

日本語韻律教育の支援を目的としたオンラインアクセント辞書と 読み上げチューターの開発

峯松 信明[†] 中村 新芽[†] 鈴木 雅之[†] 平野 宏子^{††} 中川千恵子^{†††}

中村 則子^{††††} 田川 恭識^{†††} 広瀬 啓吉[†] 橋本 浩弥[†]

[†] 東京大学, 〒113-8656 東京都文京区本郷 7-3-1

^{††} 東北師範大学中国赴日本国留学生予備学校, 〒130117 中国吉林省長春市東北師範大学浄月校区

^{†††} 早稲田大学, 〒169-8050 東京都新宿区西新宿 1-6-1

^{††††} 慶応大学, 〒108-8345 東京都港区三田 2-15-45

E-mail: mine@gavo.t.u-tokyo.ac.jp

あらまし 日本語の韻律教育を支援すべく、日本語教師と協力して自然言語処理技術、音声言語処理技術を用いたオンラインアクセント辞書と読み上げチューターを開発した。日本語アクセントの学習・教育を困難にする理由は、アクセント変形のコンテキスト依存性に十分対応した教材が存在しないことにある。我々はアクセント変形が比較的規則的な用言に着目し、その活用、及び後続語を伴った様々な用言表現に対して、アクセント核の位置を視覚的、網羅的、聴覚的に呈示する辞書システムを構築した。また、イントネーション教育に関しても、任意の句に対してアクセント核位置及びピッチパターンを（用言に限らず）推定し、視覚的に呈示するチューターシステムを構築した。次に、日本語教師を対象とした主観評価実験、日本語学習者を対象とした客観評価・主観評価実験を行なった。実験の結果、構築したシステムの非常に高い実用性を確認することができた。OJAD (Online Japanese Accent Dictionary) という名称で無償公開しており [1]、現在、世界中の日本語教育現場で使われている。

キーワード 日本語発音教育, アクセント, イントネーション, 用言, 音声合成, アクセント推定, 評価実験

Development of an online accent dictionary and a reading tutor to support teaching and learning of Japanese prosody

N. MINEMATSU[†], I. NAKAMURA[†], M. SUZUKI[†], H. HIRANO^{††}, C. NAKAGAWA^{†††}, N.

NAKAMURA^{††††}, Y. TAGAWA^{†††}, K. HIROSE[†], and H. HASHIMOTO[†]

[†] The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656 Japan

^{††} Northeast Normal University, No.2555 Jingyue St, Changchun, Jilin, 130117 China

^{†††} Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050 Japan

^{††††} Keio University, 2-15-45, Mita, Minato-ku, Tokyo, 108-8345 Japan

E-mail: mine@gavo.t.u-tokyo.ac.jp

Abstract Through tight collaboration with teachers of Japanese, an online accent dictionary and a reading tutor are developed to support teaching and learning of Japanese prosody. In this development, techniques of natural language processing and spoken language processing are effectively applied. What makes teaching and learning of Japanese accent difficult is the fact that word accent of Japanese often changes due to its context. In this paper, firstly, we focus on verbs and adjectives because their accent changes are relatively systematic. A web-based dictionary system of presenting their accent changes visually, auditorily, and comprehensively is developed. Then, another tutoring system of displaying the pitch pattern and the accent nucleus locations of an any given sentence is developed for teaching and learning of Japanese intonation. Objective and/or subjective evaluation experiments are done for these systems by using teachers and students of Japanese as subjects. The results show very high educational effectiveness of the systems. Currently, the systems are freely available in public under the name of OJAD (Online Japanese Accent Dictionary) and they are effectively used in Japanese classes all over the world.

Key words Japanese pronunciation teaching, accent, intonation, verb and adjective, speech synthesis, accent estimation, evaluation experiment

1. はじめに

日本語を学ぶ場合、国際語である英語と比較して、その動機や環境は自ずと異なる。例えば日本研究を行なうための日本語学習という状況を除けば、「学習対象言語の書籍・論文を読み書きするため」という動機付けは少なく、相手と日本語で意志疎通を図る能力の獲得が主目的となる。英語の場合、意志疎通を図る相手が非母語話者であることの方が多く [2]、互いに外国語訛りがあっても許容される傾向にある。しかし日本語の場合は相手が日本人である場合が多く、より分かり易い発音＝より母語話者のような発音、となる。更に（英語学習と比較し）、日本語を学ぶ目的が日系企業への就職であるなど、実用的な動機も多い。その場合 public speaking 能力が要求され、東京方言学習を希望する者が多くなる。TV ドラマ・アニメ・映画に代表される日本語コンテンツの多くが東京方言によるものであり、東京方言学習の要望を高めている一要因となっている。このように、英語よりも発音教育、しかも、東京方言の発音教育の需要がある中で、特に、韻律教育に割く時間が限られている。教師本人が発音（韻律）教育を受けておらず、効果的な指導法を模索する教師も少なくない。本研究ではこのような教育事情を鑑み、オンラインアクセント辞書、韻律読み上げチューターを構築し、学習者と教師を対象に主観的、客観的評価実験を行なった。

