

ポーランド日本語教師会

ワルシャワ勉強会

“OJAD”

[Online Japanese Accent Dictionary](#) について

2013年11月30日

飯田洋一

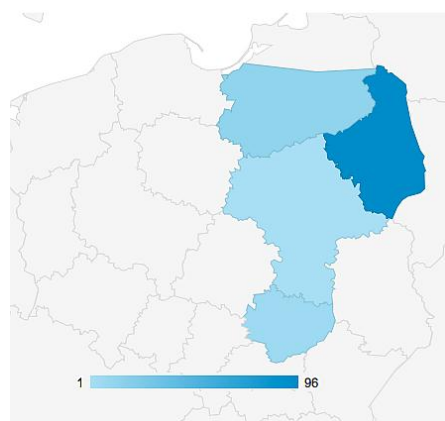
● 発表にあたって

今回、ワルシャワの勉強会で OJAD をテーマに選んだのは、去る 2013 年 10 月 5 日、ポズナンで開かれたポーランド日本語教師会の勉強会のテーマであった [“OJAD : Online Japanese Accent Dictionary”](#) を、参加できなかった会員の方々に紹介したいという思いからであった

そもそも、このポズナンでの勉強会が開かれるに至った経緯とは、ポーランド日本語教師会のウェブサイトをご覧になった東京大学大学院教授・OJAD 開発チームリーダーである峯松信明先生が、直接教師会宛てにメールを送っていただいたことがきっかけで実現した。



峯松信明先生



峯松先生曰く、「OJAD のアクセス解析によると、非アジア圏の地域からでは、ポーランドが一番のアクセス率を誇る(2013 年 5 月分の解析)。ということは、ポーランドには日本語のアクセントやイントネーションを非常に熱心に学んでいる教師や学習者がいる。そんな熱心なユーザーにぜひ会ってみたい。」とのことであった。

*←確かに、アクセス数は高い。しかし、一番高いのが、Białystok のある Podlaskie とは・・・。

それだけでなく、音声科学・音声工学の日本・ポーランド間のプロジェクトの発表会が、ちょうどその時期にポズナンであり、それに参加されていたこともある。そういった経緯で、教師会の岡崎会長にも了承をいただき、ポズナンの先生方が進んでホスト役を務めていただいたおかげで、開催に至ったという経緯である。

● 発表について

OJAD について今回ワルシャワの勉強会で発表した点は、おおまかに以下の 3 点である。そのほとんどの部分は、ポズナンの勉強会で峯松先生から教わったことである。

1. OJAD って？
2. OJAD の 4 機能
3. OJAD の今ある課題と将来

1. OJAD って？

まず、OJAD の名称であるが、“[Online Japanese Accent Dictionary](#)” の頭文字をとったものであり、「オウジャッド」と呼ばれている。「日本語教育のためのコーパスを利用したオンライン日本語アクセント辞書の開発」というプロジェクトで、国立国語研究所からの援助を受け、現在、東京大学大学院工学系研究科の峯松研究室と、同じく情報理工学系研究科、廣瀬研究室の合同チームによって開発が進められている。



OJAD開発チームのみなさん

○OJAD の目的

OJAD の研究プロジェクトには、二つの目的がある。

一つは、日本語教育のためその音声教育のインフラを作るということと、もう一つは、日本語教育を対象とした音声工学の技術を推進させるということである。

日本語教育においては、教師が学習者に日本語を流暢に運用させることを目標にして学習が進められるが、アクセントを教えるということは、時間にも制約があるし、教材も少ない。それに、音声・韻律の知識がいるので、非母語話者の教師なら、なかなかきちんと教えられないというのが現実である。そういった日本語教育の現場をサポートするため、短期間でも効果的な教育用ツールとして生まれたのが OJAD である。

また、音声工学においては、技術者・研究者が機械に日本語を教え(プログラムを構築し)、流暢に日本語を話させるようにすることにある。そして OJAD に関連した技術には、音声認識(音声をテキスト化すること)と音声合成(テキストを音声化すること)がある。今では Apple の Siri や、国会答弁のテキスト化、人間型ロボットに搭載などすでに実用化されており、これからも他の様々な分野にも応用できることが期待されている。

