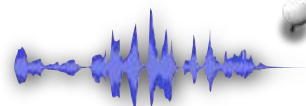




OJAD講習会@九州大学



九州大学大学院・地球社会統合科学府

言語・メディア・コミュニケーションコース 言語教育セミナー

OJADとそれを用いた音声指導

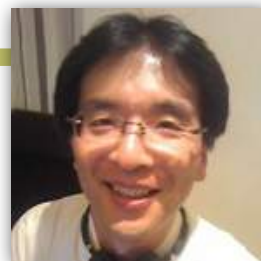


峯松 信明

東京大学大学院工学系研究科



講演者 自己紹介



- 氏名: 峯松 信明 (みねまつ のぶあき)
- 所属: 東京大学大学院工学系研究科
- : 電気系工学専攻・教授
- 専門: 音声科学・工学 (外国語学習支援を含む)
- 経歴: 高校時代に一時期英語教師を目指すも理科系に。
- : 大学1, 2年は, 英語劇の舞台の上で過ごした。
- : 89年学部4年夏から音声研究を開始する。
- : 95年博士(工学) 取得@東大工学系
- : 95年~ 豊橋技術科学大学情報工学系
- : 00年~ 東京大学大学院工学系研究科
- : 12年, INTERSPEECHにてCALL tutorial担当

大学共同利用機関法人 人間文化研究機構 ...したオンライン日本語アクセント辞書の開発

大学共同利用機関法人 人間文化研究機構
国立国語研究所
National Institute for Japanese Language and Linguistics

国語案内 研究組織 研究活動 データベース・データ集 催し物等 刊行物

トップ > 研究活動 > 共同研究プロジェクト > 領域指定型 > 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

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

- 略称 : アクセント辞書
- プロジェクトリーダー : 峯松 信明
(東京大学大学院工学系研究科教授)
- 研究分野・キーワード : <研究分野: 日本語教育学, 音声学, 音声工学>
(日本語教育) (外国語読り) (韻律的特徴) (アクセント結合)
(オンライン辞書)
- プロジェクトの成果 : 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
OJAD (Online Japanese Accent Dictionary)
<http://www.gavo.t.u-tokyo.ac.jp/ojad/>

概要

言語資源研究系
領域指定型共同研究プロジェクト

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

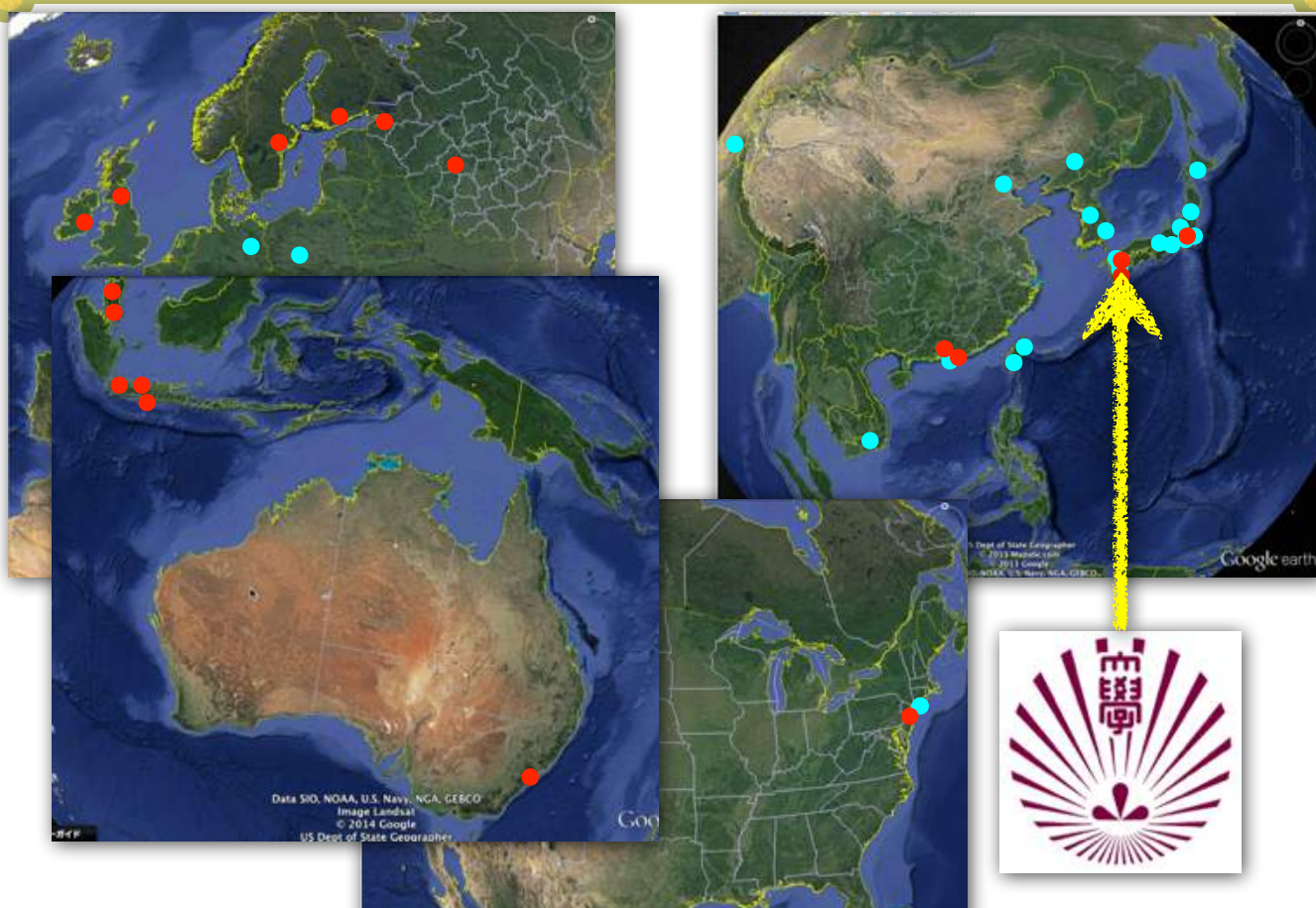
リーダー: 東京大学准教授 峯松 信明

国内外における日本語学習者の増加
→ 自然な音声コミュニケーション能力の獲得を望む学習者

ふりがな 氏 名	所属・職名	専門分野 (学 位)
みねまつのぶあき 峯松 信明	東京大学大学院情報理工学系研究 科電子情報学専攻・准教授	音声学 博士(工学)
ひろせ けいきち 広瀬 啓吉	東京大学大学院情報理工学系研究 科電子情報学専攻・教授	音声学 工学博士
ひらの ひろこ 平野 宏子	吉林華僑外国語学院 (中国) 日本語学部・専任講師	日本語教育・ 音声学 博士(科学)
なかがわ ちえこ 中川 千恵子	早稲田大学日本語教育研究 センター・非常勤講師	日本語教育 人文科学博士
なかむら のりこ 中村 則子	慶應義塾大学国際センター 非常勤講師	日本語教育 文学修士
すずき まさゆき 鈴木 雅之	東京大学大学院工学系研究科 電気系工学専攻・博士課程	音声学 修士(工学)
たがわ ゆきのり 田川 恭識	早稲田大学日本語教育研究センタ ーインストラクタ (非常勤)	日本語教育 文学修士



World Tour of the OJAD tutorial



本日のメニュー

🔊 最近の PC は上手に喋るけど、誰が日本語教えてる？

- 韻律に着目。単語アクセント・アクセント結合・アクセント句

🔊 日本人にとってのアクセントとは？

- 日本人はアクセントの間違いに敏感なの？鈍感なの？

🔊 日本人にとって分かり易い発声とはどういうものか？

- フレージング（チャンキング）とポージング

🔊 OJAD 作りました。ご紹介します。

- 単語検索 / 動詞の後続語検索 / 任意テキスト版

- 韻律音読チュータ・スズキくん



🔊 OJAD 使って教えています。指導のヒントをご紹介します。

- フレージング&ポージングによる発声指導&プレゼン指導

🔊 まとめ

日本語教育と音声工学

🔊 日本語教育の目的の一つ

- 学習者が日本語文を流暢に読み上げられるようにする。

- そのために様々な教育理論・教育戦略がある。

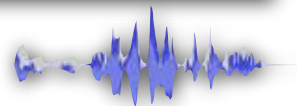
🔊 （日本語を対象とした）音声工学の目的の一つ

- 機械が日本語文を流暢に読み上げられるようにする。

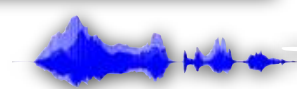
- そのために様々な音声理論・音声技術がある。



初めてバンクーバーに来ました。



初めてバンクーバーに来ました。



パソコンに日本語を読ませる技術

私も日本語教師です。でも生徒はパソコンです。

● 日本語テキスト音声合成 (text-to-speech synthesizer) の例

● <http://www.ai-j.jp>

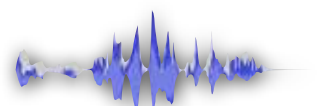
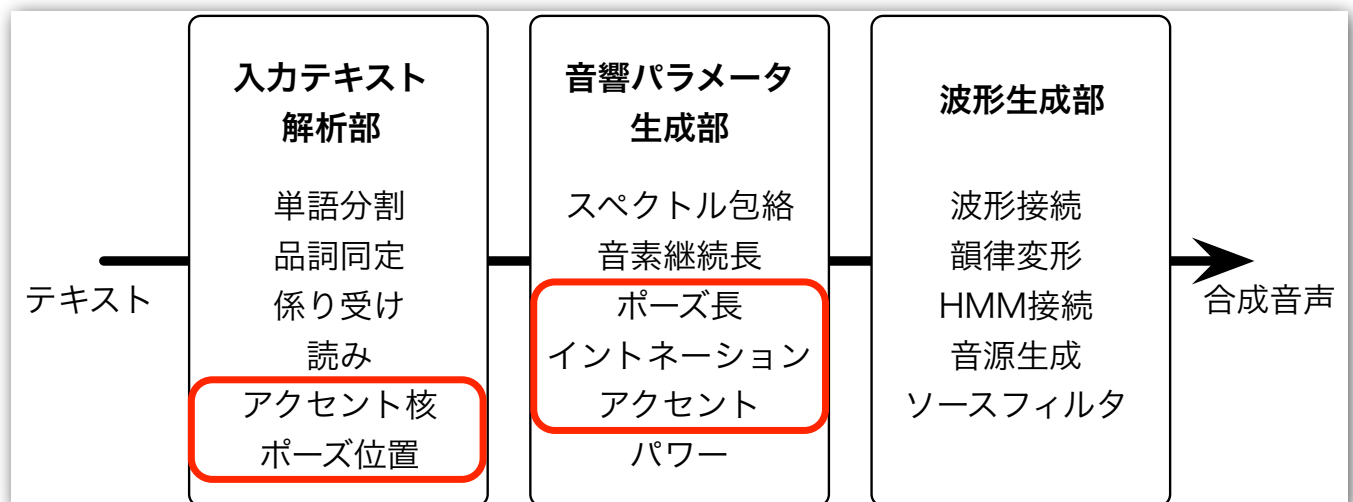


● パソコンに共通語（東京方言）で読ませるにはどうすればよい？

テキスト音声合成の全体像

まずは、全体像から始めましょう。

生まれて初めて、
バンクーバーにきました。



テキスト解析結果に基づく区切り

単語・文節・アクセント句・フレーズの区切り

生まれて初めて、バンクーバーに来ました。

↓ **単語（形態素）**に区切る

生まれ（動） | て（助） ■ 初めて（副） ■ バンクーバー（名）
| に（助） ■ 来（動） | まし（助動） | た（助）

↓ **単語境界の一部が文節境界に**

生まれて | 初めて ■ バンクーバーに ■ 来ました

↓ **文節境界の一部がアクセント句境界に**（アクセント的まとまりを作る）

生まれて初めて ■ バンクーバーに | 来ました

↓ **アクセント句境界の一部が句境界に**（イントネーション的まとまりを作る）

生まれて初めて || バンクーバーに来ました （|| ポーズを置く）

入力テキスト
解析部

単語分割
品詞同定
係り受け
読み
アクセント核
ポーズ位置

テキスト

一端ばらばらにしてから、まとめ上げる。
韻律は、まとめ上げを助ける糊のようなもの。

読みが決まれば次はアクセント

各モーラのH/Lを適切に自動推定する。

● 出力すべきは共通語（東京方言）の音声

● **モーラ**：日本語の発声の単位。拍。凡そ平仮名一つ分。

● かなだ=カ | ナ | ダ, バンクーバー=バ | ン | ク | ー | バ | ー

● Vancouver = Van | cou | ver

● とうきょう=と | ー | きょ | ー, サッカー=サ | ッ | カ | ー

● 各モーラに H/L を割り振る。

● 単語を孤立発声したときの H/L パターンは単語属性の一つ。

● カナダ=HLL, バンクーバー=LHHLLL, 東京=LHHH

● 始まる=LHHH, 許可する=HLLL, 謝る=LHHL

● 単語同士の接続／用言の活用→アクセントは変わる（**アクセント結合**）

● 大学=LHHH

● カナダ大学=LHHHLLL, バンクーバー大学=LHHHHHHLLL

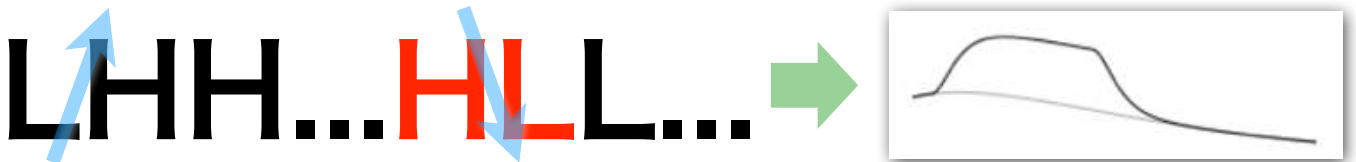
● 謝ります=LHHHHL, 謝ろうとした=LHHHLLL

● **孤立単語のアクセントの知識だけでは運用上、不十分。**

アクセントってそもそも何？

単語アクセントとは？

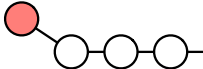
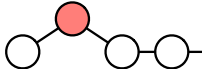
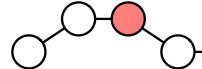
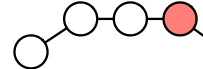
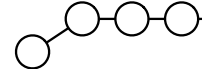
- 単語の一部を目立たせる。どうやって？
 - より強くする：強さアクセント（強勢，弱勢）
 - 英語
 - より高くする：高さアクセント（H，L）
 - 日本語
- 日本語の共通語（東京方言）における単語アクセント
 - 原則として，第1モーラから第2モーラにかけて LH と上がる。
 - そして，単語のどこかで HL と下る。下るともう上がらない。
 - どこで下るのか，によってアクセントは分類される。
 - HL と下る H のモーラ：アクセント核