2. 日本語教育が抱える問題とシステム開発指針

2.1 日本語発音教育が抱える問題

外国語教育では限られた時間内で教育する必要があるため文字に頼ることが多くなるが、その結果、発音に十分な時間を割くことが難しくなる。日本語発音を指導する場合でも、単音や特殊拍に焦点が当てられ、アクセントやイントネーション教育は見落とされがちである [3]。国内の大学で（生活言語として）日本語を学ぶ場合、単語アクセントがピッチアクセントであることを知らない学生も多い [4]。彼らは日本語音声浴びる環境にあり、教室外での発音能力の向上が期待できる。しかし地方大学の場合、地元住民の日本語音声に染まらぬよう、注意することもある [5]。国内でも韻律教育は疎かにできない。

日本語の場合、方言性がアクセントに出現し易い。そのため発音教育の中でも韻律教育が必要となる。そしてそのアクセントは（赤＋鉛筆→赤鉛筆など）前後のコンテキストによって頻繁に変化する。しかし、自らが発音（韻律）教育を受けていない日本語教師も多く、非母語話者の教師や、母語話者であっても地方出身者にとって、アクセント教育は難しい。仮に東京方言話者であっても、「発声のどこにアクセント核があるのか」を意識的に、正しく指摘できる話者となると限られる^(注1)。東京方言話者でアクセント感覚に優れた話者のみが日本語教師となっている訳ではないため、非母語話者教師であっても、適切に教えることのできる韻律教育法が求められている。

単語アクセントを調べる場合、NHK 日本語発音アクセント辞典 [7] が使われることが多い。これはスマートフォンや電子辞書でも使え、有用である。しかし、基本的に孤立単語アクセ

(注1)：アクセント制御は無意識的に行なわれるため、正しい発声はできても、アクセント核を意識的に把握することが困難となる。その一方で、不適切な位置にアクセント核がある発声に対しては、訛りを容易に感じ取る [6]。音声コミュニケーションに支障は無くても、音韻（音素）の把握に困難を示す音韻性 dyslexia が英語圏では多く見られるが、挙動としてはそれに類似している。

にほんごのぺんきょうは / むずかしいですが / だいすきです
(^①はアクセント核)

図1 フレージングとポージングに基づく韻律指導

ントを想定しており、「楽しむ」が3型であることは即座に分つても、「楽しみそうになったことがある」となるとお手上げである。その結果、コンテキスト中の単語アクセントを学習者が知りたい場合、アクセント感覚に優れた母語話者に聞くしか手段がないのが現状である。アクセント教育が、孤立単語のアクセント教育から脱皮できていない、という指摘も散見される [8]。

近年では Text-to-Speech (TTS) システムが広く流通するようになった。コンテキスト中の単語アクセントを知りたい場合、その句や文をタイプすれば読み上げてくれる。聴覚呈示でよければ Web 上にある各社の TTS システムのデモ版で用は足りる [9]。しかし、誰もが音声からアクセント核位置を把握できる訳ではなく、教育用のツールとしては不十分である。

日本語学習では「分かり易い発音＝母語話者のような発音」となる場合が多い。この時、単語アクセント以上に重要なのは、適切なイントネーションパターンをフレーズに付与し、フレーズ単位でポーズを置くことであると言われている [10]。例えば中国人日本語学習者の初級者は各モーラに四声を付与する傾向があり、また、単語単位での発声となる傾向もあり、文イントネーションに必要以上に起伏が生じ易い [11]。これに対して図1に示すように、文の意味を考えてフレーズ境界を定め、1) 各フレーズを「へ」の字を描くようなイントネーションにする、2) フレーズ間にポーズを置く、という指導を行なうことで、(日本人にとっての) 聞き取り易さは格段に向上する。当然アクセントによってピッチは上下するため、「へ」の字だけではアクセント的に不自然となる。その一方、アクセント核の付与を、全ての単語に対して学習者に求めることは負担が大きいの。より実践的な折衷案として [10] では、初級者向けに、「フレーズに最初に現れるアクセント核のみに注意を払い、その後の核は無視してよい^(注2)」という指導戦略で臨んでいる。教育用システムを構築する場合、全ての情報を呈示するのではなく、優先的に着眼すべき情報のみを呈示するなどのオプションも必要となる。