2. OJAD の 4 機能

OJAD の主な機能は、以下の 4 つである。

- (1) 単語検索
- (2) 動詞の後続語検索
- (3) 任意テキスト版
- (4) 韻律読み上げチュータスズキクン

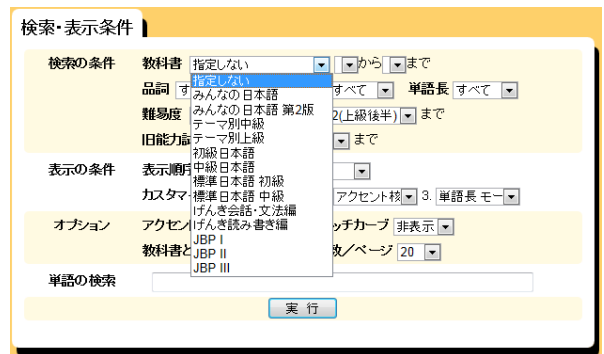
OJAD 4機能
 単語検索
 動詞の後続語検索
 任意テキスト版
 韻律読み上げチュータスズキクン

(1) 単語検索

この単語検索内には、日本語教育でいちばんよく使われている 6 種類の教科書の全単語と全品詞のアクセントが検索できるようになっている。

-掲載教科書-

- [みんなの日本語初級 I、II
- [みんなの日本語 初級 第 2 版 I II
- [テーマ別 中級から学ぶ日本語<改訂版>
- [テーマ別 上級で学ぶ日本語<改訂版>
- [初級日本語 新装版
- [中級日本語
- [新版中日交流標準日本語 初級
- [新版中日交流標準日本語 中級
- [初級日本語[げんき] 会話・文法編
- [初級日本語[げんき] 読み書き編
- [JAPANESE FOR BUSY PEOPLE I
- [JAPANESE FOR BUSY PEOPLE II
- [JAPANESE FOR BUSY PEOPLE III



みんなの日本語第 2 版 I 第五課の一類動詞の検索結果例

グループの動詞	辞書形	～ます形	～て形	～た形	～たぬい形	～ば形	使役形	受身形	命令形	～う形
① 買う・買います	かう	かいます	かって	かった	かわない	かえば	かわせる	かわれる	かえ	かう
② 聞く・聞きます	きく	ききます	きいて	きいた	きかない	きけば	きかせる	きかれる	きけ	きこう
③ 遊ぶ・遊びます	あそぶ	あそびます	あそんで	あそびた	あそびない	あそべば	あそべる	あそばれる	あそべ	あそぼう
④ 書く・書きます	かく	かきます	かいて	かいた	かかない	かかば	かかせる	かかれる	かき	かこう
⑤ 振る・振ります	ふる	ふるまいます	ふるって	ふるった	ふるまない	ふるまえば	ふるませる	ふるまれる	ふるめ	ふるもう
⑥ 飲む・飲みます	のむ	のみます	のんで	のんだ	のまない	のめば	のませる	のまれる	のめ	のもう
⑦ 読む・読みます	よむ	よみます	よんで	よんだ	よまない	よめば	よませる	よまれる	よめ	よもう

◎全ての活用のアクセントを表示

◎右クリック↓「対象をファイルに保存する」または、「Shiftを押しながらクリック」でダウンロードもできる

◎ここをクリックすることで、(男性・女性の)音声再生も可能

(2) 動詞の後続語検索

動詞(用言全般)は活用すると、そのアクセントは後続に続く文型によって多様に变化していく。(1)の単語検索機能では、動詞活用形のアクセントまで検索可能だが、この機能では、動詞の後ろにつく文型によって变化するアクセントを任意に検索できる。

後続語検索

使用例

動詞表形 明ければ

ピッチカーブ 表示 ▾

表示する列 検索した結果のみ表示 ▾

めいりやし

類型	検索した動詞「明ける」平板式 2グループ動詞	
〇〇ば		あけ れば
		あけ れば いい
類型	検索した動詞「明ける」平板式 2グループ動詞	
命令		あけ る (命令)

あけまじょう

		あけま じ ょうか
		あけま せ ん
		あけま せ んでした

あけさせてください

類型	検索した動詞「明ける」平板式 2グループ動詞	
〇〇う		あけ よう
		あけ よう と
		あけ よう ど
		あけ よう とは
類型	検索した動詞「明ける」平板式 2グループ動詞	
使役受身		あけ させ られる

アクセントを調べたい動詞とその後続語を入力し検索

検索した後続語があるところまでジャンプし表示

◎「明ける」という動詞では、327の文型(後続語)に応じたアクセントを表示

* その他の動詞も同じぐらいの文型(後続語)をカバーしている。

(3)任意テキスト版

この機能では、(1)(2)の機能においた単語単位でのアクセント検索ではなく、任意の日本語文章を入力すると、その中から用言を抽出し、用言(活用することば)の活用形のアクセントを表示する。文の中で出てくる用言の確認と、その活用形のアクセントを調べることができる。その文は、自作の文章やweb ページからのコピーなどでも可能である。