単語アクセントの分類

アクセント核の位置によって分類

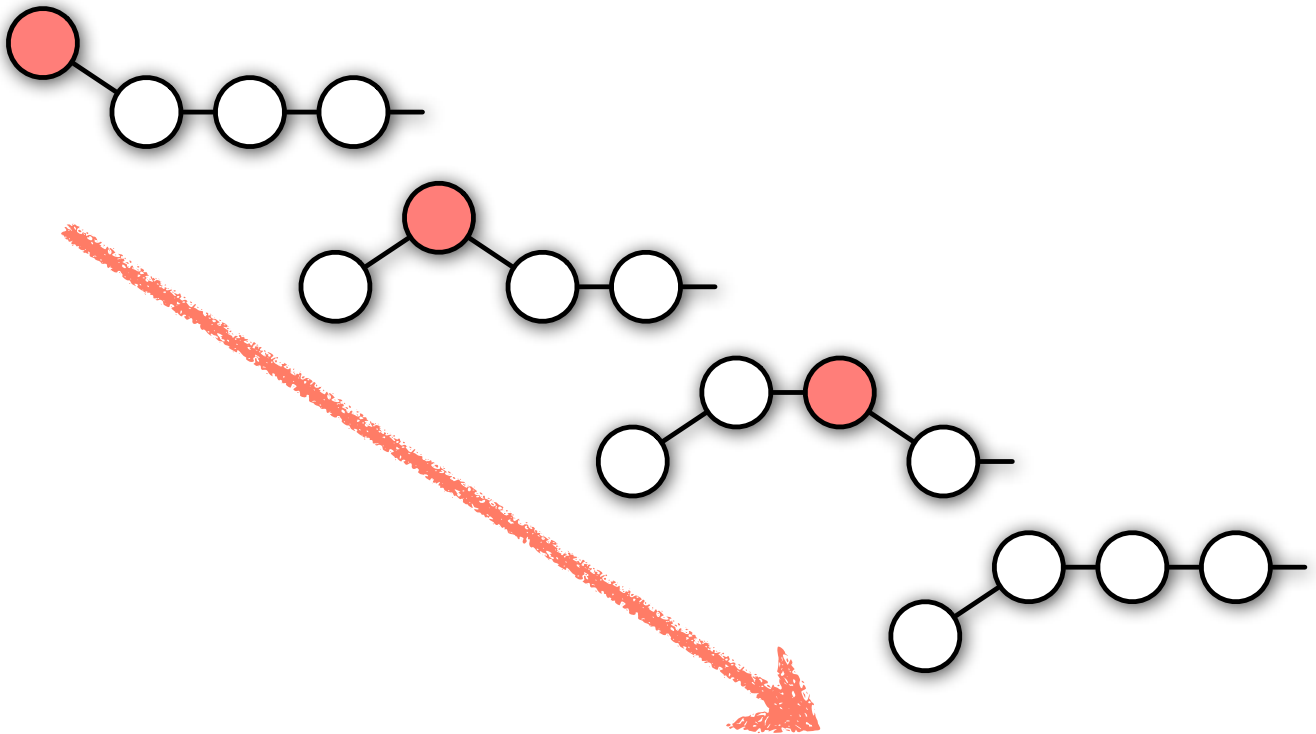
- 複数の命名法（4 モーラ単語を例にとって）

				
さんがつ	ひこーき	かんごふ	いもーと	おはなみ
頭高型	中高型		尾高型	平板型
起伏式				平板式
1 型	2 型	3 型	4 型	0 型 (∞)
-4 型	-3 型	-2 型	-1 型	

- アクセント核は単語中に高々一つ
- ちなみに音声工学の世界では「N型」という呼び方が殆ど

いろんな「ひろしま」

1, 2, 3, 0型で発声する「ひろしま」



テキスト解析結果に基づく区切り

単語・文節・アクセント句・フレーズの区切り

生まれて初めて、バンクーバーに来ました。

↓ **単語（形態素）**に区切る

生まれ（動） | て（助） ■ 初めて（副） ■ バンクーバー（名）
| に（助） ■ 来（動） | まし（助動） | た（助）

↓ **単語境界の一部が文節境界に**

生まれて | 初めて ■ バンクーバーに ■ 来ました

↓ **文節境界の一部がアクセント句境界に**（アクセント的まとまりを作る）

生まれて初めて ■ バンクーバーに | 来ました

↓ **アクセント句境界の一部が句境界に**（イントネーション的まとまりを作る）

生まれて初めて || バンクーバーに来ました （|| ポーズを置く）

テキスト	入力テキスト解析部
	単語分割 品詞同定 係り受け 読み アクセント核 ポーズ位置

一端ばらばらにしてから、まとめ上げる。
韻律は、まとめ上げを助ける糊のようなもの。

似て非なるもの

文節とは

- 自立語に付属語が繋がって構成される文の構成要素
- ある自立語から次の自立語の前までを一つの単位としたもの
- 文中で「ね」を入れて区切れる単位
- 与えられた文をどう発声しても、区切り方は変わらない。
- 生まれて | 初めて | バンクーバーに | 来ました。

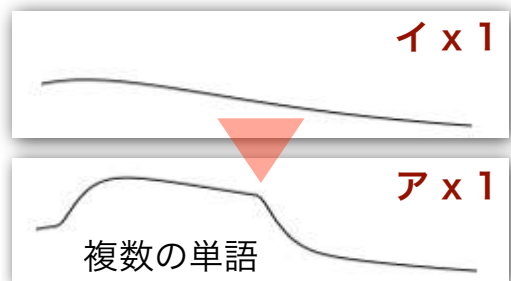
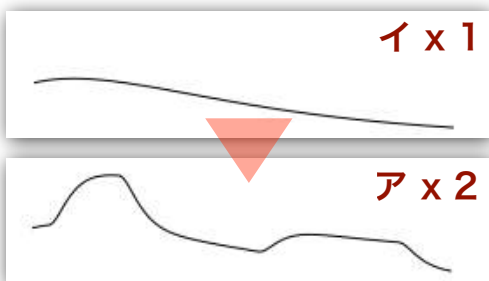
アクセント句とは（日本語独特のピッチ制御）

- 幾つかの語がひとまとまりで発話される時の韻律的単位
- 単語アクセントに見られるピッチ制御が適用される単語**系列**
 - 通常句頭でピッチは上昇し、アクセント核の直後で下落する。
 - アクセント核は通常高々一つ（副次核を認める場合もあり）。
- 与えられた文の発声の仕方で、区切り方は変わる。
 - 生まれて | 初めて | バンクーバーに | 来ました。
 - 生まれて初めて | バンクーバーに | 来ました。

複数の単位と韻律的まとまり

文節・アクセント句・イントネーション句

- 文**: うまれて | はじめて || バンクーバーに | きました。
- ア**: うまれて | はじめて || バンクーバーに | きました。
- ア**: L H H H L H L L L H H L L L L L H L L
- イ**:
- ア**: うまれてはじめて || バンクーバーにきました。
- ア**: L H H H H H L L L H H L L L L L L L L L L
- イ**:



そもそも、どこまでまとめるんだ？

アクセント句の曖昧さ・非決定性

- 生まれて初めて、カナダに来ました。



- うまれて | はじめて | かなだに | きました
- うまれてはじめて | かなだに | きました
- うまれてはじめて | かなだにきました



工学的な解決策の一つ

- ある話者に文章セットを与え、自然な話速で読んだ時に、どの文節境界がアクセント句境界になるのかをラベリング
- アクセント句境界ラベル付きのテキストコーパスを使って、任意のテキストに対して、アクセント句境界を付与する機能を実装
 - 機械学習という枠組みを用いて句境界／核の位置を推定
- その話者のアクセント句境界付与（感覚）を実装する。

しっかりアクセント制御しないと売れない

アクセント句＝常に0型として発声すると

- うまれて | はじめて | かなだに | きました
- うまれて | はじめて | かなだに | きました
 - 恐らく「地方出身者」と認識されると思います。
- 日本語は「方言の違い＝アクセントの違い」となることが多い。
 - 「どこでHLが出現するのかの違い＝方言の違い」となる。
 - 日本では public speaking は、東京方言で行なうことが多い。

単語アクセントと地方性（ちょっと極端な例？）

- 東京での東西南北

	東	西	南	北
方角として	<u>ひがし</u>	<u>にし</u>	<u>みなみ</u>	<u>きた</u>
人名として	<u>ひがし</u>	<u>にし</u>	<u>みなみ</u>	<u>きた</u>

- 浜松での東西南北

	東	西	南	北
方角として	<u>ひがし</u>	<u>にし</u>	<u>みなみ</u>	<u>きた</u>
人名として	<u>ひがし</u>	<u>にし</u>	<u>みなみ</u>	<u>きた</u>

「目立つ」ところは何処にくる？

後ろから数えることがポイント

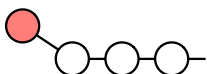
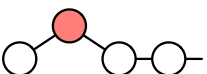
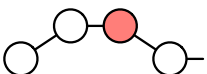
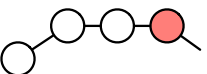
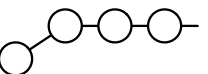
- 窪菌晴夫「日本語の音声」岩波書店（1999）
- 川越いつえ「英語の音声を科学する」大修館書店（1999）
 - 基本的なアクセント付与手法を日本語／英語（名詞）について解説
 - 当然、この手法の通りにならない単語も沢山、存在します。
- 日本語の単語アクセント
 - 語末音節を切り離し、その次から前へ数えて二つ目のモーラを含む音節にアクセント核が存在する。
- 英語の単語アクセント
 - 語末音節を切り離し、その次から前へ数えて二つ目のモーラを含む音節に語強勢が存在する。
 - 原則は日英で変わらない。音節とモーラという二種類の単位が関与。



単語アクセントの分類

アクセント核の位置によって分類

- 複数の命名法（4モーラ単語を例にとって）

				
さんがつ	ひこーき	かんごふ	いもーと	おはなみ
頭高型	中高型		尾高型	平板型
起伏式				平板式
1 型	2 型	3 型	4 型	0 型 (∞)
-4 型	-3 型	-2 型	-1 型	

- アクセント核は単語中に高々一つ
- ちなみに音声工学の世界では「N型」という呼び方が殆ど

日本語教育的には

凡その傾向・基本的なパターンを伝えたい

- 「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」
 - 簡単辞書：動詞，形容詞，複合名詞，接続詞・副詞，数詞，助数詞，更には人名などのアクセント傾向についてまとめ，活用語の場合は活用に伴うアクセント変形について示したもの。
- 動詞のアクセント
 - 平板型か-2型，まれに-3型。
 - 平板型：行く，聞く，知る，変える
 - -2型：食べ'る，読'む，来'る，見'る
 - -3型：帰る，返す，入る，通る
 - 活用形（活用語尾）によってアクセント型が決まる。
 - 平板型：行きます'す，聞きます'せん，知ります'した，変えます'んでした
 - -2型：食べま'す，読みませ'ん，来ま'した，見ませ'んでした
 - -3型：帰ります'す，返しませ'ん，入ります'した，通りませ'んでした



日本語教育的には

凡その傾向・基本的なパターンを伝えたい

- 複合名詞（前部語＋後部語）のアクセント
 - 後部語が3, 4拍で平板型と尾高型の場合は，後部語1拍目にアクセント
 - 南極＋旅行＝なんきょくりょ'こう
 - 後部語が3, 4拍で起伏式の場合は，後部語のアクセントが残る
 - 入国＋管理局＝にゅうこくかんり'きょく
 - 後部語が5拍以上の場合は，後部語のアクセントが残る
 - 原子力＋発電所＝げんしりょくはつでんしょ
 - 中央＋郵便局＝ちゅうおうゆうび'んきょく
 - 後部語が1拍か2拍の短い語の場合は，前部語の最終拍にアクセント
 - 横浜＋市＝よこはま'し
 - 神奈川＋県＝かながわ'けん
 - なお，全体が平板型になる場合もある

- 結局は，無意識的に制御できるまで繰り返し発声する？

音声工学的には

規則を用いてアクセント結合の様子を予測する

● 電子情報通信学会論文誌D, J66-D, 7, 849-856, 1983

論文

UDC 534.782:809.56-148

日本語単語連鎖のアクセント規則

正員 匂坂 芳典[†] 正員 佐藤 大和[†]

Accentuation Rules for Japanese Word Concatenation

Yoshinori SAGISAKA[†] and Hirokazu SATO[†], Regular Members

あらまし 日本語テキストからの音声合成に必要な韻律情報(アクセント, イントネーション)の生成規則を確立するため, アクセント句を構成する単語の性質からアクセント核の位置を決定する規則を検討した. 本規則化では従来の個別的な記述の整理, 分析を行い, 単語間アクセント結合において, 後続単語のアクセント属性として結合アクセント価とアクセント結合様式を提案する. これにより付属語アクセント結合の統一的な規則化を行うと共に, 複合単語についても同様の規則化を図った. また, これらの規則の妥当性を示すため, 自立語に付属語が複数個連なる文節3445文節, 複合単語4877語を用いた推定実験を行った. この結果, 各々98.3%, 95.4%の正解率が得られ, 本規則の有効性が確認された. さらに, これら推定実験結果の分析から, アクセント決定には句を構成する語の性質, 句の構造, 語の用法と意味等が種々に反映されることが示された.

音声工学的には

幾つかの場合に分けて規則を構築

● 名詞+名詞+... → 複合名詞

● 接頭語+自立語 → 複合語

● 自立語+接尾辞 → 複合語

● 自立語+付属語 → 文節

● 文節+文節 → アクセント句

● :

● これらを運用する際の適用則

● 基本的にはアクセント辞書の巻末にある規則と類似している

自立語+付属語 → 文節の場合

● 付属語 → アクセント結合様式と結合アクセント価を定める

● アクセント価=接続後, その語の何モーラ目に核が生じるか?