2.2 システム構築に向けた指針

日本語教育者（第四～第七著者）との協議の末、システム開発に対して以下の指針を立てた。

- システムが呈示する情報がほぼ 100% 正しいシステムなのか否かは、明確に区別する。
- コンテキスト中の単語アクセントを表示するシステムを開発する。比較的規則的にアクセント変形する用言に対しては、教育的配慮から別途取り扱う。
- 用言に対して、日本語教育で使われる基本活用形に対するアクセント変形を示す。活用によるアクセント変形の様子が網羅的に分かるように呈示する。
- 「楽しみそうになったことがある」など助動詞・助詞が複雑に接続される場合を考慮し、多様な後続語表現にも対応できるようにする。
- 教科書に準拠した形での情報提示が望ましい。即ち、教科書ガイド的な機能を持たせる。

(注2)：誤った位置にアクセント核を付与するよりは、なだらかに下降するイントネーションとした方が誤りが目立ち難い。

- と同時に教科書から離れて、調べたい任意の（性質を有した）用言表現の情報を検索可能にする。
- 必要に応じて、初級者・上級者間で異なった、かつ、教育上適切な情報提示を考える。
- アクセントには揺れが存在するため、必要に応じて規範型と非規範型の両方を表示する。
- 視覚表示のみならず、聴覚呈示機能も付ける。

「コンテキスト中の単語アクセント表示機能」は、TTS システムの内部モジュールを活用することで容易に構成可能である。TTS 技術を用いた語学学習支援は様々な形態で検討されている [12]。しかし筆者らの知る限り、対話システムの応答音声、シャドーイング用呈示音声、発音訓練のモデル音声としての利用など、合成音声の利用ばかりである [13], [14]。学習者がテキストを読み上げる際に遭遇する主問題は、明示的に示されない情報（多くは韻律情報）の推定である。本研究では、出力音声ではなく、それを生成するために推定される情報を視覚的に学習者に呈示し、彼らの「読み上げ」を支援する。多くの日本語教室では public speaking をシラバスの中に組み込んでいる。しかし、東京方言で読み上げたい学習者は、教師や TA にアクセント核位置を付与してもらっている。読み上げ原稿を平仮名化できても、個々の単語の孤立発声時のアクセントが分っても、母語話者抜きでは適切に音声化できない。本研究は、その必要性が叫ばれつつも現実問題として実現困難であったアクセント教育を、技術を用いて、初めて可能とする試みである。

3. 利用した要素技術

3.1 形態素解析

任意の日本語テキストを形態素に分割し、以降の処理で必要となる、品詞、読み、孤立単語発声時のアクセント型などの情報を得る。本研究では、MeCab v0.993, UniDic v1.3.12 を元に学習された形態素解析器を用いる。（UniDic とは異なるが）IPADIC と共に学習された MeCab の精度は、[15] によれば、形態素分割、品詞推定、読みを含めた全ての属性の推定に対して、99.1%, 98.7%, 97.7% である。また、CaboCha v0.62 を用いた文節境界情報も抽出、以下で利用している。

3.2 アクセント句境界位置・アクセント核位置推定

形態素解析結果を用いた、1) アクセント句境界、2) 各句内でのアクセント核位置を検出する検出器を利用する。これらは東京方言用に、筆者等の一部が開発している [16]。形態素解析結果より様々な素性を導出し、CRF (CRF++ v0.57) による識別モデルを用いてアクセント句境界ラベリング、各形態素のアクセント型ラベリングを行なう。[16] によれば、形態素解析誤りを除外した文セットに対し、句境界推定が適合率 97.4%、再現率 90.5% であり、自動推定された句境界を用いたアクセント核位置推定が、正答率 94.7% である。

3.3 ピッチパターン描画

アクセント核位置は、タノヅム、と表示することが一般的である。しかしピッチパターンの方が理解し易い、という指摘もある。仮名表記の直上に該当するピッチパターンを乗せることになるが、タノヅム、という実際の発声のピッチパターンを乗せることは不適切である。各モーラの継続長は同一ではない。また、抽出誤りも避けられない。この場合、パターンを生成する数値モデルを用意し、そのモデル制約の下で、教師が示したいパターンの“イメージ”を描く必要がある。

