

機関番号：13901

研究種目：若手研究(B)

研究期間：2009～2010

課題番号：21720205

研究課題名(和文)

プロソディ情報の利用を評価するためのオンライン型英語音声認識テストの開発

研究課題名(英文)

Development of Web-Based Speech Recognition Test for Assessing the Use of Prosodic Cues

研究代表者：

村尾 玲美 (MURAO REMI)

名古屋大学・国際言語文化研究科・准教授

研究者番号：80454122

研究成果の概要(和文)：

学習者が英語音声聞き取るにあたり、プロソディ情報を利用できているかどうかを自己診断できるような音声認識テストを開発するため、まず英語母語話者に対して「プロソディ情報を利用した文構造認識テスト」を課し、正解率が100%に近い項目を選定して、英語学習者用のテスト項目として使用することにした。テストのオンライン化はHot Potatoes Ver.6およびSakai Projectによる学習支援システム(NUCT)を利用して行った。大学生20名にオンライン型英語音声認識テストを受験させた結果、弁別力の最も高い項目はLate/Early Closure文であることが明らかになった。この項目は英語母語話者と上位27%の学習者の正解率は100%であるのに対し、下位27%の学習者の正解率は0%であった。

研究成果の概要(英文)：

In this study, the ability of native speakers of English (NS) to use prosodic cues to recognize sentence structure was first examined. The purpose of examining the performance of NS was to establish the standard of what NS can do with prosodic cues and to determine the validity of test items, which will be used to develop a Web-Based Speech Recognition Test for language learners. Hot Potatoes Ver.6 and Course Management System by Sakai Project were used to develop the online test. The result of 20 Japanese College Students who took the online test showed that Late/Early closure type of sentence pairs had the highest discriminative power; All NS and upper 27% of language learners marked 100% for these items, whereas lower 27% of language learners marked 0%.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	600,000	180,000	780,000
2010年度	900,000	270,000	1,170,000
総計	1,500,000	450,000	1,950,000

研究分野：第二言語習得論

科研費の分科・細目：言語学・外国語教育

キーワード：音声言語認識テスト、プロソディ、英語教育

1. 研究開始当初の背景

村尾(2009)は、英語の聞き取りが得意な学習者と不得意な学習者が、音素情報とプロソディ情報をどのように利用しているかを調べるため、ゲーティング法を応用した音声言

語認識実験を行った。実験の結果、英語の聞き取りが得意な学習者は、心的辞書に個々の単語の音素情報のみならず、高頻度表現の音韻情報が余剰的に入っているため、入力音声との照合を行う際、より多くの情報を利用

きることが明らかになった。

これに対し、英語の聞き取りが不得意な学習者の心的辞書には、単語が個別に入っており、音声インプットとの照合ができるのは、ほんの一部の音素だけであることが示唆された。つまり、入力音声は豊富な情報を含んでいるのに、学習者が手がかりとして利用できる情報は僅かだということである。

また、Pennington & Ellis (2000)の実験では、日本語はピッチアクセント言語であるため、文レベルのプロソディの特徴や、統語構造との関連を認識することは、日本人にとって困難であることが指摘されている。一方、英語母語話者は、プロソディ情報を利用することにより、文の統語構造を推測することができるという可能性が示唆されている(村尾, 2009)。

これまで国内外で行われてきた研究により、プロソディ情報を音声認識の手がかりとして授受することの重要性が明らかになりつつある。しかしながら、プロソディの認識を評価するテストとしては、現状では単語のアクセント位置に丸を打たせる形式しか存在しない。このテスト形式では、単語のアクセントがどこに置かれるかという「知識」を測っており、実際の音声を聞かせずに紙面上で回答させている。入力音声と心的辞書内の語彙項目とを照合する際に、与えられた音韻的手がかりを利用して聞き取ることができるかという「運用」を評価するようなテストは行われていない。そこで本研究では、これまでの実験手法と研究成果を生かし、プロソディ情報の利用に焦点を当てたテストを開発する。

2. 研究の目的

本研究の目的は、単語のアクセント位置などの知識を測るテストではなく、実際にプロソディ情報を利用して音声言語を認識できるかという言語運用を測るテストを開発することである。村尾(2009)の研究で利用された高頻度表現の認識を評価するテスト項目に加え、先行研究で示唆された文構造の認識についても評価できるようなテスト項目を作成する。具体的な研究目的は以下の二点である。

- (1) プロソディ情報を利用した文構造の認識判断について英語母語話者に対して実験を行い、テスト項目を精査する。
- (2) すでに用意してある高頻度表現のテスト項目と、(1)により精査された文構造の認識にまつわるテスト項目をまとめ、音声言語認識テストとしてウェブ上で受験できるようにする。