使用例

任意テキスト版

入力エリアの表示

■1グループの動詞 ■2グループの動詞 ■3グループの動詞 ■い形容詞 ■な形容詞
※色のついた語にマウスオーバーするとその語の情報が見れます

自然言語処理技術を用いた自動処理を行なっているため精度は100%ではありませんが、
 日本語の学習にぜひお役立てください。

見出し	品詞	辞書形	～ます形	～て形	～た形	～ない形	～なかった形	～ば形	使役形	受身形	可能形	～う形
ある	動1	ある	あります	あって	あった	ない	なかった	あれば				
くださる	動1	くださる	くださります	くださって	くださった	くださらない	くださらなかった	くだされば	くださらせる	くださられる		くださろう
行なう	動1	おこなう	おこないます	おこなって	おこなった	おこなわない	おこなわなかった	おこなえば	おこなわせる	おこなわれる	おこなえる	おこなおう
用いる	動2	もろいる	もろいます	もろいて	もろいた	もろいない	もろいなかった	もろいれば	もろいさせる	もろいられる	もろいられる	もろいよう
役立てる	動2	やくだてる	やくだてます	やくだてて	やくだてた	やくだてない	やくだてなかった	やくだてれば	やくだてさせる	やくだてられる	やくだてられる	やくだてよう
いる	動2	いる	います	いて	いた	いない	いなかった	いれば	いませる	いられる	いられる	いよう

文中の用言を表示

見出し	品詞	辞書形	～ます形	～て形	～た形	～ない形	～なかった形	～ば形	使役形	受身形	可能形	～う形
用いる	動2	もろいる	もろいます	もろいて	もろいた	もろいない	もろいなかった	もろいれば	もろいさせる	もろいられる	もろいられる	もろいよう

◎文中の色のついたことばにマウスオーバーすると、個別にアクセントがポップアップで表示

*ただし、自然言語処理技術を用いた自動処理を行なっているため、精度は100%ではないとのこと。

(4)韻律読み上げチュータスズキクン

この機能では、学習者に読ませたい任意の日本語文章を入力することで、アクセントの核とピッチパターンの2つを表示させることができる。その2つに注意させ、目で追わせながら音読練習をさせると、文章単位で自然な東京アクセントの体得につながる。

次のページに、その使用例を紹介する。

韻律読み上げチュートスズキケン

使用例

文を入力し検索

◎アクセントの核が赤で表示され、文章の上に、ピッチパターンがカーブで表示される

*ただし、ピッチパターンは特定の語に強いフォーカスを置かずに読み上げることを想定している。

*疑問文のイントネーションを表示させる場合は、文末に「？」を入れる。

*驚きを示すイントネーションなどについては対応していない。

*形態素解析、アクセント句境界推定、アクセント核推定などの技術を用いているため、精度は100%ではない。

この機能を活用して、ピッチパターンとポージング、フレージングを明記した原稿で学習者が練習すれば、弁論大会やプレゼンの練習に抜群な効果を発揮するだろう。

まず、実際にOJADのサイトから「OJAD 4機能のご紹介」と「使ってみようOJAD」を自分で使ってみて、どんなものか試してみるといい。

OJAD

Online Japanese Accent Dictionary

INTERSPEECHデモ展示用の 1-min PV を作成しました。日本語版/英語版

OJADとは？

OJADは日本語教師・学習者のためのオンライン日本語アクセント辞書です。約3,000の東京方言アクセントを調べることができます。約3,500の用言(動詞、い形容詞、名詞)について、基本12活用、約42,300のアクセントを調べることができます(男女別、年齢別、職業別)。それ以外の約300の継続語表現に関するアクセントも調べることができます。更に任意の単語に対して、特定の

[「使ってみよう OJAD」](#)

* 例題を使いながら OJAD の機能を学べるページ

[「OJAD 4機能のご紹介」](#)

* OJAD の取扱い説明書

3. OJAD の今ある課題と将来

以上見たように、OJAD の機能を活用すれば、短期間で効果的な学習ができ、非母語話者の教師でも音声指導が可能になる。

しかし、100%完璧な技術を生み出すのは不可能であり、そのため、4つ目の機能「韻律読み上げチュータスズキクン」においては、まちがったアクセントの表示をしてしまうときもある。

以下が、現在わかっているスズキクンの弱点とその対処法である。

○スズキクンの弱点と対処法

-数字(記号)が読めない。

(例) ”2006年”と書くと、”にゼロゼロろくねん”と読んでしまう。

(対処法)漢数字に書き換える。(例)2006年→二千六年

-漢字の読み方が複数あることばでは、決まった読み方で表示される。


(例) ”日本”→”にほん” ”私”→”わたし”