本研究では、基本周波数パターン生成過程モデル [17] を用いた。基本周波数パターンをフレーズ成分（大局的な変化パターン）とアクセント成分（アクセントに伴う局所的な変化パターン）の足し合わせとして捉え、両成分を少数のパラメータで制御する。本モデルでは、アクセント成分に対応する制御パラメータが、アクセント核位置と直接対応がとれるため都合が良い。教師のイメージに沿ったパターンニングとなるよう、教師と協議しつつ各種パラメータの値を設定した。OJAD の「韻律読み上げチューター」利用時の「ピッチパターン表示用パラメータ」を ON にすると、値の詳細が表示される。

4. 単語及び後続語検索システムの開発

4.1 単語検索システム

用言におけるアクセントのコンテキスト依存性（アクセント結合）は比較的規則的であるため、ここでは、任意の用言（動詞、い形容詞、な形容詞）をクエリとし、基本活用（12 種類）に伴うアクセント変形を呈示する検索システムを構築した。

代表的な教科書を 7 種類選定し、出現する全ての用言を、その教科書で初めて出現する課の情報とともに抽出し、用言に読みを振り、基本活用時のアクセント型を求めた。対象となった用言は約 3,500 種類である。活用時のアクセント型は次のようにして定めた。まず、活用後の用言をアクセント句一つと解釈し、句中のアクセント核位置を自動推定した（3. 節参照）。当然誤りが含まれるため、全ての推定結果を検査する web システムを構築し、日本語教師 3 名（第四～六著者）に合計三回検査させた。活用数は 12 種類であるため、検査すべき項目数は約 42,000 である。検出器は一通りのアクセント型のみを呈示するが、実際には揺れが存在する。揺れが許容される用言の場合は、許容されるアクセント型についても併記させた。

男女二名の声優に約 42,000 全ての（活用後の）用言を発声させ、遮音室で収録した。12 活用/用言を発声単位としてファイル化し、音声パワーに基づいて用言頭・尾を自動検出し、前後に 200msec のポーズを付けて切り出した。切り出し発声は一度ヘッドホン聴取し、切り出し誤りは手動で修正した。

各用言に対して、二種類の難易度ラベルを付与した。一つは旧日本語能力試験に基づく難易度表を用いたものであり、他方は [18] で検討されている難易度表を使ったものである。以上の情報を用いて、用言活用に伴うアクセント情報を、用言をクエリとして検索する web システムを、MySQL v5.1.63 及び CakePHP v2.1.3 を用いて構築した。

図 2 に検索・表示条件を示す。検索対象とする用言は個別指定もできるが（「単語の検索」窓に入力）、用言の属性を用いた検索もできる。属性としては 1) 教科書とその課、2) 品詞、3) （孤立発声時の）アクセント型、3) （孤立発声時の）単語長、4) 二種類の難易度である。また、表示する項目群の表示順序についても、アクセント型（平板、頭高、中高、尾高）、アクセント核位置、単語長、難易度、五十音順のいずれを優先させて表示するのかわる変更できるようにした。更にはオプションとして、アクセントの揺れ表示、ピッチパターン表示の ON/OFF も指定できるようにした。図 3 に実際の表示の様子を示す。

音声再生は、個々の項目単位、用言単位（ある用言の全活用形、横読み）、活用形単位（ある活用形の（そのページ内の）全用言、縦読み）を再生させるなど、種々の再生モードを用意し、便宜を図った。音声ファイルはダウンロードもできる。

検索・表示条件

検索の条件 教科書 みんなの日本語 第13課 から 第16課 まで

品詞 い形容詞 アクセント型 すべて 単語長 すべて

難易度 A1(初級前半) から A2(初級後半) まで

旧能力試験 すべて から 1級 まで

表示の条件 目的別表示順序 アクセント型順に並べる

表示順序 1. アクセント型 2. アクセント型 3. 単語長 モー

オプション アクセントの揺れ 表示 ピッチカーブ 表示

左メニュー 表示 行数/ページ 20

単語の検索 カンマ、スペースで区切ると複数の単語を同時に検索できます。

実行

図2 単語検索システムの検索・表示条件

い形容詞	～い+N形	～くて形	～かった形	～くない形
長い・長いです み初10, 前日5, 前初10 初級後半 4級	ながい FEMALE MALE	ながくて FEMALE MALE	ながかった FEMALE MALE	ながくない FEMALE MALE
短い・短いです み初10, 前日5, 前初10 初級後半 4級	みじかい FEMALE MALE	みじかくて FEMALE MALE	みじかった FEMALE MALE	みじくない FEMALE MALE
全体を一括再生 FEMALE MALE				