● アルク+マス → アルキマス (マスは1価)

(N_1 モーラ M_1 型+ N_2 モーラ $\tilde{M}_2$ 価→ N_c モーラ M_c 型)
(a) 付属語アクセント結合規則

結合様式	文節のアクセント型 M_c	
	$N_1 = 0$	$M_1 \neq 0$
(F1) 従属型		M_1
(F2) 不完全支配型	$N_1 + \tilde{M}_2$	M_1
(F3) 融合型	M_1	$N_1 + \tilde{M}_2$
(F4) 支配型	$N_1 + \tilde{M}_2$	
(F5) 平板化型	0	

最近出版されたアクセント辞典

新明解・日本語アクセント辞典@2014/3/21

● CD付き, 「アクセント習得法則」を解説した冊子付き



パソコンに日本語を読ませる技術

私も日本語教師です。でも生徒はパソコンです。

● 日本語テキスト音声合成 (text-to-speech synthesizer) の例

● <http://www.ai-j.jp>



● パソコンに共通語（東京方言）で読ませるにはどうすればよい？

本日のメニュー

最近の PC は上手に喋るけど、誰が日本語教えてる？

- 韻律に着目。単語アクセント・アクセント結合・アクセント句

日本人にとってのアクセントとは？

- 日本人はアクセントの違いに敏感なの？鈍感なの？

日本人にとって分かり易い発声とはどういうものか？

- フレージング（チャンキング）とポージング

OJAD 作りました。ご紹介します。

- 単語検索 / 動詞の後続語検索 / 任意テキスト版

- 韻律音読チュータ・スズキくん



OJAD 使って教えています。指導のヒントをご紹介します。

- フレージング&ポージングによる発声指導&プレゼン指導

まとめ

興味深い聴取実験とその結果

日本人はアクセントの誤りに敏感です。

- 但し、無アクセント地域出身の方を除きます。
- 日本語・米語のバイリンガル話者の音声を使った実験
 - ある単語に対して、「共通語発音」と「米語訛り発音」を収録
 - 音声技術を使って、「共通語発音」から「米語訛り発音」へ音響的に、かつ段階的に変化させた（morphing）。



- 音の高さだけ、音の強さだけ、音の音色だけ、訛らせてみる。
 - 音のある側面だけ「共通語」を「米語訛り」にする。
 - 段階的に「共通語」を「米語訛り」にする。
- どれくらい「米語訛り」にすると「訛ってる！」と分るのか？
 - 被験者は日本人とオーストラリア人



vs.



Use of morphed utterances

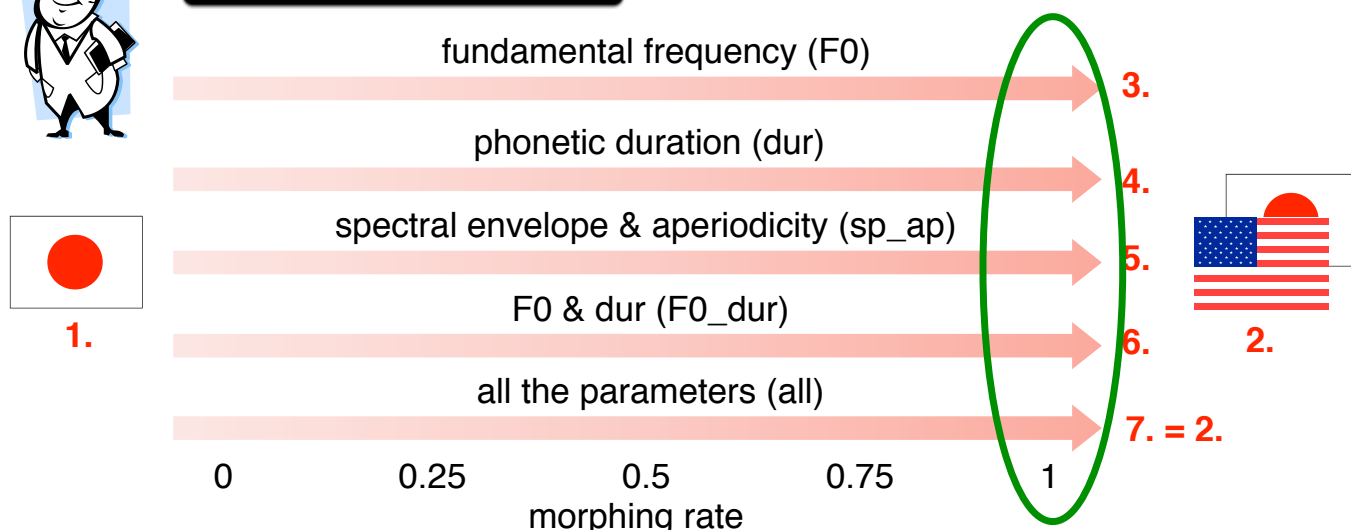
- Morphing of a native utterance and its accented version [Kato+'11]

- Use of a pair of word utterances spoken by a bilingual speaker

- Normal Tokyo Japanese
- Heavily American accented Japanese



igaku (medical science) 医学



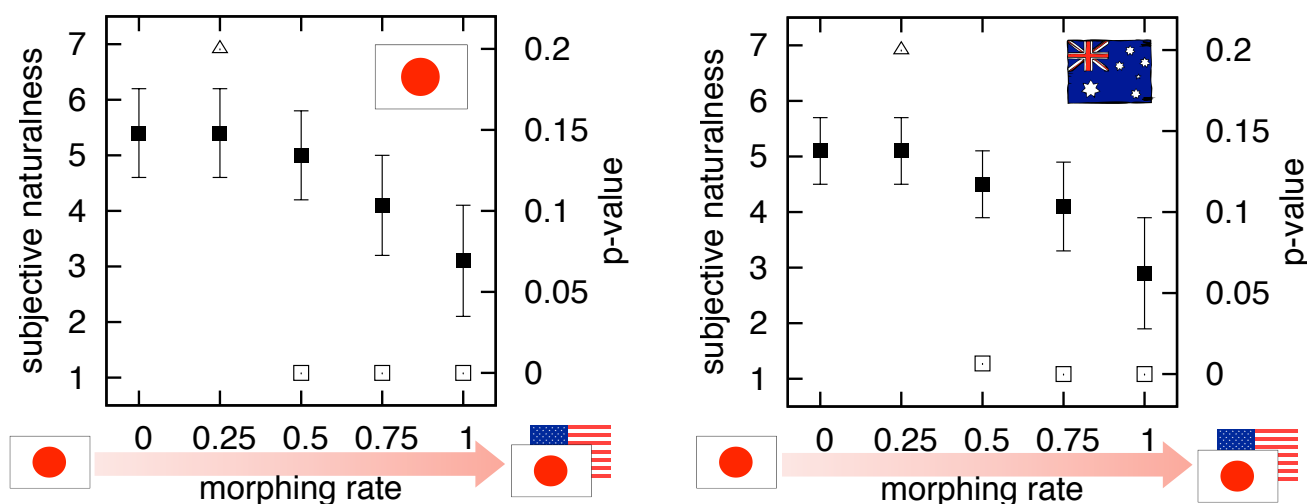
Use of morphed utterances

- Prosodic insensitivity of foreign listeners [Kato+'11]

- 42 Japanese listeners
- 15 Australian listeners
- Judgement of naturalness as Tokyo Japanese



Morphing only in terms of **spectrum & periodicity**

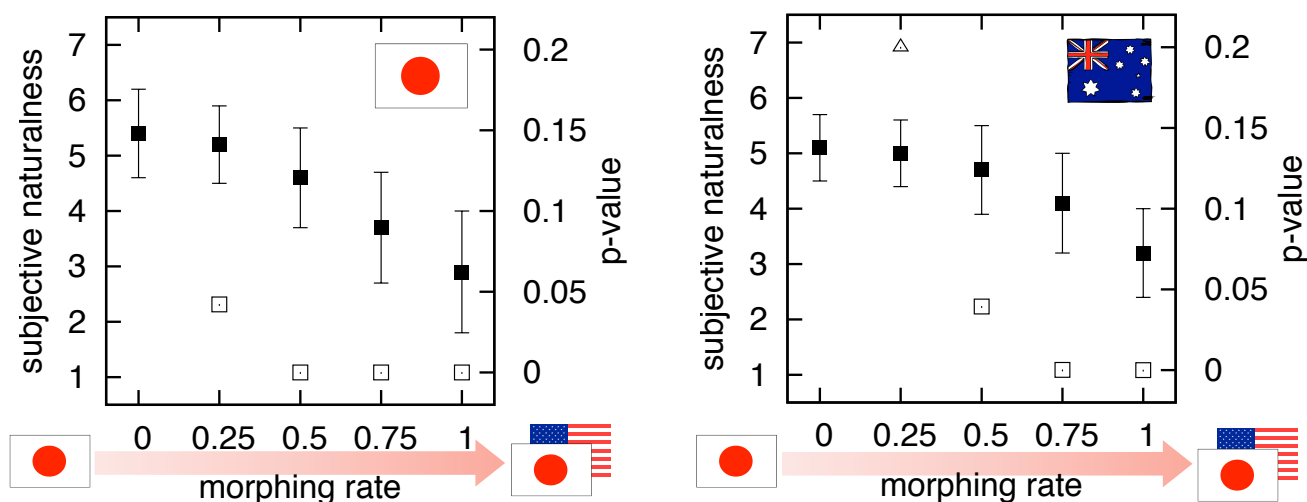


Use of morphed utterances

- Prosodic insensitivity of foreign listeners [Kato+'11]

- 42 Japanese listeners 
- 15 Australian listeners 
- Judgement of naturalness as Tokyo Japanese

Morphing only in terms of **duration**

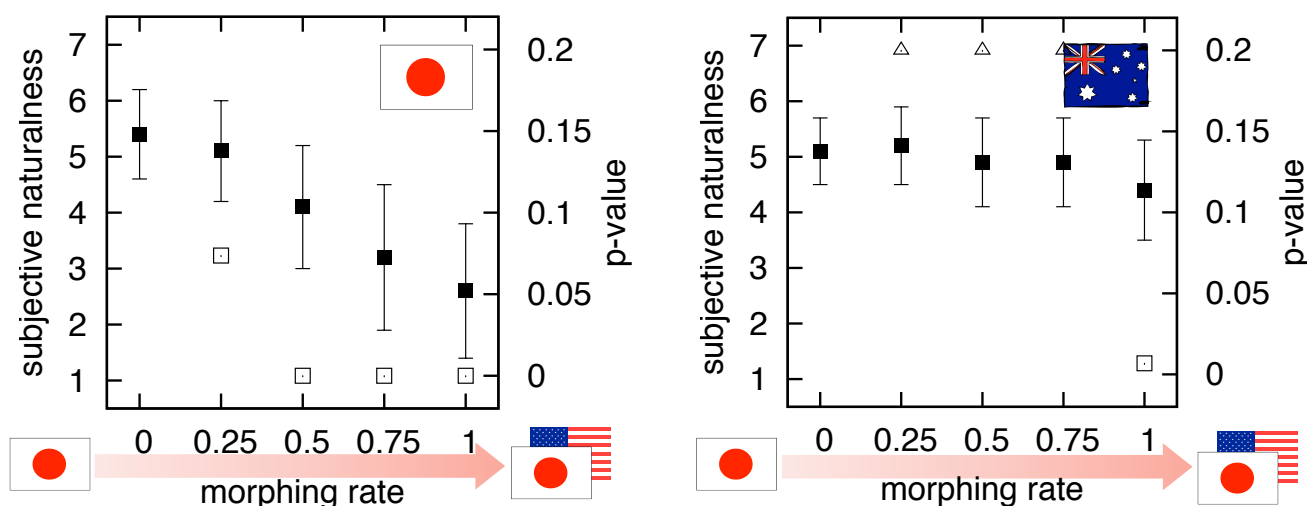


Use of morphed utterances

- Prosodic insensitivity of foreign listeners [Kato+'11]

- 42 Japanese listeners 
- 15 Australian listeners 
- Judgement of naturalness as Tokyo Japanese

Morphing only in terms of **F0**



日本人はアクセントに敏感？鈍感？

日本人なら誰でも各モーラに LH を付与できるのか？

- 日本人にとって LH の制御は無意識的に行なわれる。
- だから、LH を意識的に捉えることが難しい。
 - 当然、上手に自然な日本語（共通語、東京方言）を話します。
 - 小学校でも、音声文法は通常扱っていない。
- 母語話者の日本語教師でも LH を意識的に捉えることは難しい。
- 句（フレーズ）の先頭のアクセント核を指摘するテスト
 - 留学生（多くは中国人）：68.2% 日本人：**61.6%**

yī yí yǐ yì
mā má mǎ mà

Nǐ hǎo!
Xièxie!
Zài jiàn!
Rèliè huānyíng!

