

英語母語話者を対象とした日本語CALL システムの構築と発音の評価*

◎ 羅 徳安（東京大・情報理工） 峯松 信明（東京大・新領域）
鶴谷 千春（グリフィス大・言語） 山内 豊（東京国際大・情報システム）
広瀬 啓吉（東京大・情報理工）

1 はじめに

近年、日本語学習熱が世界的に高まり[1], 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が開発されている。[2] 本研究では, 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の開発と評価を試みた。本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検出ができること, を念頭に置き, a) 当該日本語教師が考える発音エラーカテゴリ, 及び, b) 学習者毎の評定項目, 更には, c) 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行った。

2 開発した CALL システム

2.1 システム概要

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, サーバに録音した音声を送る。サーバは音声の分析を行うとともに, Web サーバと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。

2.2 エラーカテゴリの定義

本システムの開発に当たって, オーストラリア人日本語学習者の発音エラー分析を行った。学習者初級 27 名, 中級に 19 名に日本語例文 30 文を発声してもらい, 発音誤り検出能力が比較的高い日本人母語話者を選定し, その全文を平仮名で書き起こさせ, 「検出すべき発音誤り」を定義し,

以下の分析で用いた。また, 日本語教師 (第三著者) の考える「検出すべき発音誤り」と書き起こし結果から得られた誤りとを統合する形で, 表. 1 のように優先順位を付けて, 最終的に「本研究で対象とする発音誤り」19 種類を定義した。

表 1 : 対象とする発音誤り

1	促音挿入 hito → hitto
2	短母音の長母音化 a → aa, u → uu
3	促音の脱落 itte → ite
4	長母音の短母音化 aa → a
5	拗音の単純化 CyV → CV
6	拗音における母音挿入 CyV → CiyV
7	子音置換 R (flap) → l (lateral approximant)
8	子音置換 ts (alveolar affricate) → s (alveolar fricative)
9	母音の置換 i → ai
10	鼻音の挿入, 脱落
11	子音置換 F (bilabial fricative) → f (dental fricative)
12	語頭での有声音の無声化, 摩擦音の破裂音化 b → p, s → ts,
13	連続母音の二重母音化 o.i → oi
14	無声子音, 文末の母音の無声化 shimas → shimasu
15	母音脱落 shiyō → shō,
16	ください → ku:dsai
17	d → r の置換ミス
18	促音「っ」と長音の置換っ → ー
19	促音「っ」と長音の置換ー → っ

これとは別に, 初, 中, 上級者相手に行なわれたエラー分析では, 図. 1 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。[3] 継続長の情報は音響モデルの中に十分に反映されないため, 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。

*Development of a CALL system for Australian learners of Japanese pronunciation, by Dean Luo, Nobuaki Minematsu (The University of Tokyo), Chiharu Tsurutani (Griffith University), Yutaka Yamauchi (Toyo International University), and Keikichi Hirose (The University of Tokyo)

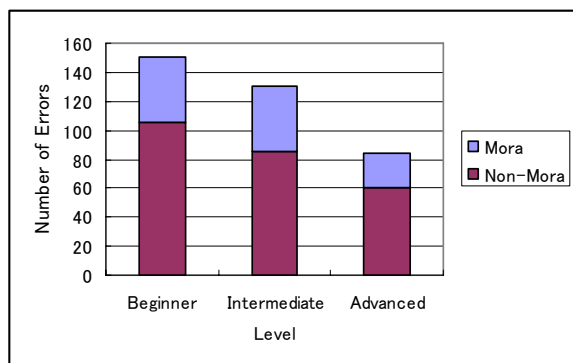


図. 1 学習者レベル別誤りの総数

2.3 学習者依存型検出ネットワーク文法

発音誤りは、予め予測される発音誤りを含むネットワーク文法を作成し、音声認識することで同定される。ここで、ネットワーク文法としては過不足なく発音エラーをカバーするものが望ましいが、エラーの多様性は学習者に依存する。結局、学習者の習熟度に応じてチェックすべき項目は異なるため、ここでは、教師が学習者の習熟度に応じて「読み上げる文」と「考えるべき発音誤り」を指定する形で、自動的にネットワーク文法が生成されるツールを整備することとした。こうすることで、ネットワーク文法の肥大化と、それに起因する同定率の下落を防ぐ事ができる。図. 2 にその、GUI を示す。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幾つかの技術的問題の解決を図ることにした。

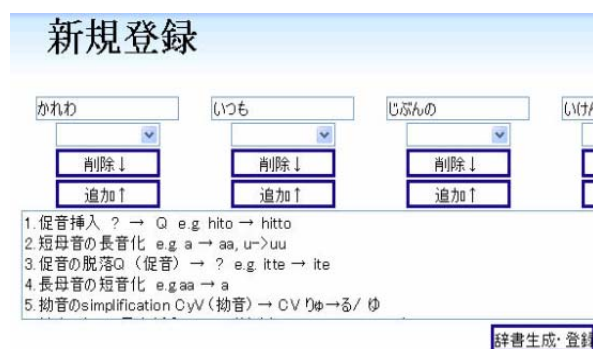


図. 2 ネットワーク文法生成 GUI

2.4 音響モデルと音素継続長に基づく後処理

ATR 多数話者音声 DB, 豪語音声 DB (ANDSOL) により学習した性別非依存 HMM モデルを使用した。既述したように、音響モデルは通常、音声スペクトル特性が主なモデル化対象となるため、/a/ と /a:/ の音響モデルによる長母音／短母音の区別のみでは、その弁別精度が十分でない可能性がある。ここでは、音素継続長に基づく後処理を導

入することによって、この検出誤りの低減を試みた。

2.5 システム精度評価実験

4 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。

G1: エラー対象となる音素が、単語のどこに位置するの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。

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。

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を指定する。

G4: G3 で処理した結果に対し、音素継続長に基づく後処理を行う。G1~G4 の各文法を用いた際の検出結果を図. 3 に示す。

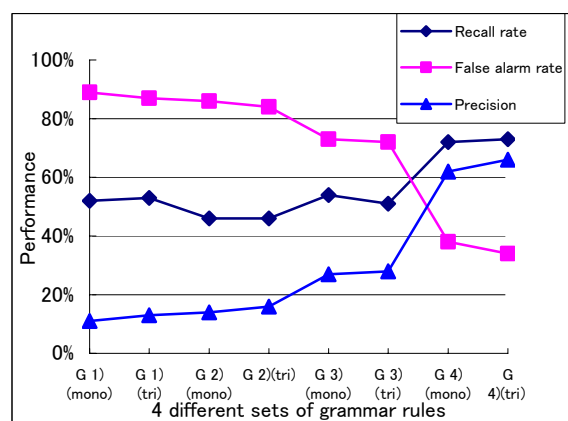


図. 3 システム評価結果

3 考察と今後の課題

英語話者のための日本語 CALL システムの開発と評価を行った。学習者依存型ネットワーク文法と音素継続長に基づく後処理を導入することによって、検出精度の向上を試みた。今後はシステムにおける音素識別の改良を検討しながら、評価点算出における主たる原因を解明するために、分析実験を通して、システム信頼度を向上させる予定である。

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] 小俣他, “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, 2003
- [3] 鶴谷他, “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備案, 2006