図3 単語検索結果の例

4.2 後続語検索システム

1種類の教科書を取り上げ、「動詞+その後続語」を全て抽出した。動詞に続く後続語系列(～たかったの、など)として、約320種類が得られた。後続語検索システムは「動詞+複雑な後続語系列」(「楽しみそうになったことがある」など)をクエリとして入力し、アクセント核位置を呈示するシステムである。

クエリとして入力された「動詞+後続語系列」を形態素解析し、動詞を抽出する。次にその動詞を(起伏型、平板型)×(1グループ、2グループ、3グループ)^(注3)と6種類のいずれであるかを自動分類する。この6カテゴリの何れかが分れば、320種類の後続語系列が接続された時のアクセント型は一意に定まる。各後続語系列接続時のアクセント型は、6カテゴリの動詞を各々一つ用意し、各後続語系列を実際に接続した句を4.1節同様、自動処理し、それを人手で修正することで確定した。

結果出力例を図4に示す。単語検索システムと異なり、入力結果に対する形態素解析処理を行なうため、必ずしも精度は100%ではないと思われる。しかし入力表現から動詞を検出し、得られた言語属性から6カテゴリの何れかと判定する処理のみが自動処理であるため、実用上の問題は起きていない。

320種類の後続語系列以外の語句がクエリとして入力される場合がある。この場合、クエリを平仮名表記したものと、320種類の平仮名表記された動詞+後続語系列を先頭から比較し、最長一致となるものを代案として表示している。完全一致の場合は結果を赤枠で、代案呈示の場合はピンク枠で呈示し、情報の不確実性についてもユーザーに明示している。図4にその様子を示す。なお、検出結果のみならず、関連する表現のアクセントについても示している(図参照)。

(注3): 日本語教育では、五段動詞を1グループ動詞、上一段・下一段動詞を2グループ動詞、不規則動詞(～する、と、来る)を3グループ動詞と呼んでいる。

つくり <u>ぞう</u> になったことがある	あそび <u>そう</u> に <u>な</u> ったことがある	たのしみ <u>ぞう</u> になったことがある
つくり <u>ぞう</u> になったことがある	あそび <u>そう</u> に <u>な</u> ったことがある	たのしみ <u>ぞう</u> になったことがある

入力動詞「楽しむ」(起伏式1グループ動詞)に対して、「作る」(起伏式1グループ動詞)、「遊ぶ」(平板式1グループ動詞)も常に示している。

図4 後続語検索の例(上:完全一致,下:不完全一致)

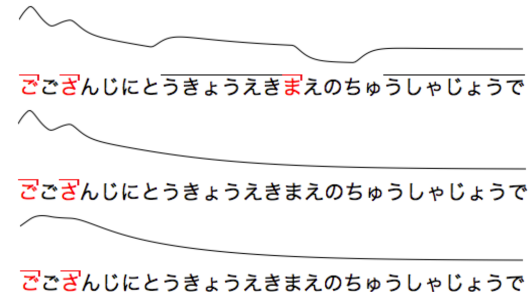


図5 3種類のピッチパターンと核位置表示

5. 韻律読み上げチューターの開発

フレーズング&ポージングに基づく発声訓練法[10]に準拠し、任意の文に対してアクセント核位置やピッチパターンを呈示する韻律読み上げチューターを設計した。[10]ではフレーズを「意味の区切り、呼吸の区切りなどによって形成される一息で発声される単語系列」と定義している。読み上げチューターは、フレーズ区切り(“/”)が与えられた日本語テキストに対して、各フレーズにアクセント核位置を必要な箇所呈示する、との方針をとった。なお、文中に含まれる句読点(とそれに準ずる記号)及び改行は、自動的にフレーズ区切りと解釈している。

フレーズを単位として形態素解析を行ない、アクセント句境界検出を行なうと、通常、複数のアクセント句が出力される。つまり、フレーズの中には複数のアクセント核が観測されることが多い。しかし全てのアクセント核を常時呈示するのは学習者の負担も大きいと、上級者モードでは全てのアクセント核を、初級者モードでは第一アクセント核及び、(頭高型アクセントに対する知覚的敏感性[19]を考慮し)3モーラ以上の頭高型アクセント句の核のみを示すこととした。