(対処法)”にっぽん”や”わたくし”と表示させたい場合は、ひらがなで書く。


-たまに間違ったアクセントを表示してしまう。

(例) ”ちりょうが”と書くと、 ちりょうが となってしまう。

*形態素解析(単語分割)処理は、ひらがな入力に弱いため

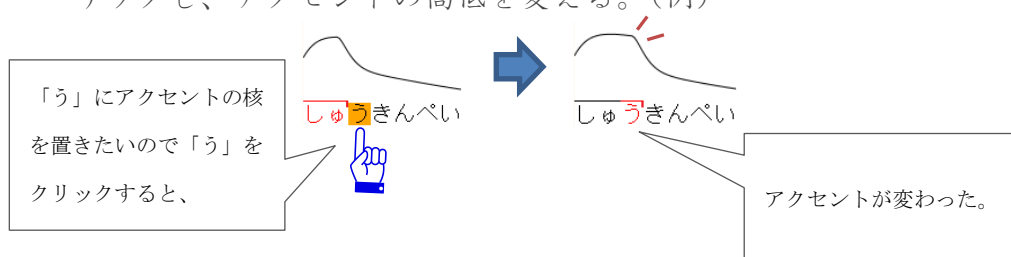
(対処法)漢字に書き換えてみる。(例)ちりょうが→治療が 

-カタカナの人名や地名など、知らないことばがあると、適当なアクセントを表示したり、無視・黙ってしまう。

(例) ”習近平”のアクセントを  しゅうきんぺい と表示してしまう。

“朴槿恵”は、無視して表示しない。

(対処法)アクセントが表示された場合、アクセントを変えたい文字の上をクリックし、アクセントの高低を変える。(例)



● OJAD の将来

今後 OJAD に期待できるのは、外国人学習者が初級でも年齢が若くても、単語検索でアクセントを自分で調べられるようになるため、多言語化が進められている。2014 年 1 月の時点で、日本語以外では、英語、インドネシア語、韓国語、中国語版のページが開設されている。ポーランド語版が英語の次に来るかと思われたが、残念ながら、他言語に先を越されてしまったようである。ポーランド語担当者にも一刻も早い対応を期待する。

それ以外では、チュータズズキクンをグレードアップさせることであろうか。前述したようなエラーが減ることが考えられる。

現在では、テキストを入力する時点で、「,、.。.. ? ! ? ! /」の記号、及び改行をいれることによって、ズズキクンはポーズするところを認識するが、近い将来、これが自動的に入るようになる。

さらに、音声合成(Text to speech synthesizer)の技術と組み合わせて、ズズキクンを喋らせるようにし、聴覚的にアクセントを呈示させることも進められている。

● まとめ

日本語はピッチアクセントで音節もモーラ単位なので、その響きには、ストレスアクセントやトーンアクセントの言語のそれとはまた異なった美しい趣がある。私の学校にも、アニメや歌などで映像や音楽に絡んでくるその日本語の音が好きで、日本語を勉強してみるきっかけになったという学習者も少なくない。音声面だけでなく、文化や代々積み上げられてきた日本人の精神が絡んだことばを、物心つかないころ親から教わり、それが今は私たち日本人の心と身体に染みついている。それを次の世代、しかも外国人学習者に伝播できる立場にいられるという状況に感謝せざるを得ない。

私がポーランドにいる目的とは、将来、日本語を運用することで日本とポーランドの両社会を支えるような人材を育てる土地の、ごくわずかな一部分を耕すことである。正しいアクセントで話すことは、コミュニケーションを円滑にするための重大な要素である。そして、こういったプログラムを開発している技術者の日々努力してくれるおかげで、日本語教育だけでなく、日本の技術分野の発展にもつながる。今後、自身の授業にも、OJAD を積極的に取り入れていくつもりである。

追記：はるばるポーランドまでお越しいただき、貴重なツールをご紹介くださった峯松先生、ポズナン勉強会でホスト役を担ってくださった UAM の先生方に、今一度心からお礼申し上げます。

(参考資料)ポズナン勉強会での峯松先生のハンドアウト：OJAD_Poznan_4up.pdf

OJAD ホームページ <http://www.gavo.t.u-tokyo.ac.jp/ojad/>

OJAD フェイスブック <https://www.facebook.com/#!/pages/OJAD/346579645441929>

峯松信明先生 & 峯松研究室 <http://www.gavo.t.u-tokyo.ac.jp/~mine/japanese/>