おはよう

ありがとう

げんきですか？

本日のメニュー

最近の PC は上手に喋るけど、誰が日本語教えてる？

- 韻律に着目。単語アクセント・アクセント結合・アクセント句

日本人にとってのアクセントとは？

- 日本人はアクセントの間違いに敏感なの？鈍感なの？

日本人にとって分かり易い発声とはどういうものか？

- フレージング（チャンキング）とポージング

OJAD 作りました。ご紹介します。

- 単語検索 / 動詞の後続語検索 / 任意テキスト版
- 韻律音読チュート・スズキクン



OJAD 使って教えています。指導のヒントをご紹介します。

- フレージング&ポージングによる発声指導&プレゼン指導

まとめ

中国人留学生の日本語音声の例

とある訓練に対する before & after を聞き比べる

● 「フード・マイレージ」について

● = 「食物の重さ x 輸送距離」で計算される量

● before

● after

後者の方が聞きやすいと誰もが分るのか？

● 留学生の一部は、それが分らない。



何をやったら after のようになるのか？

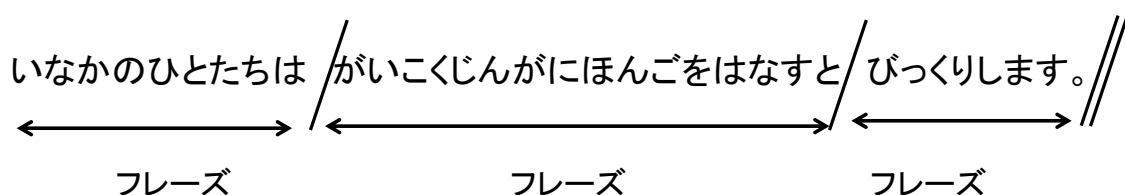
A 簡単なフレージングができる

「句切り」と「への字」(チャンキングとポージング)

⇒大きいリズムをとることができる

1) 文の意味を理解し、(聞きながら)句切りを入れる

2) ピッチカーブを描く



何をやったら after のようになるのか？

A 簡単なフレージングができる
「句切り」と「への字」(チャンキングとポージング)
⇒大きいリズムをとることができる

1) 句切りを入れる

2) ピッチカーブを描く(「へ」の字にする)

いなかのひとたちは / がいこくじんがにほんごをはなすと / びっくりします。 //

何をやったら after のようになるのか？

c 下がり目があるとしたら、それはどこにあるのかに注意して発音することができる
(頭高型アクセント⇒複合名詞のアクセント⇒動詞のアクセント)

1) 文の意味を理解し、句切りを入れる

2) (聞きながら)ピッチカーブを描く(「へ」の字にする: 山か丘)

3) (アクセント辞典を使って)

フレーズの中の最初のアクセント核を見つける

いなかのひとたちは / がいこくじんがにほんごをはなすと / びっくりします。 //

上級学習者・テキストを作る人が使いやすい辞書

テキスト解析結果に基づく区切り

単語・文節・アクセント句・フレーズの区切り

生まれて初めて、バンクーバーに来ました。

↓ **単語（形態素）**に区切る

生まれ（動） | て（助） ■ 初めて（副） ■ バンクーバー（名）
| に（助） ■ 来（動） | まし（助動） | た（助）

↓ **単語境界の一部が文節境界に**

生まれて | 初めて ■ バンクーバーに ■ 来ました

↓ **文節境界の一部がアクセント句境界に**（アクセント的まとまりを作る）

生まれて初めて ■ バンクーバーに | 来ました

↓ **アクセント句境界の一部が句境界に**（イントネーション的まとまりを作る）

生まれて初めて || バンクーバーに来ました （|| ポーズを置く）

一端ばらばらにしてから、まとめ上げる。
韻律は、まとめ上げを助ける糊のようなもの。

入力テキスト
解析部

単語分割
品詞同定
係り受け
読み
アクセント核
ポーズ位置

テキスト

本日のメニュー

最近の PC は上手に喋るけど、誰が日本語教えてる？

● 韻律に着目。単語アクセント・アクセント結合・アクセント句

日本人にとってのアクセントとは？

● 日本人はアクセントの間違いに敏感なの？鈍感なの？

日本人にとって分かり易い発声とはどういうものか？

● フレージング（チャンキング）とポージング

OJAD 作りました。ご紹介します。

● 単語検索 / 動詞の後続語検索 / 任意テキスト版

● 韻律音読チュータ・スズキクン

OJAD 使って教えています。指導のヒントをご紹介します。

● フレージング&ポージングによる発声指導&プレゼン指導

まとめ



OJAD 開発の背景

日本語の音声教育 ～ 需要 > 供給という現実 ～

- 学習者「聞き取り易い、自然な発声を身に付けたい」
- 教師1「でも、授業時間にも制約があるし」
- 教師2「私自身、十分な音声（韻律）教育を受けてないし」



音声（韻律）教育のインフラ作りの重要性

- 非母語話者の教師でも十分に実行できる簡便な教育法
- なるべく短期間で十分な効果が期待できる教育法
- 一つの解として、フレーズ＋ポージング法（中川ら'09）

フレーズ＋ポーズで聞きやすい日本語にはなるが

- より自然な日本語 → フレーズ内単語アクセントの勉強も必要
- アクセント（含変形）を初級者教育へ導入（平野ら'11）
- アクセント（含変形）を教育するための教材不足
 - アクセント辞典は孤立単語のアクセント型がメイン



OJAD 開発の背景



何時・何処でも使える web アクセント辞書

- 単語アクセントやその変形を視覚的／網羅的／聴覚的に呈示
 - アクセントはよく揺れる。（必要に応じて）揺れも含めて呈示
- まずは動詞・い形容詞・な形容詞とその基本活用を対象
 - 用言の基本活用に伴うアクセント変形は比較的規則的
 - 多種多様な用言の後続語表現にも対応する必要あり
- 名詞の（孤立発声時の）アクセントも呈示
- 各種教科書に準拠した形での情報提供（教科書の副読本的）

文音読時の自然なピッチパターンの視覚呈示

- 単語や文節という単位を超えて、文としての韻律が知りたい
- フレーズ＋ポージング法（中川'09）との融合

Online Japanese Accent Dictionary

<http://www.gavo.t.u-tokyo.ac.jp/ojad>

大学共同利用機関法人 人間文化研究機構 ...したオンライン日本語アクセント辞書の開発

国立国語研究所
National Institute for Japanese Language and Linguistics

国語案内 研究組織 研究活動 データベース・データ集 催し物等 刊行物

トップ > 研究活動 > 共同研究プロジェクト > 領域指定型 > 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

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

- 略称 : アクセント辞書
- プロジェクトリーダー : 峯松 信明
(東京大学大学院工学系研究科教授)
- 研究分野・キーワード : <研究分野: 日本語教育学, 音声学, 音声工学>
(日本語教育) (外国語教育) (韻律的特徴) (アクセント結合)
(オンライン辞書)
- プロジェクトの成果 : 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
OJAD (Online Japanese Accent Dictionary)
<http://www.gavo.t.u-tokyo.ac.jp/ojad/>

概要

言語資源研究系
領域指定型共同研究プロジェクト

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

リーダー: 東京大学准教授 峯松 信明

国内外における日本語学習者の増加
→ 自然な音声コミュニケーション能力の獲得を望む学習者

ふりがな氏名	所属・職名	専門分野(学位)
みねまつのぶあき 峯松 信明	東京大学大学院情報理工学系研究科電子情報学専攻・准教授	音声学博士(工学)
ひろせ けいきち 広瀬 啓吉	東京大学大学院情報理工学系研究科電子情報学専攻・教授	音声学工学博士
ひらの ひろこ 平野 宏子	吉林華僑外国語学院(中国) 日本語学部・専任講師	日本語教育・音声学博士(科学)
なかがわ ちえこ 中川 千恵子	早稲田大学日本語教育研究センター・非常勤講師	日本語教育人文科学博士
なかむら のりこ 中村 則子	慶應義塾大学国際センター 非常勤講師	日本語教育文学修士
すずき まさゆき 鈴木 雅之	東京大学大学院工学系研究科電気系工学専攻・博士課程	音声学修士(工学)
たがわ ゆきのり 田川 恭識	早稲田大学日本語教育研究センターインストラクタ(非常勤)	日本語教育文学修士



2013年度日本音声学会学術研究奨励賞授賞報告

2014年3月15日

2013年度日本音声学会学術研究奨励賞選考委員会
委員長 田中ゆかり(日本大学)

2013年度は4件の応募があった。慎重に審査した結果、5人の選考委員全員一致にて、下記研究成果が本奨励賞に値するものという結論を得た。

» 成果の名称: オンライン日本語アクセント辞書の開発と普及
<http://www.gavo.t.u-tokyo.ac.jp/ojad/>

» 受賞者: 峯松信明(東京大学) 他9人

» 研究メンバー(研究成果サイト掲載者)

- » 平野宏子(東北師範大学中国語日本語留学生予備学校)
- » 中川千恵子(早稲田大学)
- » 中村則子(東京外国語大学)
- » 田川恭識(早稲田大学)
- » 中村新芽(東京大学)
- » 広瀬啓吉(東京大学)
- » 橋本浩一(東京大学)
- » 水上智子(東京大学)
- » 鈴木雅之(元・東京大学 現・IBM Research - Tokyo)

» 授賞理由:
本研究成果の授賞理由は、以下の5点とまとめられる。

1. 「オンライン日本語アクセント辞典」の開発は、音声合成の知見を教育に応用したもので、教育的に非常に価値が高い。アクセントの基礎研究と応用が結びついた成果として評価できる。
2. 公開後も、教員や学習者のフィードバックを踏まえて改良を重ねてきた実績をもち、継続性・発展性という点において評価できる。
3. 教室内外での活用方法の提案や学習者の自律的学習の促進、講習会などの開催による普及活動の面で実績がある点も評価できる。

ふりがな氏名	所属・職名	専門分野(学位)
みねまつのぶあき 峯松 信明	東京大学大学院情報理工学系研究科電子情報学専攻・准教授	音声学博士(工学)
ひろせ けいきち 広瀬 啓吉	東京大学大学院情報理工学系研究科電子情報学専攻・教授	音声学工学博士
ひらの ひろこ 平野 宏子	吉林華僑外国語学院(中国) 日本語学部・専任講師	日本語教育・音声学博士(科学)
なかがわ ちえこ 中川 千恵子	早稲田大学日本語教育研究センター・非常勤講師	日本語教育人文科学博士
なかむら のりこ 中村 則子	慶應義塾大学国際センター 非常勤講師	日本語教育文学修士
すずき まさゆき 鈴木 雅之	東京大学大学院工学系研究科電気系工学専攻・博士課程	音声学修士(工学)
たがわ ゆきのり 田川 恭識	早稲田大学日本語教育研究センターインストラクタ(非常勤)	日本語教育文学修士

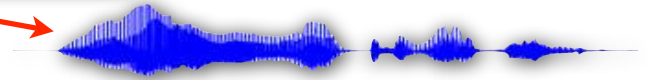


結局、何が必要なのか？

「ひらがな表記」は「読み」になってない。

- 「ひらがな列」の裏に隠された情報をあぶり出す必要あり。
- 「ひらがな列→音声」するために必要な情報を示す必要あり。

ひと がいこくじん にほんご はな
いなかの人たちは、外国人が日本語を話すとびっくりします。



- 平仮名列（音韻列，音素列）を単音列に代えるような作業？
/arajurugeNzituo/ → [ɾɛɹjirigeẽnd͡zitsiɔʔ]

- もちろん，韻律も示しないとダメ

- 音声合成技術が裏で行なっている作業を表舞台に出すことで解決！
- 音声合成器を使わせれば解決？視覚的な呈示が求められている！
- ひらがな列の裏に隠された情報（主には韻律）の視覚化
- 学習者の母語によって，隠されたどの情報が知りたいのかは異なる

結局、何が必要なのか？

「ひらがな表記」は「読み」になってない。

- 「ひらがな列」の裏に隠された情報をあぶり出す必要あり。

2006 年の調査によると，日本全国で，約 33%の家庭がペットを買っているそうです。

ニセ'ン/ロク'ネンノ/チョ'ーサニヨルト_ニホ'ン/ゼ'ンコクデ_ヤ'ク/サ'ンジュー/サンパーセ'ントノ/カテ_ガ_ペ'ットオ/カ'ッテ/イルソ'ーデス%.

グッズの専門店もでき，お洒落な服を着た犬も，よく見かけます。
グッズノ/センモ'ンテンモ/デ'キ_オシャ'レナ/フ%ク'オ/キ%タイヌ'モ_ヨ'ク/ミカケマ'ス%.

':アクセント核，/:アクセント句境界，_:ポーズ，%:母音の無声化

- ひらがな列の裏に隠された情報（主には韻律）の視覚化
- 学習者の母語によって，隠されたどの情報が知りたいのかは異なる

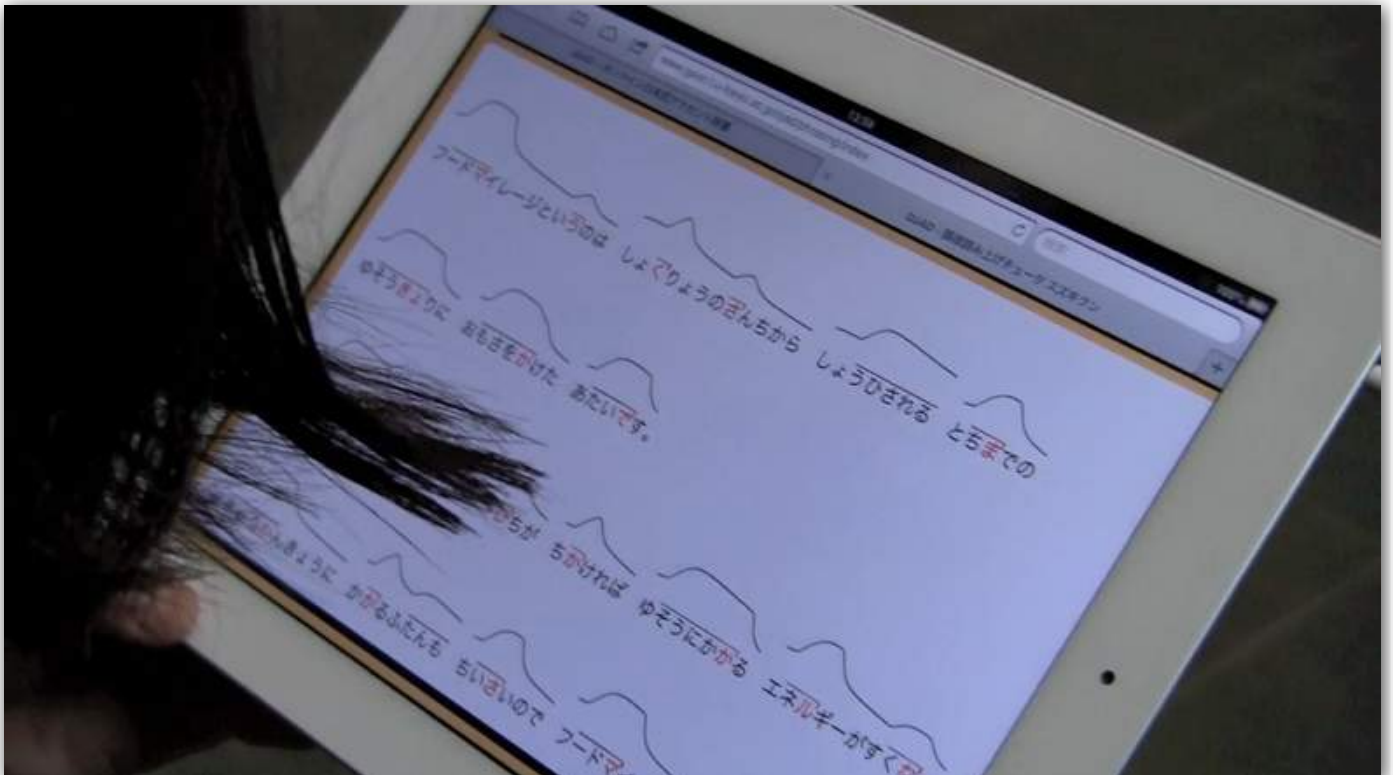
一つお願い

● 技術は完全ではありません。時として誤ります。

- 誤ることがある技術に関しては、正直にそう言います。
- ですので、それを承知の上で使っていただけたら嬉しいです。



1min PV を作りました



OJAD 4機能

単語検索

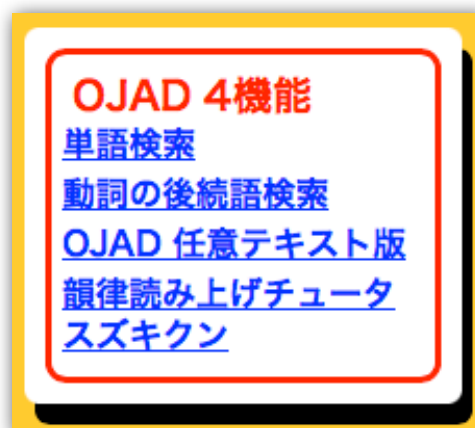
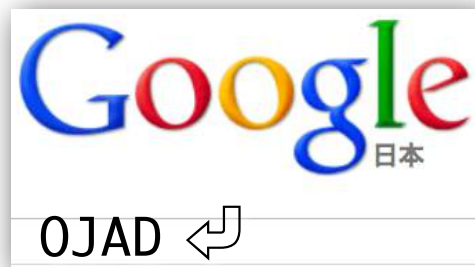
動詞の後続語検索