更に[10]では山フレーズと丘フレーズという概念を導入している。前者はアクセント核を有するフレーズのピッチパターンであり、後者は有さないフレーズのピッチパターンを意図している。前者は、アクセント核によるピッチの急速な下落を実現するために(後者に比べ)事前により大きなピッチの立ち上がり形成することを意図しており、これを山と表現している。後者はそれが無いため、丘となる。これはアクセント核の位置は正しく把握できていないが、アクセント核があることだけは分っている学習者が発声する場合に、高低差のより大きい「へ」の字を描くように発声指導することが効果的であるという、教育経験から生まれた実用的な便法である(図1参照)。

以上の検討に基づき、フレーズを単位としたアクセント核表示について、3種類のモードを用意した。図6には「午後三時に東京駅前の駐車場で」を一フレーズとして入力した場合の処理を示している。実際の出力結果を図5に示す。なお、フレーズが長すぎて一息で発声困難な場合は、フレーズ境界記号“/”を挿入して、2フレーズとして解析すればよい。山・丘ピッチパターン表示の際のスムージング処理は、基本周波数パターン生成過程モデルの制御パラメータの値を変更して実装している。

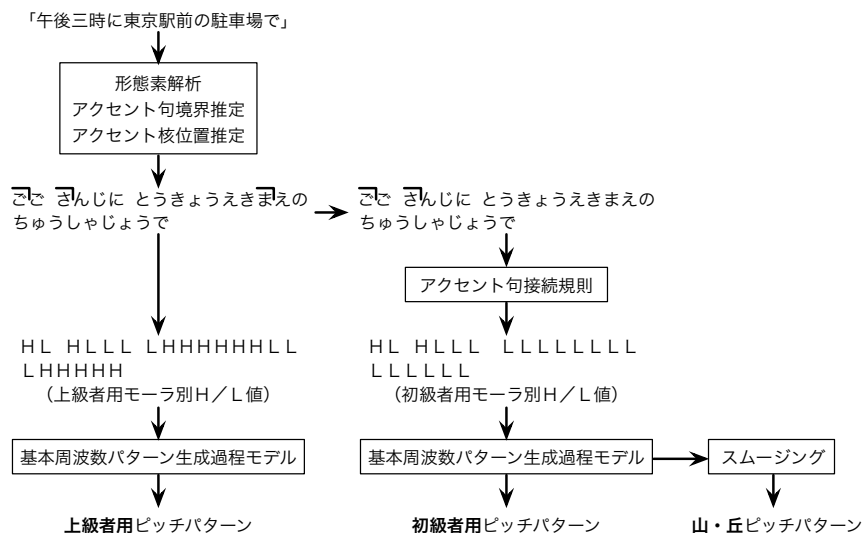


図6 フレーズに対する各モーラのアクセント属性の3種類の推定方法

表1 単語検索システムに対する教師の評価 (%)

a) 学習者に役立つと思うか?		b) 授業で使うか?	
非常に役立つ	71.0	是非使いたい	38.7
少し役立つ	29.0	必要があれば	59.7
あまり役立たない	0.0	必要ない	1.6
全く役立たない	0.0		

表2 後続語検索システムに対する教師の評価 (%)

a) 学習者に役立つと思うか?		b) 授業で使うか?	
非常に役立つ	54.8	是非使いたい	29.0
少し役立つ	45.2	必要があれば	64.5
あまり役立たない	0.0	必要ない	6.5
全く役立たない	0.0		

表3 韻律読み上げチューターに対する教師の評価 (%)

a) 学習者に役立つと思うか?		b) 授業で使うか?	
非常に役立つ	62.7	是非使いたい	42.6
少し役立つ	28.8	必要があれば	50.0
あまり役立たない	8.5	必要ない	7.4
全く役立たない	0.0		

6. 開発したシステムの評価実験

6.1 2種類の検索システムの評価

各国の80名の日本語教師に「使ってみよう OJAD」メニューを実行させ、アンケート調査を行なった。約2/3は海外で日本語を教える教師である。2種類の検索システムに関するアンケート調査結果を表1、表2に示す。韻律教育は日本語教育全体の中の一部門であることを考えると、検索システムの教育的実用性は十分に認めてもらえたものと考えている。

6.2 韻律読み上げチューターの評価

6.2.1 教師を対象としたアンケート調査

教師によるアンケート調査の結果を表3に示す。こちらも同様、高い実用性が認められる。3種類のシステムの中では、一番「授業で是非使いたい」との反応が得られる一方、「学習者にあまり役立たない」という否定的な回答も8.5%ほど見られた。