3. 研究の方法

3. 1. プロソディ情報を利用した文構造の認識判断

3. 1. 1. テスト項目

先行研究から、英語母語話者は文のプロソディ情報を利用することにより、文構造を推測できるという可能性が示唆されているものの、直接は実証されていない。そこで、まず、英語母語話者はプロソディ情報を特定の文構造と結びつけて認識することができるのかどうかを調査した。実験項目には、Pennington & Ellis (2000)で用いられているような、文構造の違いによってアクセント位置が変わる文ペアを用いた。アクセント位置以外が完全に同一である文ペアを複数作成するのは困難であるため、先行研究を参考に、発音の似た単語を使った文ペアや、同じ単語で語順の異なる文ペアを5種類作成した(Adams, 2007; Hwang & Schafer, 2006; Soderstrom et al., 2003)。

①Early Closure 文と Late Closure 文のペア

- 1a. Since I always DRINK, wine is gone.
- 1b. Since I always drink WINE, it's gone.

②SVOC 文 (結果構文) と SVO 文のペア

- 2a. He pushed the door CLOSED.
- 2b. He pushed the CLOSED door.

③句動詞のペア

- 3a. Put on your SHOES.
- 3b. Put your SHOES on.

④複合語と名詞句のペア

- 4a. He's at the GREENhouse.
- 4b. He's at the green HOUSE.

⑤内容語 (強音節語) と機能語 (弱音節語) の同音異義語ペア

- 5a. Cats jump WOOD Benches.
- 5b. Cats would jump BENCHES.

3. 1. 2. 音声データの収集と加工

音声データを収集するため、実験項目 46 文を英語母語話者 6 名に音読させ、録音した。国籍はイギリス人 4 名とアメリカ人 2 名である。音読は 2 回行われた。1 回目は、46 文をペアで提示するのではなく、ランダムに提示し、自由に音読させた。2 回目は、46 文を 5 種類のカテゴリーごとにペアで提示し、ペア文の違いを意識するように読ませた。6 名の英語母語話者から収集した音声データを聞き比べ、文ペアの違いを最も顕著に表していたアメリカ人話者の音声を本実験で使う音声として採用した。英語母語話者が文のプロソディを聞くことにより文構造を認識す

ることができるか否かを明らかにするため、音声全体をローパスフィルタでマスキングして、音素情報を排除した。

3. 1. 3. 実験参加者

本実験には、音声データを提供した6名の母語話者とは異なる英語母語話者が6名参加した。国籍はイギリス人1名、アメリカ人3名、オーストラリア人2名である。

3. 1. 4. 実験方法：文構造認識判断タスク

文構造認識判断タスクは、音素情報が排除され、プロソディ情報のみが利用できるよう加工された音声聞き、二つの選択肢のうちどちらの文構造であったかを判断するタスクである。文構造を正しく判断するには、アクセント位置やポーズ位置の違いといったプロソディの手がかりを利用する必要がある。実験は被験者一人ずつに対し個別に行った。被験者には、23個の文ペア（合計46文）が書かれた紙が用意された。文ペアは、カテゴリがわからないようにランダムに並べられている。被験者はローパスフィルタのかかった音声を23文聞き、ペアのうちどちらの文であったかを判断して、解答用紙のaかbに丸をつけた。被験者6名は、3名ずつ二つのグループに分けられ、一方のグループがペアのうちa文を聞いた場合、もう一方のグループはb文を聞くように配慮された。音声は解答に自信が持てるまで、何度でも聞き直してよく、聞いた回数（音声呈示回数）は記録された。

3. 1. 5. 分析方法

5種類のカテゴリおよび実験項目ごとに、正答率と音声呈示回数に違いがあるかどうかを分析するため、(1) 全体的な正答率、(2) カテゴリごとの正答率と平均音声呈示回数、(3) ペア文ごとの正答率と平均音声呈示回数、(4) 項目ごとの正答率と平均音声呈示回数をそれぞれ計算した。

3. 1. 6. 結果と考察

文構造認識判断タスクの全体的な正答率は93.48%と極めて高く、英語母語話者はプロソディ情報を利用して文の構造を認識できることを示した。カテゴリ別の正答率（図1）と音声呈示回数（図2）の結果から、Early/Late Closure文は最も早く正確に認識されることがわかった。これらの文の違いを認識するには、アクセント位置だけでなく、ポーズ位置も手がかりとして利用でき、他のカテゴリよりもプロソディ情報が豊富であったためだと考えられる。