OJAD 任意テキスト版

● 上記三つはアクセント学習がメイン

韻律音読チュータ・スズキクン

● イントネーション+アクセント学習



三種類の OJAD 利用法解説ページ

OJAD 4機能のご紹介

● いわゆる「取り扱い説明書」的なページです。

使ってみよう OJAD → 使ってみよう XXX

● 以下の3機能について「使うことで学ぶ」ためのページです。

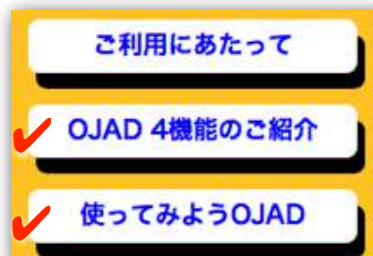
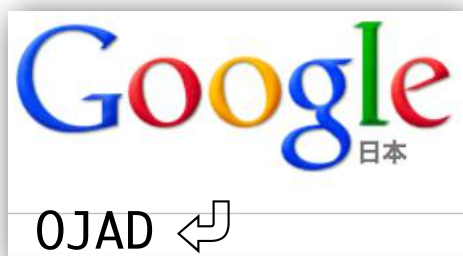
● 使ってみよう単語検索

● 使ってみよう後続語検索

● 使ってみようスズキクン

使ってみよう OJAD → タブレット端末用 OJAD デモ

● 学会デモ発表用のページ, 「手っ取り早く学ぶ」ページです。



教科書ガイドとしての単語検索



動詞, い形容詞, な形容詞, 名詞, 約12,700単語

動詞の活用とアクセント

※ 1グループの動詞	辞書形	～ます形	～て形	～た形	～ない形	～なかった形	～ば形	使役形	～う形
弾く・弾きます	ひく ひくひく	ひき ま す ひくひく	ひいて ひくひく	ひいた ひくひく	ひかない ひくひく	ひか な かった ひくひく	ひ け ば ひくひく	ひかせる ひくひく	ひ ご う ひくひく
洗う・洗います	あらう あらう	あら い ます あらう	あらって あらう	あらった あらう	あらわない あらう	あら わ なかった あらう	あら え ば あらう	あらわせる あらう	あら お う あらう
歌う・歌います	うたう うたう	うた い ます うたう	うたって うたう	うたった うたう	うたわない うたう	うた わ なかった うたう	うた え ば うたう	うたわせる うたう	うた お う うたう
泊まる・泊まります	とまる とまる	とま り ます とまる	とまって とまる	とまった とまる	とまらない とまる	とま ら なかった とまる	とま れ ば とまる	とまらせる とまる	とま ろ う とまる
登る・登ります	のぼる のぼる	のぼ り ます のぼる	のぼって のぼる	のぼった のぼる	のぼらない のぼる	のぼ ら なかった のぼる	のぼ れ ば のぼる	のぼらせる のぼる	のぼ ろ う のぼる
なる・なります	な る なる	なり ま す なる	な って なる	な った なる	な ら ない なる	な ら なかった なる	な れば なる	ならせる なる	な ろ う なる
※ 2グループの動詞	辞書形	～ます形	～て形	～た形	～ない形	～なかった形	～ば形	使役形	～う形
換える・換えます	かえる かえる	か え ます かえる	かえて かえる	かえた かえる	かえない かえる	か え なかった かえる	か え れば かえる	かえさせる かえる	か え よう かえる
捨てる・捨てます	すてる すてる	す て ます すてる	すてて すてる	すてた すてる	すてない すてる	す て なかった すてる	す て れば すてる	すてさせる すてる	す て よう すてる
出来る・出来ます	で き る できる	でき ま す できる	で きて できる	で きた できる	で き ない できる	で き なかった できる	で き れば できる	できさせる できる	で き よう できる
集める・集めます	あ つ める あつめる	あ つ め ま す あつめる	あ つ めて あつめる	あ つ めた あつめる	あ つ めない あつめる	あ つ め な かった あつめる	あ つ め れ ば あつめる	あ つ めさせる あつめる	あ つ め よ う あつめる
0型 or -2型 0型 or -2型 0型 or -3型 0型 or -3型 0型 or -3型 -4型 or -5型 -2型 or -3型 0型 or -2型 -2型									

単語検索の検索・表示条件

何を検索し、どう表示したいのか？

検索・表示条件

検索の条件 教科書 指定しない から まで
品詞 すべて アクセント型 すべて 単語長 すべて
難易度 すべて から 上級後半 まで
旧能力試験 すべて から 1級 まで

表示の条件 表示順序 アクセント型順に並べる
カスタマイズ 1. アクセント 2. アクセント 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 非表示
教科書と難易度 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

何を検索するのか？
絞り込み方法の指定

検索結果をどう表示
するのか？

その他のオプション

検索語の個別指定

さあ、これで今日から貴方も検索上手？

韻律音読チュータの二種類のモード

入力文=フレーズ1+フレーズ2+フレーズ3+・・・

● 生まれて初めて、バンクーバーに来ました。

とても住みやすい所と、聞いています。

● フレーズ=複数のアクセント句

● || 生まれて | 初めて || バンクーバーに | 来ました ||

|| とても | 住みやすい | 所と || 聞いています ||

● || フレーズ句区切り | アクセント句区切り

● 上級者用と初級者用

● 上級者用：全てのアクセント句に必要な核を置く。

● 初級者用：自然性を保ちながら核の数を減らすモード

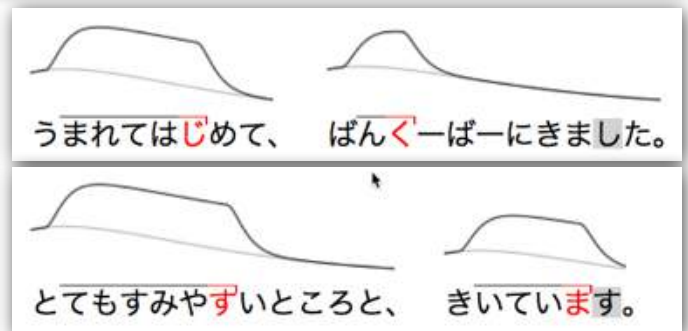
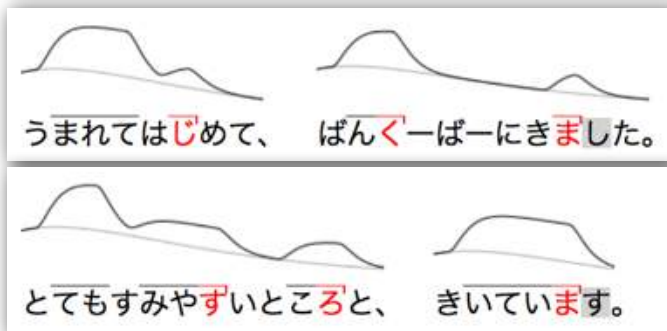
=フレーズ冒頭アクセント核+

3モーラ以上の1型アクセント句の核だけを残して、
後は核を消失させる。

韻律音読チュータの二種類のモード

上級者モードと初級者モードの表示

生まれて初めて、バンクーバーにきました。とても住みやすい所と、聞いています。



- 初級者モード＝自然性を落とさずに、フレーズをアクセント句に
- アクセント句に沿ってL/Hを割り当てると、アクセント句境界でHL（ピッチの下降）となることが多い。
 - これを「アクセント核」と呼ぶべきかどうかは分りません。
- 無声化すべきモーラを網かけ表示
 - 無声化モーラがアクセント核になる場合は、有声に戻しています。

ピッチエディタとしてのスズキくん

アクセント推定の間違いに気付いたら

- そのモーラをクリック。H/L が反転します。
- 但し、合成音声に反映されるのは上級&上級の時のみ
- 視覚表示だけなら、様々な方言アクセントも示せる!?

例えば、「ありがとう」を例にすると、

ありがとう

「り」に高アクセントをおくと、標準語、

ありがとう

「が」に高アクセントをおくと、名古屋弁、

ありがとう

「と」に高アクセントをおくと、関西弁、

ありがとう

「とう」に高アクセントをおくと、鹿児島弁。

と、面白いように方言になってしまいますよね！



ありがとう

ありがとう

ありがとう

ありがとう

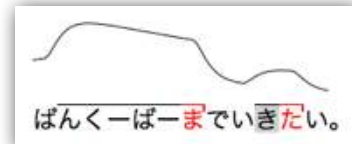
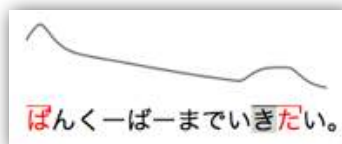
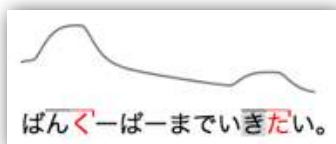
全てのピッチパターンで読み上げる訳ではありません

● ピッチ及びアクセントの条件

- ピッチパターン＝**上級**，テキスト上のアクセント＝**上級**
- ピッチパターン＝**初級**，テキスト上のアクセント＝**初級**
- 但し「初&初」でも，呈示される音声は「上級」です。

● ピッチエディタによる H/L の修正について

- ピッチ＝上，アクセント＝上の場合のみ，修正は反映されます。
- ピッチ＝初，アクセント＝初の場合，合成音は「元の上級」のまま。
 - 画面上では H/L は反転します。
- 東京方言として不適切なアクセントを指定すると，合成できません。
 - Nモーラアクセント句＝0, 1, 2, ..., N 型のいずれかのみ
- 無声化母音をアクセント核として指定すると，合成できません。



二種類のアクセント句境界推定

● アクセント句境界推定の二つの方針

工学的な解決策の一つ

- アクセント句境界ラベル付きのテキストコーパスを使って，任意のテキストに対して，アクセント句境界を付与する機能を実装
 - その後，個々のアクセント句内でアクセント核位置を推定する。
 - **ア句＝適度な長さ**，となるが，境界／核位置推定を誤ることもある。
- 文節＝アクセント句，と解釈して文節境界を句境界とする
 - 各文節（最短のアクセント句）に対して，アクセント核位置を推定。
 - 誤りは減るが，アクセント句が短くなり，ドタバタしたパターンに。
 - 但し，**初級者モードにすればフレーズ句がアクセント句に**。
- どちらを使うべき？
 - 合成音声は，自然性が**僅かに**高い前者（機械学習）
 - アクセント制御を勉強する場合は，**僅かに**誤りが少ない後者（文節）

うまれては**じ**めて | **か**なだに | き**ま**した
うまれて | は**じ**めて | **か**なだに | き**ま**した

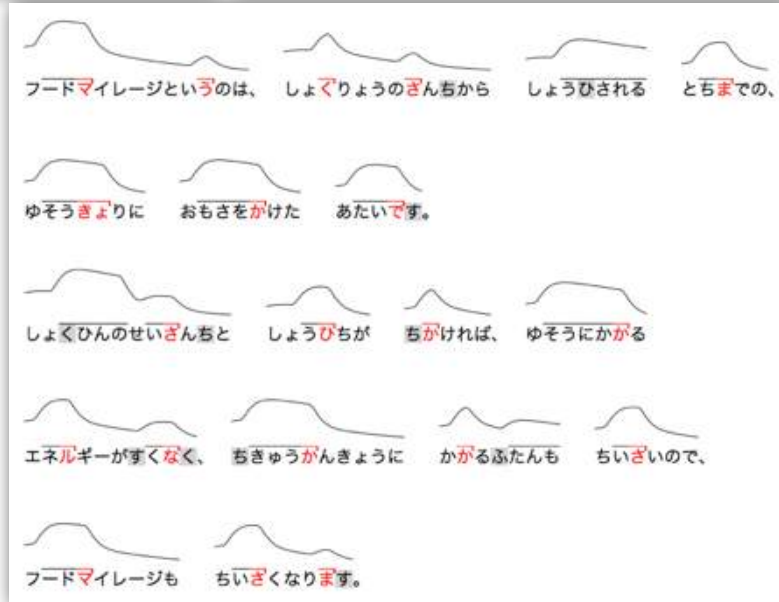
アクセント句境界推定

✓ 機械学習による句境界推定
文節境界を利用

あの留学生が聞いていた音声



フードマイレージというのは、食糧の産地から / 消費される / 土地までの、輸送距離に / 重さをかけた / 値です。食品の生産地と / 消費地が / 近ければ、輸送にかかる / エネルギーが少なく、地球環境に / かかる負担も / 小さいので、フード・マイレージも / 小さくなります。