6.2.2 学習者を対象とした客観的評価実験計画

Public speakingを控えた学習者が、母語話者の助けが得られない状況で、読み上げ原稿にアクセント位置を振る状況を考える。この状況下で、アクセント辞典、音声合成器、読み上げチューターを使わせて作業させ、結果を比較する。具体的には、a) PC上のNHKアクセント辞典のみ、b) アクセント辞典と合成器の併用、c) アクセント辞典とチューターの併用、の三者を

比較した。合成器としてはHOYAサービスのSAYAKAを使用した[9]。a)の場合はコンテキスト中の単語アクセントが表示困難であること、b), c)の場合は合成器、チューター共に誤ることがあることを被験者に伝え、自分の持つ日本語の知識に照らし合わせて、各システムを参照して作業するよう伝えた。

被験者としては1)単語アクセントがピッチアクセントであること、2)コンテキストによって容易に変化すること、を知っている学習者を集めた。被験者数は35であり、約8割が日本語能力試験1級を取得しており、上級者と呼ばれる学習者である。

読解教材から旧日本語能力試験2級程度に相当すると思われる読解文を四つ選んだ(文章0~3)。アクセント付与のみに着眼するため、文章0~3に、フレーズ区切りを事前に挿入した。各文章のフレーズ数は73, 68, 73, 70であった。被験者のタスクは「各フレーズの先頭のアクセント核を指摘する」である。実験はPC上で行なわれ、アクセント付与web、アクセント辞典、合成器、チューター共に同一のPC上で行なえるよう、環境を構築した。各々のシステムは事前に使い方を十分に習得させた。PC画面の例を図7に示す。各システム使用時の迅速性を見るため、web上のクリックは全てログとして記録した。学習者にはアクセント位置に関する判断をクリックに即座に反映するよう、依頼した。開始後30分を目安に作業は終了させた。

文章0~3は次のように使い分けた。自らの日本語知識のみでどの程度正解できるかを見るため、被験者全員に対し、まず文章0を用いた実験を行なった。次に文章1~3を用いて以下を行なった。今回の実験では、合計36通りの文章及びシステムの組み合わせが存在するので、これを被験者に順次割り当てた。

被験者によるアクセント付与が行なわれた文章に対する回答の正誤判定は、日本語教師(第五、六著者)が行なった。アクセントの揺れも考慮して判定した。



アクセント核の指摘は、該当する平仮名をクリックする。

図 7 アクセント核位置検出実験

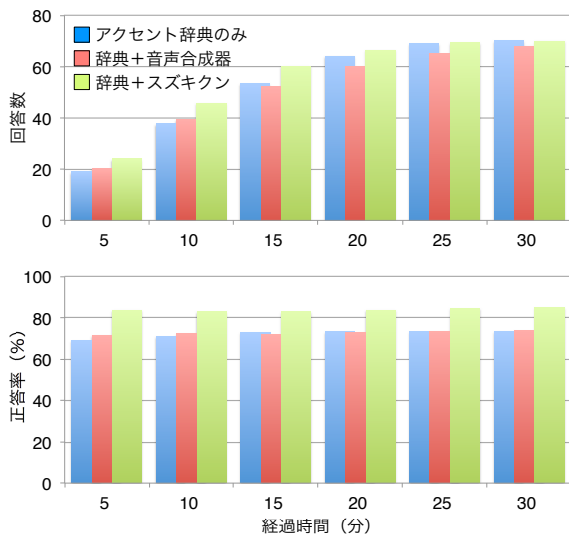


図 8 客観的評価実験結果

6.2.3 客観的評価実験結果とその考察

文章 0 に対する精度は平均 68.2%であった。なお、同じタスクを母語話者の工学部学生 10 名（いずれも関東出身者）に対して行なったところ、平均 61.6%であった。この結果は、2.1 節で述べたように何も驚くべきことではない。読み上げチューターは 93.2%の正解率を示した。身近にいる東京方言話者に依頼するよりは、遥かに高い精度で付与することができる。

文章 1～3 を用いた実験に関する結果を図 8 に示す。何れも横軸は実験開始からの経過時間である。上図の縦軸は、各システム使用時の開始後 x 分までの回答数の平均（回答の正誤は無視している）であり、回答の迅速性を意味する。下図は、正答率（正答数／回答数）の平均であり、回答の正確さを意味する。

