345-350, 2003
- [13] A. Neri, C. Cucchiaroni, and H. Strik, “ Automatic speech recognition for second language learning: How and why it actually works, ” Proc. Int. Congress of Phonetic Sciences (ICPhS '2003), pp.1157-1160 (2003)
- [14] International Phonetic Association,, “ Handbook of the International Phonetic Association”, Cambridge University Press (1999)
- [15] 川越いつえ, “ 英語の音声を科学する ”, 大修館書店 (1996)
- [16] 窪園晴夫, 太田聡, “ 音韻構造とアクセント ”, 研究社 (1998)

- [17] 中川聖一，“語学学習における音声・言語処理技術の利用”，電子情報通信学会誌，Vol.85，No.12, pp.942-943（2002）
- [18] L. Bachman, “Fundamental considerations in language testing,” Cambridge University Press, 1990
- [19] J. Bernstein, M. Lipson, G. Halleck, and J. Martinez, “Comparison of oral interviews and automatic tests of spoken language,” Language Testing Research Colloquium, 1999
- [20] Microsoft Corporation, “ENCARTA インタラクティブ英会話”，2000
- [21] メディア教育開発センター，Listen to Me!，NHK エデュケーショナル（2000）
- [22] Transparent Language, “English Now! Ver.6.0J,” スリー・エー・システムズ, 1998
- [23] 山田玲子，英語リスニング/スピーキング，<http://exp.atrcall.jp/>（2000）
- [24] Yamauchi, Y. Multimedia and EFL Classroom in the IT era, Kenkyusha, Tokyo, 2001
- [25] A. Raux et al., “Automatic Intelligibility Assessment and Diagnosis of Critical Pronunciation Errors for Computer-Assisted Pronunciation Learning,” Proc. ICSLP, pp.737-740, 2002
- [26] H. Franco et al, “Automatic pronunciation scoring for language instruction,” Proc. ICASSP, pp.1471-1474, 1997
- [27] 太田, 中川:“日本語学習者の日本語発音評価法”，日本音響学会春季研究発表会，2-1-13（2005）
- [28] 小俣修一他“韓国人の日本語発声に対する発音誤りの自動検出・評定法”，日本音響学会 2003 年秋季研究発表会講演論文集,2-8-8,pp.117-20, 2003
- [29] 多田 圭，中川 聖一，“日本語学習者の文発話の発音評価法”，日本音響学会秋季講演論文集，3-8-13, pp.867-868, 2005
- [30] 峯松信明，富山義弘，吉本啓，清水克正，中川聖一，壇辻正剛，牧野正三，“英語 CALL 構築を目的とした日本人及び米国人による読み上げ英語音声データベースの構築”，日本教育工学論文誌，Vol.27, No.3, pp.259-272,2003
- [31] 森一将，中川聖一，“日本人の英単語発音の評価法，日本音響学会講演論文集”，1-6-2，pp.211-212,2002
- [32] C. Tsurutani, “Acquisition of Yo-on (Japanese contracted sounds) in L1 and L2 phonology in Second Language,” vol 3. 27-47, 2004

-
- [33] Tsurutani, C. “ Speech rate normalization strategies of moraic timing -The study of English learners of Japanese- Journal of Japanese Language Teaching No. 119, 51-64, 2003
- [34] 鶴谷千春, 山内豊, 峯松信明, 羅徳安, “ 日本語音声の難易度 ~英語を母語とする学習者の場合~”, 日本音声学会全国大会予稿集, pp.51-56 (2006-10)
- [35] 毛利太郎他, “ 母音持続時間に着目した日本語 CALL システムにおける矯正フィードバックの考察 ”, 電子情報通信学会音声研究会, SP2002-191, pp.73-78 (2003-3)
- [36] Akinobu Lee et al, “ Real-time word confidence scoring using local posterior probabilities on tree trellis search,” Proc. ICASSP2004, Vol.I, pp.793-796, May 2004.

謝辞

音声試料の収集・分析をはじめとする数々のご協力に加え，日本語音声学に関するご指導下さった豪グリフィス大学の鶴谷千春先生，そして，東京国際大学の山内豊先生に深く感謝します．

本研究を進めるにあたり，峯松信明助教授，そして広瀬啓吉教授，には度重なる御指導を賜り，深く感謝致します．御多忙中にも関わらず，学会発表の練習や論文の推敲，研究の方向性など細かな気配りをして頂き，感謝の言葉では言い尽くすことも出来ません．

また，日頃の研究生生活において様々な面で支えてくださった高橋登技官，秘書の武田祥子さん，笹島恵さんに深く感謝致します．最後に研究のみならず，多くの面でお世話になりました広瀬・峯松研究室の皆様にも心より深く感謝致します．