本日のメニュー

最近の PC は上手に喋るけど、誰が日本語教えてる？

- 韻律に着目。単語アクセント・アクセント結合・アクセント句

日本人にとってのアクセントとは？

- 日本人はアクセントの間違いに敏感なの？鈍感なの？

日本人にとって分かり易い発声とはどういうものか？

- フレージング（チャンキング）とポージング

OJAD 作りました。ご紹介します。

- 単語検索 / 動詞の後続語検索 / 任意テキスト版
- 韻律音読チュータ・スズキクン



OJAD 使って教えています。指導のヒントをご紹介します。

- フレージング&ポージングによる発声指導&プレゼン指導

まとめ

OJADを使った音声教育

中村則子

メニュー

- 1. はじめに
- 2. OJADを使った授業案
 - ①スピーチ
 - ②『初級日本語』 会話
- 3. 終わりに

メニュー

- ・ 1. はじめに
- ・ 2. OJADを使った授業案
 - ①スピーチ
 - ②『初級日本語』 会話
- ・ 3. 終わりに

1. はじめに

音声指導をしていますか

- ・ 毎日の進度がきつく、音声指導に時間が割けない。
- ・ 文法などと違って、何が間違いか、はっきり言えないし、分からない。
- ・ 何を指導すればいいか、指導法が分からない。
- ・ 教材が少ない。研究が少ない。
- ・ 教師が習った経験がないので、学生に教えられない。
（非母語話者教師、日本人教師）
- ・ 方言の問題をどうするか。
- ・ そもそも指導する必要があるのか。

日本語音声の特徴

- ・**単音**＝分節的特徴；
言語の**明瞭性**に関与
ざ、じゃ、な、ら、清濁音、特殊拍（促音、長音、撥音）
- ・**韻律**＝超分節的特徴；
言語の**自然性**に関与
アクセント、イントネーション、リズム、ポーズ

個人的な体験：**音声教育は難しい**と思っていました

- ・**単音が大事？**（ミニマルペア/矯正）

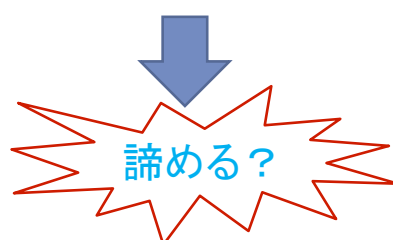
てんき/でんき（どっち？）

おはようございます（→ざ！）

} →直りにくい→無力感

母語によって問題点異なる

→クラスで扱いにくい

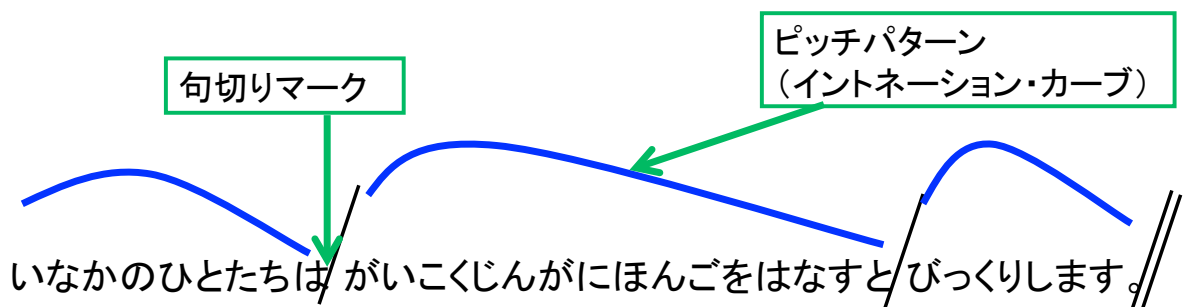


個人的な体験: **韻律**に注目した指導
(**フレージング指導法**)

- ・韻律は**どの学習者にとっても問題**(クラスで扱える)
- ・単音に比べて、**直りやすい**(例: ポーズは誰にでもできる)
- ・**「句切り」と「への字」**(チャンキングとポーシング)

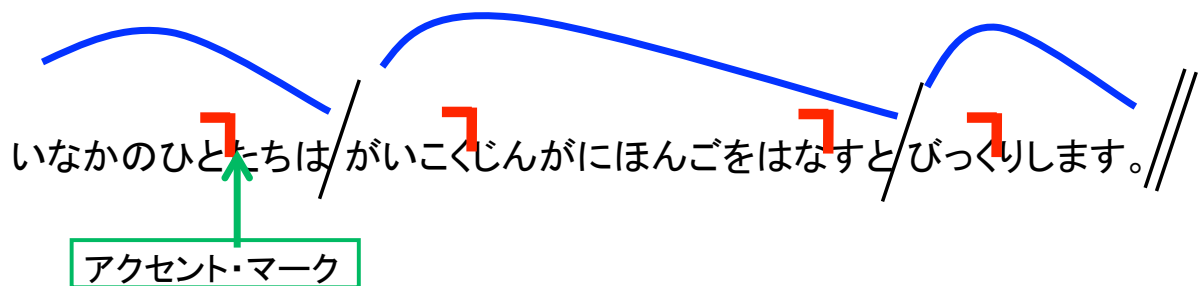
一番 簡単なフレージング指導法

- 1) 句切りを入れる
- 2) ピッチパターンを描く(「への字」の字にする)



次の段階：アクセントで下げる

- 1) 文の意味を理解し、句切りを入れる
- 2) (聞きながら)ピッチパターンを描く(「へ」の字にする)
- 3) (アクセント辞典を使って)
フレーズの中の**アクセント核**を見つける



しかし

- ・一語一語アクセント辞書で調べるのは面倒！
- ・動詞や形容詞の活用形アクセントは複雑！

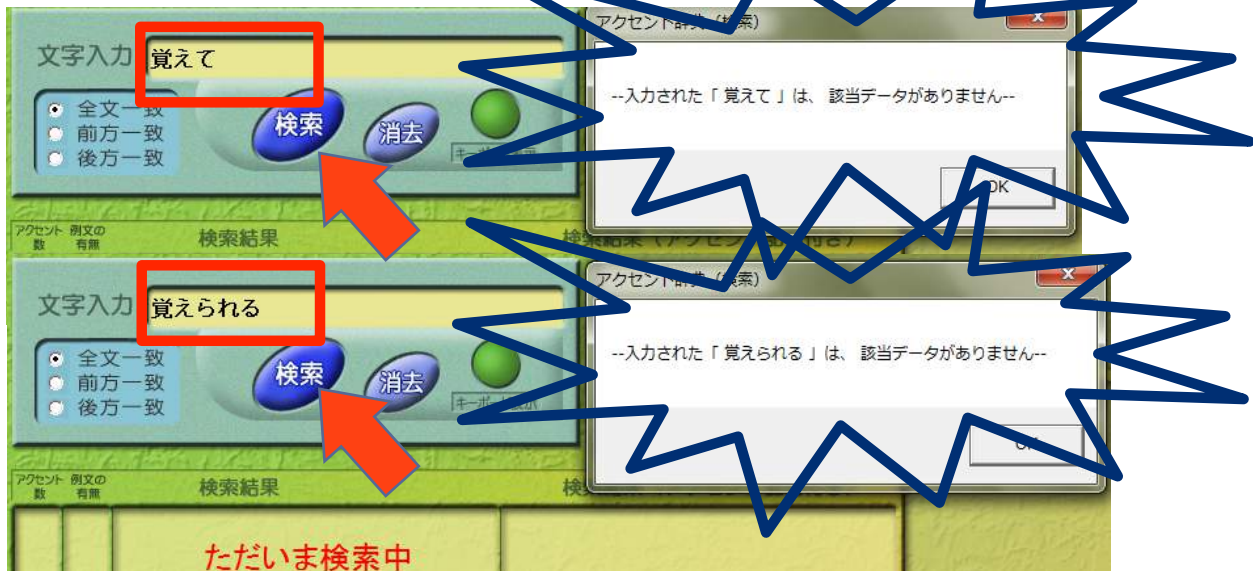
例えば、

「覚える」「覚えて」「覚えられる」

のアクセントを調べたい



アクセント辞書を使う



12

1. はじめに

巻末の規則の説明 学習者向きではない

		動詞に助動詞が付いたときのアクセント								
		動詞の終止形・連体形に付く助動詞			動詞の未然形に付く助動詞			動詞の連用形に付く助動詞		
名詞1	平板式動詞	2拍語	1段活用	イル 居る	イルヨーダ	イルダロー	イルマイ	イサセル	イナイ	イヨー
		5拍語	活用	ナル 鳴る	ナルヨーダ	ナルダロー	ナルマイ	ナラセル	ナラナイ	ナロー
		3拍語	1段活用	ハレル 腫れる	ハレルヨーダ	ハレルダロー	ハレルマイ	ハレサセル	ハレナイ	ハレヨー
		5拍語	活用	アラウ 洗う	アラウヨーダ	アラウダロー	アラウマイ	アラワセル	アラワナイ	アラオー
		4拍語	1段活用	クラベル 比べる	クラベルヨーダ	クラベルダロー	クラベルマイ	クラベサセル	クラベナイ	クラベヨー
		5拍語	活用	オコナウ 行う	オコナウヨーダ	オコナウダロー	オコナウマイ	オコナワセル	オコナワナイ	オコナオー
		サ行 変格活用		カンズル 感ずる	カンズルヨーダ	カンズルダロー	カンズルマイ	カンジサセル	カンジナイ	カンジヨー
		2拍語	1段活用	イル 射る	イルヨーダ	イルダロー	イルマイ	イサセル	イナイ	イヨー
		5拍語	活用	ナル 成る	ナルヨーダ	ナルダロー	ナルマイ	ナラセル	ナラナイ	ナロー
		3拍語	5拍語	カエル 帰る	カエルヨーダ	カエルダロー	カエルマイ	カエラセル	カエラナイ	カエロー
名詞2	起伏式動詞	3拍語	5拍語	ナラウ 言う	ナラウヨーダ	ナラウダロー	ナラウマイ	ナラワセル	ナラワナイ	ナラオー
		4拍語	1段活用	ハレル 腫れる	ハレルヨーダ	ハレルダロー	ハレルマイ	ハレサセル	ハレナイ	ハレヨー
		5拍語	活用	シラベル 調べる	シラベルヨーダ	シラベルダロー	シラベルマイ	シラベサセル	シラベナイ	シラベヨー
		4拍語	5拍語	テツダウ 手伝う	テツダウヨーダ	テツダウダロー	テツダウマイ	テツダワセル	テツダワナイ	テツダオー
数詞1	中高型	4拍語	サ行	シンズル 値する	シンズルヨーダ	シンズルダロー	シンズルマイ	シンジサセル	シンジナイ	シンジヨー
		5拍語	活用	シラベル 調べる	シラベルヨーダ	シラベルダロー	シラベルマイ	シラベサセル	シラベナイ	シラベヨー
数詞2	中高型	5拍語	活用	テツダウ 手伝う	テツダウヨーダ	テツダウダロー	テツダウマイ	テツダワセル	テツダワナイ	テツダオー
		サ行	活用	シンズル 値する	シンズルヨーダ	シンズルダロー	シンズルマイ	シンジサセル	シンジナイ	シンジヨー

新明解 日本語アクセント辞典 第2版
2014年4月10日発行



アソブ 遊ぶ アソバナイ、アソポー、アソビマス、
アソンデ、アソベバ、アソベ

但し、すべての動詞の活用形アクセントが
出ているわけではない

OJAD 4 機能

- 単語検索
- 動詞の後続語検索
- 任意テキスト版
- 韻律音読チュータ・スズキクン

OJAD 4機能

単語検索

動詞の後続語検索

OJAD 任意テキスト版

韻律読み上げチュータ

スズキクン



OJAD ↩

メニュー

- ・ 1. はじめに
- ・ 2. OJADを使った授業案
 - ①スピーチ
 - ②『初級日本語』 会話
- ・ 3. 終わりに

OJADを使った授業案

①スズキクンを使ったスピーチ指導

- ・ 指導の流れ
 1. 「区切り」と「への字型イントネーション」の重要性
 2. 区切りの入れ方
 3. スピーチ原稿作成（必要に応じて添削）
 4. 区切りを入れる
 5. **スズキクン**に入力
 6. ピッチパターンを見ながら音声聞いて発音練習
（個人→グループ）
 7. 発表

1. 音声を聞きましょう。

2.OJADを使った授業案



どちらが、聞きやすいですか。なぜですか。

2.OJADを使った授業案

「音声を聞いて、／で切ってみましょう」

フード・マイレージというのは、食糧の産地から消費される土地までの
 輸送距離に重さをかけた値です。食品の生産地と消費地が近ければ、輸送にかか
 るエネルギーが少なく、地球環境にかかる負担も小さいので、フード・マイレー
 ジも小さくなります。

／で1，／／で2数えるつもりで読みましょう

普通、学習者は、漢字の読みがわからないときにポーズを置きます

フード・マイレージというのは、／^{しょくりょう}食糧の^{さんち}産地から／^{しょうひ}消費され
 る^{とち}土地までの／^{ゆそうきより}輸送距離に／^{おも}重さをかけた値です。／／^{しょうひん}食品
 の^{せいさんち}生産地と／^{しょうひち}消費地が近ければ、／^{ちか}輸送にかかる^{ゆそう}エネルギー
 が^{すく}少なく、／^{ちきゅうかんきょう}地球環境にかかる^{ふたん}負担も／^{ちい}小さいので、／^{ちい}フード・
 マイレージも／小さくなります。／／

読みに集中するために、ルビがあったほうがいい

への字を意識して読みましょう

フレーズの終わりは低く、弱く

フード・マイレージというのは、／^{しょくりょう}食糧の^{さんち}産地から／
^{しょうひ}消費される^{とち}土地までの／^{ゆそうきより}輸送距離に／^{おも}重さをかけた値です。／
^{しょうひん}食品の^{せいさんち}生産地と／^{しょうひち}消費地が近ければ、／^{ちか}輸送にかかる^{ゆそう}エネルギー
 が^{すく}少なく、／^{ちきゅうかんきょう}地球環境にかかる^{ふたん}負担も／^{ちい}小さいので、／
 フード・マイレージも／^{ちい}小さくなります。／／

2. 区切りの入れ方

- ・ 句読点 「。」の後は特にゆっくり休む
- ・ 列挙文 「と」「や」の後
ど サムギョプサルや／タッカルビな
- ・ 接続詞の後 しかし／ それで／ では／
- ・ 「は」の後 私が言いたいのは／そういうこと
じゃないんです。
- ・ 強調したいところの前後 締め切りは／今週じゃなくて／
来週の／土曜日です。