まず回答数を見る。事前に音声合成器及びチューターの使い方を習得させる時に、“/”で区切られた文章全体を合成器、チューターにコピー＆ペーストして実行すれば数秒後には、合成音声やアクセント核位置が呈示されることは説明している。しかし図 8 上図は、チューターが示した核位置を各々吟味して回答している様子が窺われる。これは、「誤る可能性がある」という事前知識によるものと考えられる。分散分析の結果、統計的有意差（危険率 1%）は、何れの場合も観測されなかった。

次に正答率を見る。分散分析の結果、実験開始後の経過時間によらず、c) において、a) および b) に対する統計的な有意差

（危険率 1%）が見られた。これらの結果より、チューターは他のシステム利用時よりも、（回答の速度ではなく）回答の質を上げることには貢献していると言える。その一方で合成器の利用（聴覚呈示）は、a) と比べて回答の質を向上できていない。

なお、各システム利用時の最終的な精度平均 a)73.1%, b)73.9%, c)84.8%となった。文章 1～3 に対するチューター単体の精度は 91.0%であり（合成器単体は 91.3%）、実は、c) よりも高い。これはチューターの出力に対して、学習者がアクセント辞典や自らの知識を参照して吟味した結果は、改悪する方向に働いていることを意味する。即ちチューターの結果に対して「正しい箇所は即座にそれを利用し、不確かな箇所のみ吟味する」ことが難しいと解釈される。今後、技術的な精度向上を目指すと共に、後続語検索システムのように、呈示する情報の「確からしさ」を視覚化することで、この問題は（部分的には）解決できると思われる。回答の速度向上も期待できる。

6.2.4 学習者を対象としたアンケート調査

実験後、アンケート調査を行なった。「三種類のシステムはアクセント付与作業にどの程度貢献したのか」に関する主観評価であり、非常に用役立つと答えた率は a)37.5%, b)30.0%, c)82.5%であり、圧倒的にチューターが有利であった。合成音声は公のサービスで広く使われている高品質の合成音声を用いているが、アクセントの把握を支援するまでには至っていない。

7. ま と め

日本語教師と非常に密な協力を図り、日本語韻律教育を支援する目的で、オンラインアクセント辞書及び韻律読み上げチューターを開発した。教師及び学習者を対象にした評価実験の結果、優れた教育的効果を示すことができた。2012 年 8 月中旬より運用を開始しているが、過半数は海外からのアクセスである。また、未採択教科書の出版社から、本システムでの採択を希望する連絡も来ており、日本語教育業界に対する一定のインパクトも示すことができたと考えている。アンケート調査の結果幾つかの要望も来ており、今後も改良を加えていきたい。

文 献

- [1] OJAD, <http://www.gavo.t.u-tokyo.ac.jp/ojad/>
- [2] D. Crystal, *English as a global language*, Cambridge University Press, New York, 1995.
- [3] 轟木他, 香川大教育実践総合研究, 18, 45–51, 2009.
- [4] 阿他, 信学技法, SP2009-151, 19–24, 2010.
- [5] 船本, “留学生の方言意識 ～熊本方言テキスト作成のためのアンケート調査から～”, 科研費 (基盤研究 (B))18320082, 成果物
- [6] S. Kato, et al., *Proc. Speech Prosody*, 198–201, 2012.
- [7] NHK 日本語発音アクセント辞典新版, NHK 出版, 1998.
- [8] 松崎, 広島大日本語教育研究, 18, 35–41, 2008.
- [9] 例えば, <http://voicetext.jp> など。
- [10] 中川他, さらに進んだスピーチ・プレゼンのための日本語発音練習帳, ひつじ書房, 2009.
- [11] 平野他, 日本音響学会誌, 65, 2, 69–80, 2009.
- [12] M. Eskenazi, *Speech Communication*, 51, 832–844, 2009.
- [13] A. Black, *Proc. SLaTE*, CD-ROM, 2007.
- [14] Z. Handley, et al., *Language Learning & Technology*, 9, 3, 99–120, 2005.
- [15] MeCab の開発経緯, <http://mecab.googlecode.com/svn/trunk/mecab/doc/feature.html>
- [16] 鈴木他, 電子情報通信学会論文誌, vol.J96-D, no.3, 2013.
- [17] 藤崎他, 日本音響学会論文誌, 27, 9, 445–453, 1971.
- [18] 基盤研究 (A) 「汎用的日本語学習辞書開発データベース構築とその基盤形成のための研究」, <http://jisho.jpn.org>
- [19] N. Minematsu et al., *J. ASJ(E)*, 16, 5, 311–320, 1995.