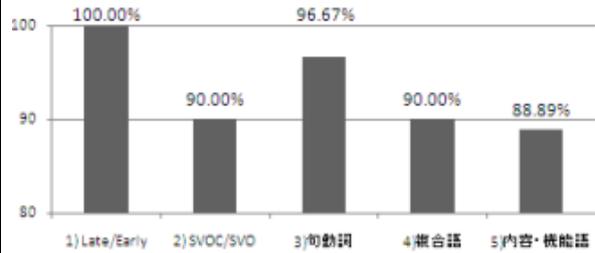


図1. カテゴリごとの正答率

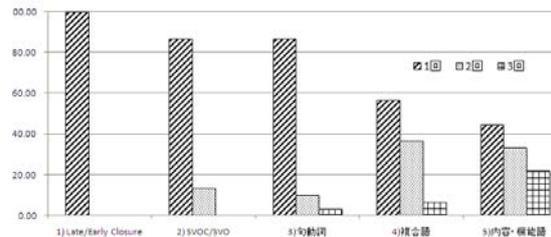


図2. カテゴリごとの音声呈示回数の割合

SVOC（結果構文）とSVOの文ペアは、正答率が90%だったものの、英語母語話者でも1割が正しく認識されないのであれば、第二言語学習者用に使う実験項目としては再検討する必要がある。1割が聞き間違えられた理由として、SVOC（結果構文）とSVOのペアには、韻律的に大きな違いがない可能性が考えられる。実験に用いた音声には韻律的な違いがあったとしても、被験者自身が自然な状態で発話する場合に韻律的な違いが無いのであれば、判断するのは難しいだろう。Noh (2005)は結果構文のアクセント位置について、本研究とは異なる見解を示している。Noh (2005)によれば、結果構文では必須項(argument)のみにアクセントが置かれ、本動詞に付加している結果述語(resultative predicate)にはアクセントが置かれない。例えば、he painted the door blueの場合、本動詞paintedと形容詞blueは共に述語を形成しており、the doorを必須項としているため、アクセントが置かれるのはthe doorのみとなる。Noh (2005)の指摘を確かめるため、今後の課題として、収集した6名の音声データを音響的に分析し、結果構文のアクセント位置を明らかにする必要がある。

複合語と名詞句のペアも正答率が90%であり、1割が正しく認識されなかった。しかし、正しく認識されなかった項目を見てみると、ひとは複合語の問題では無いことがわかる。“She painted her EYEBrows”と“She painted her eye BROWN”のペアは、前者の複合語の文は正しく認識されており、後者の文が聞き間違えられている。後者の文は先に

述べた結果構文であるため、Noh (2005)の指摘が正しければ、聞き間違える理由の説明がつく。“She’s a LIGHThouse keeper”と“She’s a light HOUSEkeeper”のペアは、2名が後者を聞き間違えている。このペアは、両方に複合語を含んでおり、形容詞と名詞という修飾関係があるため、他の項目よりも複雑で判断が難しかったと考えられる。複合語を含んだカテゴリとなっているが、品詞の並びが統一されていないため、再検討の必要がある。実験終了後の被験者へのインタビューにおいて、“She’s a light HOUSEkeeper”を認識できなかった被験者は、この文の意味的な曖昧さや不自然さを指摘していた。文法的には正しくても不自然さのある文は、音声化した時のイメージがわからず、そのためにプロソディを手がかりとした認識判断を難しくしている可能性も考えられる。

内容語と機能語の文ペアは、同音異義語を用いて実験文を作成する必要があったため、文法的に不自然さのある実験項目も含まれていた。正しく認識されなかった2文については、実験終了後のインタビューにおいて、被験者からその文法性が指摘された。“Let’s take on our CHALLENGE”と“Let’s take ONE HOUR CHALLENGE”のペアでは、前者が正しく認識され、後者が聞き間違えられている。後者の文は文法的には“Let’s take a one-hour challenge”と冠詞を入れる必要があるが、いずれにせよ表現として不自然であると指摘された。先に録音を行った6名の英語母語話者のうち5名も、同じように冠詞の「a」を入れるべきであると文法的に不自然さを指摘している。また、“I have bought sandwich”と“I bought half sandwich”のペアは両方も成績が悪かったが、6名の英語母語話者のうち6名ともが、前者は“I have bought a sandwich”と修正し、後者は“I bought half a sandwich”と修正している。冠詞などの機能語は、意味の伝達においてはあまり重要ではないかもしれないが、音韻的な観点からは、文のリズムを生成するという点で極めて重要である。母語話者へのインタビューで、文法的に不自然な文や、表現としてあまり聞かない文の場合、どのように音声化されるかというプロソディのイメージがわからないため、認識が困難になると指摘された。このことから、英語母語話者は文法的に正しく、頻繁に耳にする文構造であれば、頭の中にあるプロソディのイメージと照合して、インプット音声の文構造を認識判断していると考えられることができる。つまり、プロソディは計算によって処理されるばかりではなく、記憶によっても処理されることが示唆された。