ポーズの長さ

スズキクンは、句読点(、。、.)と／を区切りと認識する。
句読点の方が、／よりポーズが長い。

フード・マイレージというのは、食糧の産地から／

消費される土地までの／輸送距離に、重さをかけた値です。



3. スピーチ原稿作成

- ・ 時間、文字数などを指定して、原稿作成
300字／1分 程度
- ・ 文法的誤り、チェック
- ・ 区切りを入れる

今日は学習者のスピーチ原稿を使う

4. 原稿をスズキクンに入力

OJADトップページ >「使ってみようOJAD」>「タブレット端末用

OJADデモ」>4-4サンプル文章 フレーズ区切り入り>9.「納豆キム

チ」の下「入力済みのスズキクンを表示する」をクリック

モードを選択

・ピッチパターン

山、丘のみを表現したカーブ
初級者用

上級者用

非表示

・テキスト上のアクセント

初級者用

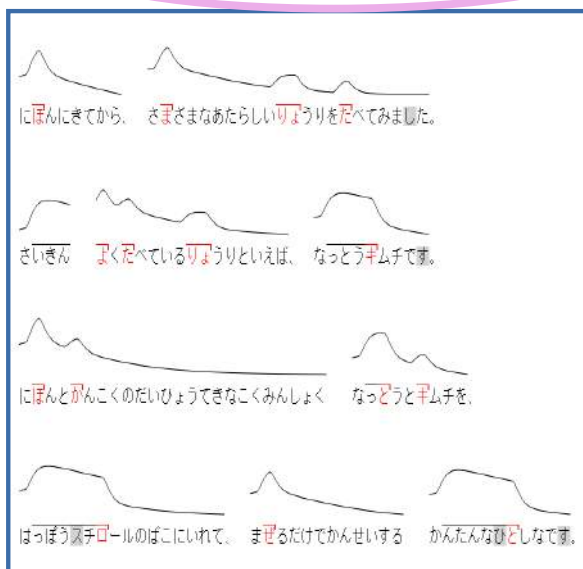
上級者用

非表示

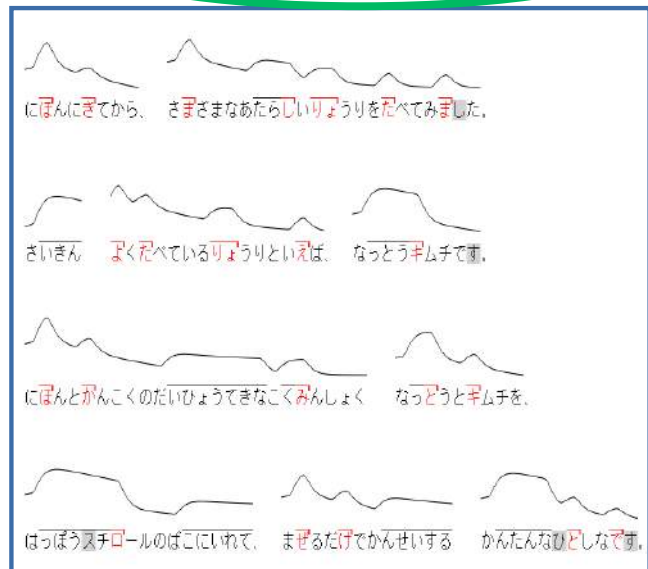
テキストも含めて非表示

(音声はピッチパターン、アクセントともに上級者用)

初級者用ピッチパターン



上級者用ピッチパターン



初級者用

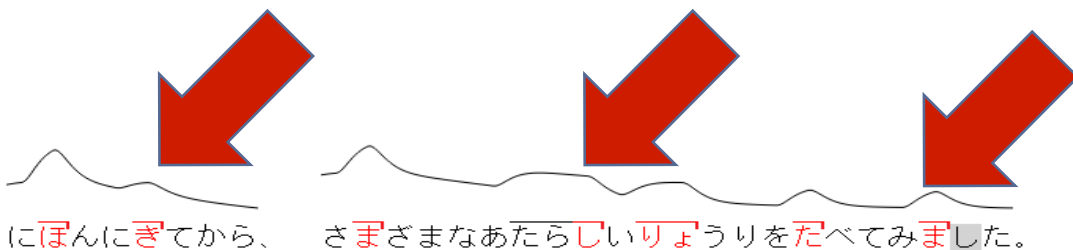
フレーズの1つ目＋頭高(3拍以上)

頭高



上級者用

すべてのアクセント



どちらを選択すべきか

初級者用

- ・声の上げ下げに慣れていない人
- ・とりあえず、簡単に、うまく聞けるように話したい人

上級者用

- ・正確なアクセントを知りたい人
- ・完璧な発音を目指す人

名前に問題あり？

フレーズ末に向かってどんどん下げるのがポイント

入力して「実行」をクリック

日本に来てから、様々な新しい料理を食べてみました。最近よく食べている料理と言えば、納豆キムチです。日本と韓国の代表的な国民食、納豆とキムチを、発泡スチロールの箱に入れて、混ぜるだけで完成する、簡単な一品です。調味料の方は、キムチの唐辛子だけでも、口がピリピリするほどからさを感じて、からしを加えると余計につんつんしますから、これ以上、特に加える必要はないと思います。また個人的には、タレのあっさりした味が好きですから、いつも入れています。このねばねばしたものをそのまま食べても、さっぱりした食感で美味しいですが、ご飯にかける食べ方を、おすすめします。ぐちゃぐちゃの見た目は、なんだか気持ち悪くて、食欲が湧かないかもしれませんが、栄養効果が抜群だそうです。まずは、血液をサラサラにする効果が、期待できそうです。特に、玉ねぎをみじん切りにして一緒に混ぜた場合、効果は、より顕著だと思われます。つぎに、納豆の納豆菌と、キムチの乳酸菌は、相互に増殖させる特性がありますから、消化を助ける効果が著しくて、ダイエット食品にも使われているそうです。そのほかには、がんの予防効果や、育毛の効果も期待できそうです。安くて、手軽に出来上がる割に、すごい栄養効果のある納豆キムチを、ぜひ試してみてください。

ピッチパターン

 テキスト上のアクセント

 アクセントマーク

 アクセント句境界推定

 読み・アクセント推定

 ピッチパターン表示用パラメータ

ところが。。。

よく見ると、

修正が必要(>_<)

スズキクンの弱点1

ときどき、漢字が読めない

はっぼうスチロールの**はこ**にいれて、

たま**ね**ぎをみじん**ぎ**りにして

→解決法: ひらがなで入力し直す

発泡スチロールの**はこ**に入れて

たまねぎをみじん**ぎ**りにして

スズキクンの弱点2

- ・アクセントの誤りがあるかも。。。。

なっ**ど**うのなっ**と**う**ぎ**んと

ギムチのにゅうさん**ぎ**んは

画面上で
アクセントが直せます(*^_^*)

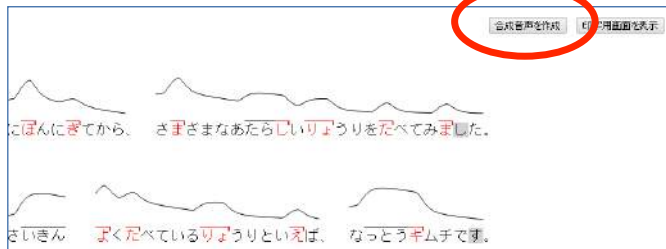
修正方法

ひらがな1文字をクリック→HがLに
LがHに

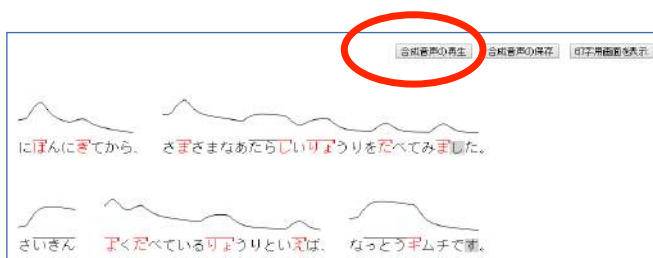
納豆菌と乳酸菌のアクセントを直す

音声を聞く

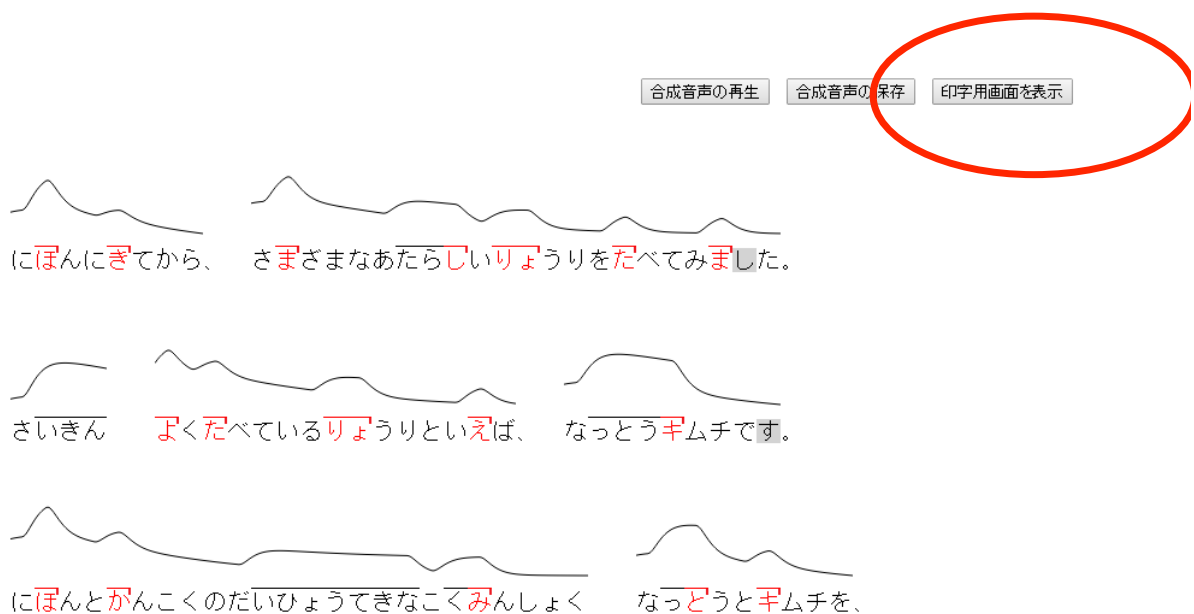
- 画面右上の「音声合成を作成」をクリック



- 「音声合成作成中」→「合成音声の再生」をクリック



印刷



印刷画面

にほんにきてから、さまざまあたらしいりょうりをたべてみました。

さいきん よくたべているりょうりといえば、なっとう芋モチです。

にほんとかんこくのだいひょうてきなこくめんしよく なっとうと芋モチを、はっばうスチロールのばこにいれて、まぜるだてでかんせいする

かんたんなひょうじなです。

ちょうみりょうのほうは、芋モチのとうがらしだてでも

おくとくわえるむつようはないとおもいます。

右クリックして
印刷

ひらがなの下に漢字を
書いて、読みやすくし
ている学習者が多い
です

にほんにきてから、さまざまあたらしいりょうりをたべてみました。

さいきん よくたべているりょうりといえば、なっとう芋モチです。

にほんとかんこくのだいひょうてきなこくめんしよく なっとうと芋モチを、はっばうスチロールのばこにいれて、まぜるだてでかんせいする

かんたんなひょうじなです。

ちょうみりょうのほうは、芋モチのとうがらしだてでも

あるいは、通常
画面をSnipping
Toolで切り取り、
Wordに貼り付け

6. 読む練習

- ・ 画面または、印刷した紙を見ながら練習

個人練習→グループ練習

チェックポイントは、

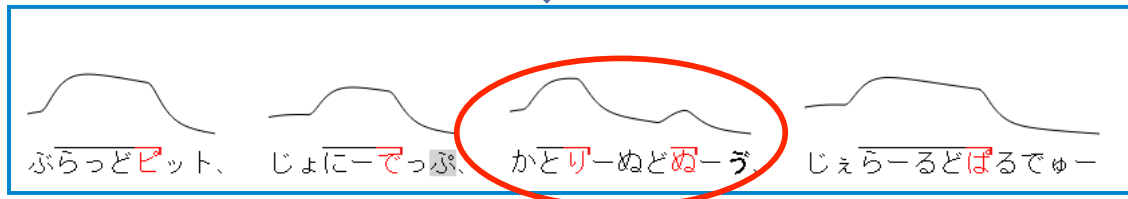
1. ポーズは十分か
2. への字になっているか

7. 発表

スズキクンの弱点3 名前のアクセント

- カタカナの名前のアクセントをときどき間違える

ブラッド・ピット、ジョニー・デップ、カトリーヌ・ドヌーヴ、ジェラルド・ドパルデュー



- アクセントが違う場合は 教師が対応
- 外国人名は名字と名前が1つのまとまり

スズキクンの弱点3 名前のアクセント

どうやって、アクセントをつけているか

➡ 未知語が出てきたら、「最後の音節を取外し、後ろから数えて…」の規則を適用

- 私は、ウンジャラペンと言います。

わたしは、ウンジャラペンと言います。

- 私は、ナンジャラペンと言います。
何+じゃら+ペン(?)