ありがとうございました．

発表文献

- [1] 羅徳安, 峯松信明, 鶴谷千春, 山内豊, 広瀬啓吉, “ 英語話者を対象とした日本語 CALL システムにおける発音評価”, 日本音響学会秋季講演論文集, 3-P-14, pp.363-364 (2006-9)
- [2] 羅徳安, 峯松信明, 鶴谷千春, 山内豊, 広瀬啓吉, “ 日本語 CALL システムのための学習者発音分析とその自動評価”, 電子情報通信学会音声研究会, SP2006-78, pp.13-18 (2006-11)
- [3] 鶴谷千春, 山内豊, 峯松信明, 羅徳安, “ 日本語音声の難易度 ~ 英語を母語とする学習者の場合 ~”, 日本音声学会全国大会予稿集, pp.51-56 (2006-10)
- [4] C. Tsurutani, Y. Yamauchi, N. Minematsu, D. Luo, K. Maruyama, and K. Hirose, “ Development of a program for self assessment of Japanese pronunciation by English learners, ”Proc. Int. Conf. on Spoken Language Processing (ICSLP’2006), pp.841-844 (2006-9)

付 録 A

読上げ音声収録用テキスト

日本語教師が発音誤りの生起を考慮して作成した、読上げ音声収録用日本語例文（初級，中級用それぞれ30文）を以下に示す。

初級用読上げテキスト30文

1. あした また きても いいですか。
2. ゆうびん きょくで きてを 3まい かって きて ください。
3. よっかかん がっこうを やすみました。
4. うちから だいがくまで くるまで じゅっぶん かかります。
5. ごはんの まえに ビールを いっぼん のんだ。
6. その ざっしは せんえん でした。
7. ゆきが ふって やまも みちも まっしろに なった。
8. おばさんの つくった りょうりは とても おいしい。
9. あめのひに きいろい かさを さして だいがくに いった。
10. この こうえんは さくらが きれい なので ゆうめい です。
11. おねえさんは こうこうの とき じゅうどうと すいえいを していた。
12. この ようふくは わたし には おおきい ですね。
13. おきゃくさんに おちゃを だした。
14. どうして えいごを べんきょう しているん ですか。
15. きょうしつに あたらしい じしょを わすれた。
16. でんしゃに のって うみまで 行きますよ。
17. こんしゅう しゅくだいが 三つある。
18. わたしは ぎゅうにゆうが あまり すきじゃない。
19. じゅぎょうの あとで びょういんに 行くつもり です。
20. 日本ごが じょうずに なって 日本に りゅうがく したい。
21. いろいろな くにに りょこうした ことがある。
22. おばあさんの たんじょう日 だから ケーキを つくった。
23. ともだちと いっしょに けいざいの べんきょうを した。
24. その ちずは はっぴゃくえん くらい でしょう。
25. きんのうの よるは つきが きれい だった。
26. やすみは どうする よてい ですか。
27. テストの まえは よく べんきょう したほうが いいですよ。
28. バスの じこくひょうを かして ください。
29. つぎの しゅうまつは もっと ゆっくり しましょう。
30. まいしゅう どうようびは とても いそがしい。

中級用読上げテキスト 30文

1. おきゃくさんに ちょっと まって もらって ください。
2. きょうじゅうに はっぴょうの じゅんびを しなくちゃ いけない。
3. いしゃに しょうかいじょうを もらって びょういんに 行った。
4. かれは いつも じぶんの いけんを しゅちょうする。
5. そつぎょうしたら 日本の会社で しごとが したい。
6. いちりゅう きぎょうに しゅうしょく するのは むずかしい。
7. 日本りょうり では どうふと さしみが すきです。
8. そんな ひょうげんは 聞いたことが ない。
9. がんになる 前に きんえんして じびょうを なおしたほうが いい。
10. 出張が あるので 電車の じこくひょうを しらべて おいた。
11. ゆきが ふったら スキーに 行こうと 思う。
12. つまらない 本を よんで しゅうまつを すごして しまった。
13. どうも としょかんに めがねを わすれた らしい。
14. 社内りょこうで おんせんに 行きました。
15. おばあさんは ようが が できて きゅうに 来られなく なった。
16. 来週は こむので ホテルを よやくして おいた。
17. 一日中 一人で じゅうに あそびたい。
18. 六百メートル きょうそうで さんいに なった。
19. ぼくの 弟は すうがくの きょうしだ。
20. りゅうが わからない けど その 店員は おこって いる。
21. 来年 東京の 大学に りゅうがく しようと思ふ。
22. きゅうに やくそく した 人が 来れなく なって がっかりした。
23. それは きっと おねえさんの くつ でしょう。
24. おばさんが くれた しゃしんは きれいに うつつて いる。
25. 小さい とき きんぎょ 5ひきと 犬を いっぴき かっていた。
26. 日本人は しょくじちゅうに よく りょく茶を のむ。
27. えんりょ しないで もっと 食べても いい ですよ。
28. 学生の うちに ゆめを じつげん したい。
29. こうこうの 時の しんゆうは ゆにゆう会社の しゃちょうに なった。
30. びじゅつかんに きょうりゅうの えがある。