わたしは、なんジャラペンと言います。

形態素解析がどの単語をどのように解析したのかは、ユーザには表示していません。

日本人の名前

峯松 信明


 みねまつのぶあき

峯松／信明


 みねまつ
 
 のぶあき

日本人の場合 名字と名前が別

スズキクンの弱点4 その他アクセントの誤り


 もうだいた？


 もういちど おわがいます


 だいいちに


 だいさんもくようび


 どうきょうだいがく

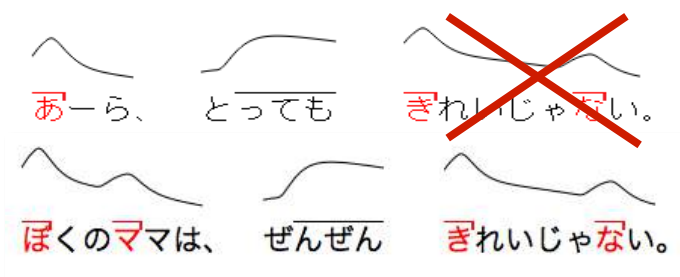

 どうきょうだいがくこうがくぶ

スズキクンの弱点5 読みの誤り

- 文脈に依存して読みが変わる語は、やっぱり読めないことも
 何が見えますか？ → **なん**が見えますか？
 生物はこの袋に捨てて下さい。 → **せいぶつ**はこのふくろに・・・
 木の下で、雨宿りをした。 → **きのもと**で・・・
- 読み推定は、全て形態素解析器に任せています。
- 解決**→**平仮名**での入力を試してみてください。
- 数字＋助数詞表現も同様、読み／アクセントを間違いやすい
 - 数字そのものはかなり読めます。

スズキクンの弱点6 文脈依存性

- 意味の曖昧性(文脈依存性)とイントネーション制御
- 「あーら、とっても／きれいじゃない」
- 「僕のママは、全然／きれいじゃない」



イントネーション
によって逆の
意味に

- スズキクンは文脈に応じて意味を理解し、アクセント／イントネーションを付けている訳ではない。
- 原理的に、処理不可能な問題。**解決**→ **手で修正**

入門クラス スピーチ例 超初級者用

はじめ^まして。

初級の場合、切れ目は教師が入れた方がいい

い ^みんちえです。

が^んこくから き^ました。

初級の場合、切れ目とへの字で十分

ど^うぞ よろしく。

メニュー

- ・ 1. はじめに
- ・ 2. OJADを使った授業案
 - ①スピーチ
 - ②『初級日本語』 会話
- ・ 3. 終わりに

OJADの使い方② 会話『初級日本語』9課

タン: もしもし、ゴーさんですか。タンです。こんばんは。

ゴー: ああ、タンさん、こんばんは。おげんきですか。

タン: はい、げんきです。ほっかいどう大学のじゅぎょうは／どう
ですか。せんもんのじゅぎょうは／むずかしいですか。

ゴー: いいえ、むずかしくないです。とてもおもしろいです。

タン: 大学では／日本ごや／えいごをべんきょうしますか。

ゴー: 日本ごはべんきょうしますが、えいごはべんきょうしません。

タン: ちゅうごくごもべんきょうしませんか。

ゴー: はい、ちゅうごくごもべんきょうしません。

スズキクンに入力

タン、もしもし、ゴーさんですか。

タンです。

こんばんは。

ゴー、ああ、タンさん、こんばんは。

おげんきですか。

疑問イントネーション

疑問イントネーション

- ももし、ゴーさんですか？
- おげんきですか？




 ダン、 ももし、 ゴーさんですか？


 おげんきですか？


 ダン、 ももし、 ゴーさんですか？


 ダンです。 こんばんは。


 ゴー、 ああ、 たんさん、 こんばんは。 おげんきですか？

網掛けは無声化音


 ダン、 はい、 げんきです。 ほっかいどうだいがくのじゅぎょうは どうですか？


 せんもんのじゅぎょうは むずかしいですか？


 ゴー、 いいえ、 むずかしくないです。

メニュー

- ・ 1. はじめに
- ・ 2. OJADを使った授業案
 - ①スピーチ
 - ②『初級日本語』 会話
- ・ 3. 終わりに

OJAD 便利なこと

スズキクン:

- ・スピーチや会話の練習が一人でもできる(音声も聞ける)

単語検索:

- ・動詞、形容詞の活用アクセントが調べられる(音声も聞ける)
- ・クラスで音声を聞きながら活用アクセントが練習できる
- ・教科書の課を指定して、新出語彙のアクセントが調べられる
 - ┌ 授業前に教師が確認
 - └ 学習者が予習

成果 学習者の発音 (2013年前期のクラスから)

- ・ 1. 中上級の学習者 中国
- ・ 2. " 香港
- ・ 3. " フランス
- ・ 4. " ロシア
- ・ 5. " 韓国
- ・ 6. 上級の学習者 ニュージーランド
- ・ 7. " 中国
- ・ 8. " フィンランド



最後に一言

- ・ OJADは、教師と学習者が**ともに学ぶ**ためのツール
- ・ 教師がすべて知っている必要はない
- ・ 一緒にOJADで**遊び**ながら、新たな発見をしていただきたい

ありがとうございました。

本日のメニュー

最近の PC は上手に喋るけど、誰が日本語教えてる？

- 韻律に着目。単語アクセント・アクセント結合・アクセント句

日本人にとってのアクセントとは？

- 日本人はアクセントの間違いに敏感なの？鈍感なの？

日本人にとって分かり易い発声とはどういうものか？

- フレージング（チャンキング）とポージング

OJAD 作りました。ご紹介します。

- 単語検索 / 動詞の後続語検索 / 任意テキスト版

- 韻律音読チュータ・スズキくん



OJAD 使って教えています。指導のヒントをご紹介します。

- フレージング&ポージングによる発声指導&プレゼン指導

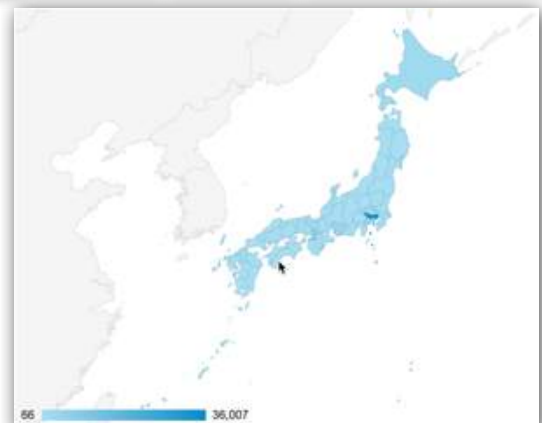
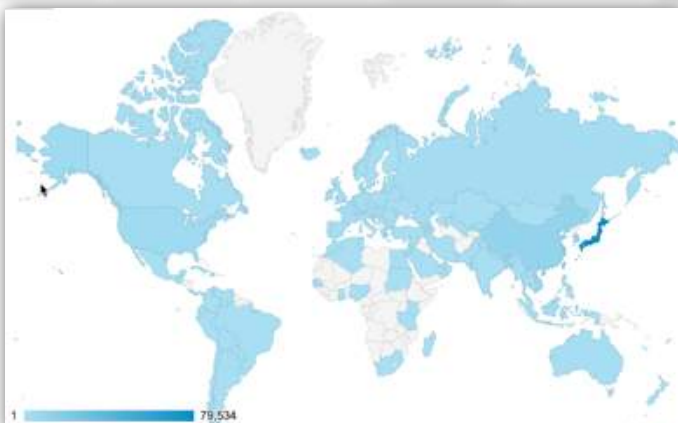
まとめ

OJADの利用状況

2012年8月に一般公開

- それ以来、Google Analytics を使ってアクセス状況を調査。

- 合計約 15.3 万のアクセス。約 5 割は海外からのアクセス。

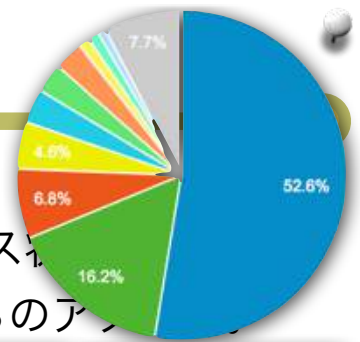


(Google Analytics が提供するアクセスマップ)

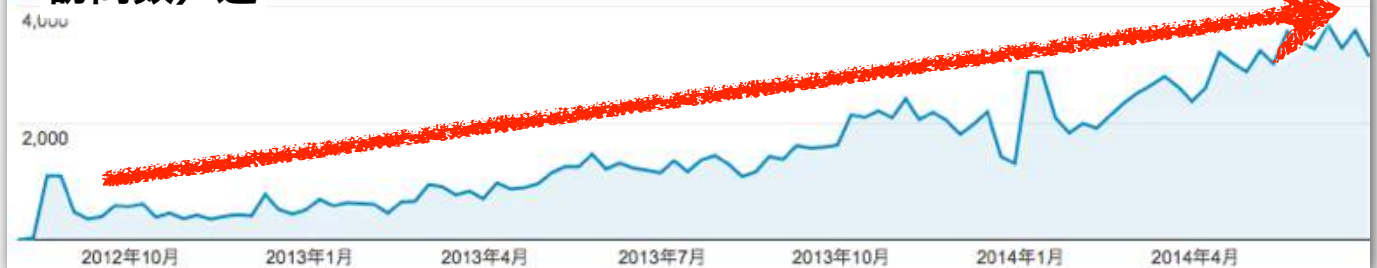
OJADの利用状況

2012年8月に一般公開

- それ以来, Google Analytics を使ってアクセス
- 合計約 15.3 万のアクセス。約 5 割は海外からのアクセス



訪問数/週



(Google Analytics が提供するアクセスマップ)

OJAD on Facebook

是非「いいね」を宜しくお願いします。

OJAD - オンライン日本語アクセント辞書

www.gavo.t.u-tokyo.ac.jp/ojad/

OJAD

ツイート 99 いいね! 1,477

ご利用にあたって

OJAD 4機能のご紹介

使ってみようOJAD

1-min PV JPN / ENG

f OJAD on Facebook

OJAD 4機能

単語検索

動詞の後続語検索

任意テキスト版

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9,000の名詞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、な形容詞）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1名ずつの音声サンプルを聴取可）。それ以外の約300の後続語表現に関しても、アクセントを調べることができます。更に任意の肯定文・疑問文に対して、特定の語に強いフォーカスを置かず読み上げた際に予想されるピッチパターンを表示することもできます。学習者の自主学習、日本語教室や日本語教師養成講座での韻律学習にお使い下さい。日本語母語話者の東京方言参照用にもお使いいただけます。

利用される場合は、まず「ご利用にあたって」をご覧ください。

更新履歴

20140709 スズキクンに喋る機能を付与しました。スズキクンの挨拶は[こちら](#)。喋るスズキクンを詳しく知りたい方は[こちら](#)。

20140704 [アクセス履歴](#)に2014年6月のアクセス履歴を追加しました。教科書「まるごと日本のことばと文化 入門A1 かつどう(三修社)」を追加しました。

OJAD on Facebook

是非「いいね」を宜しくお願いします。



OJADの今後

多言語化



● ポーランド/タイ語版は、夏休み明けには公開か

新教科書の更なる追加

● トップページ「教科書の追加」をご覧ください。

アクセント制御に関する基礎知識の提供

● アクセント結合やアクセントの揺れに関する簡単な解説

OJAD 講習会のビデオ化とオンライン視聴

● 可能なら、字幕をつけたいところだが

スズキクンの喋りを鍛える

● 初級者モードでの読み上げのサポート & 話速の制御

● 複数の声色のサポート（複数人の会話を音声で呈示）

● フレーズ区切りの自動挿入機能

● 助数詞、固有名詞など、細かなバグフィックス



今後とも、OJADを宜しくお願い致します。
ご意見などは ojad@gavo.t.u-tokyo.ac.jp まで
(開発チーム一同)



今後とも、OJADを宜しくお願い致します。
ご意見などは ojad@gavo.t.u-tokyo.ac.jp まで
(開発チーム一同)



ご清聴，有り難うございました



日本語（非国際語）と英語（国際語）

日本語（非国際語）の音声学習

- 日本語を話す場合，**相手は殆ど日本人**
- 日本人にとって聞きやすい日本語を学ぶ必要がある

英語（国際語）の音声学習

- 英語を話す場合，**相手が母語話者であることは稀**
- 非母語話者同士の共通語（国際語）として存在する英語
 - 公用語としての英語の存在（インド，フィリピン，ナイジェリア・・・）
 - 多様な英語を認めざるを得ない状況（＝**世界諸英語，WEs**）

多様な英語発音を前提とした学習支援策の一つ

- 自分の英語は世界諸英語の中でどのように位置づけられるのか
- 自分と似た／異なる発音の話者は世界中のどこに存在するのか



- 
- **英語話者（＝15億人）を話者を単位として分類，地図化する！**

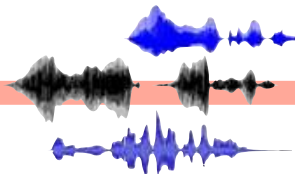
Speech Accent Archive

特定パラグラフを読ませた世界諸英語音声コーパス

- 米語の音素（対）の coverage を考えたパラグラフ設計
- 発話者参加型のコーパス構築（現在約1,900名）
- IPA narrow transcription の提供（約150種類のIPA記号）
- <http://accent.gmu.edu>



Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.



tolis kool stela ask ha tu bün	
tolis kool stela ask ha tu bün	tol
[plis kool stela ask ha tu bün]	tol
di: ʔiŋks wəʃ ha fəʃm ʔu stə	ks
siks sʔiŋs of fəʃ s no pi:ʔ	u
fəʃ ʔi:ʔ k:ʔ əʃləps əv blə	ʔi
ʔi:ʔ ʔən meibi e snæk fə ʔu	
bəʔde bəp wi ʔoləs nɪ ʔe	ʔi
sməl pləstik snæk ʔi:ʔ e bɪk	n
təʃ fəʃm de kɪs ʔi kəʃn	
skəp ʔəʃ ʔiŋs tɪnu ʔri: ʔe:	ʔi
bəks ʔən wi wɪl go tɪn e	
wɛnzde ʔəʃ ʔe mɪn stəʃ	ʔi

Best Paper Award

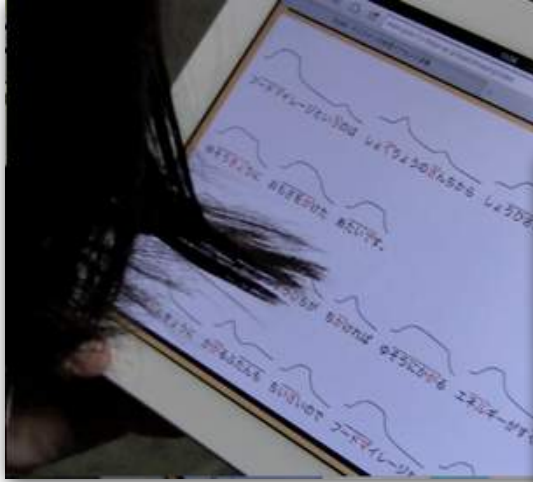
IEEE Int. Conf. Information Science and Technology2014



OJAD と WE, 実は繋がってます・・・

OJADの開発

- 任意の日本語テキストに対して、アクセント・イントネーションの情報を呈示する, **世界で最初&唯一の**教育インフラ



Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.



ご清聴, 有り難うございました