母語話者の正答率と判断の早さを確認し、第二言語学習者向けに採用できるテスト項目を決定するという目的から実験の結果を

まとめると、以下の四点が明らかとなった。

- (1) Late/Early Closure 文と句動詞文は、母語話者による正答率の高さが確認できたため、第二言語学習者向けのテスト項目として利用できる。
- (2) 結果構文はアクセント位置が不確かであるため、再検討の必要がある。
- (3) 複合語のペアは、品詞の並びを考慮した上で再検討の必要がある。
- (4) 文法的正確さを欠いた文および表現として不自然な文は、音声的なイメージがわからないため、プロソディ情報を利用することができない。

4. 研究成果

4. 1. オンライン型英語音声認識テスト

4. 1. 1. Hot Potatoes Ver. 6 を利用したテスト

第3節で説明した実験の結果により、文構造の違いを聞き分けるテスト項目については最終的に38文を使用することとなった。高頻度表現を認識するテスト項目については、村尾(2009)に基づき、24文をテスト項目として利用することとした。テストのオンライン化についてはHot Potatoes Ver. 6のJQuiz形式を利用した。Multiple-choiceで答えを選択した後、不正解の項目については強弱の違いが視覚的にわかるようなフィードバックが与えられ、最終的な正解率とコメントが呈示される。図3はHot Potatoes Ver. 6によるテスト画面である。



図3. Hot Potatoes Ver. 6 を利用した「韻律の使用を評価するための音声言語認識テスト」

受験者は画面上の指示文（「音声聞いて、AとBどちらの語句を言っているか選びなさい。わからない時はCを選びなさい。」）を読んだあと、ボタンを押して音声聞き、AからCまでの選択肢を選ぶ。回答ボタンを押すと直後に正解・不正解の情報が与えられる。最後の問題まで回答すると、全部の問題に対するフィードバックが与えられる。フィードバックでは、ペア文の意味の違いが提示され、アクセントの位置が大文字で示される。

4. 1. 2. 学習管理システム (NUCT) を利用したテスト

Hot Potatoes Ver. 6 で作成したのと同様のテストを、名古屋大学で使用している学習管理システム (NUCT: Nagoya University Collaboration and Course Tools) を利用して作成した。このシステムは Sakai Project により提供されているソフトウェアにより作成されている。図 4 は NUCT のテスト画面である。

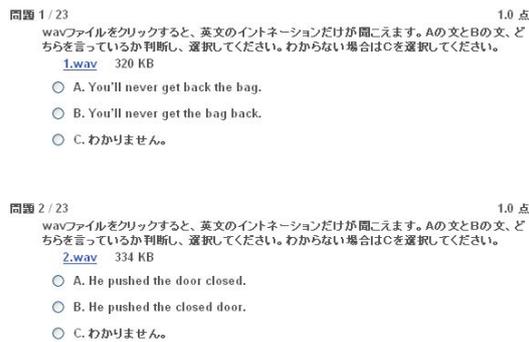


図 4. NUCT を利用したテスト

このシステムではテストの統計情報を得ることができる。また、項目分析により各項目の弁別力を知ることもできる。実際に文構造認識判断テストを 20 名の学生に受験させたところ、正解率 65% から 74% の間に位置する学生が最も多いことがわかった。最高点は 91 点であった。

項目分析の結果、最も弁別力の高い問題 (弁別力 1.0) は “Whenever the guard checks the DOOR, it’s locked.” と “Whenever the guard CHECKS, the door is locked.” の文ペアおよび、 “Because her mother KNITTED, sweaters kept Mary warm.” と “Because her mother knitted SWEATERS, Mary kept warm.” の文ペアであることが明らかになった。この問題は上位 27% の学生は 100% 正解しているのに対し、下位 27% の学生の正解率は 0% である。更に、二番目に弁別力の高い問題 (弁別力 0.8) も “When John leaves CATHY, we’ll be upset.” と “When John LEAVES, Cathy will be upset.” の文ペアであることがわかった。この問題は上位 27% の学生は 100% 正解しているのに対し、下位 27% の学生の正解率は 20% である。つまり、最も弁別力の高い上位三つの問題が、すべて同じカテゴリーの問題であった。Late/Early Closure 文のペアは、英語母語話者の正答率は 100% で、認識までの音声

呈示回数も 1 回であったが、日本人英語学習者の場合、平均正解率は 65% から 94% まで幅があり、平均は 78% であった。今回の調査により、聞き取りが得意な学習者と苦手な学習者で差のでる項目であることが明らかになった。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計 2 件)

- ① 村尾玲美、「韻律の手がかりによる文構造の認識判断—英語母語話者を対象とした調査—」ククロス：国際コミュニケーション論集、査読有、第 8 号、2011、57-72.
- ② Remi Murao & Michiko Nakano, Towards the assessment of the use of prosodic cues in speech recognition, Proceedings of the 14th Conference of Pan-Pacific Association of Applied Linguistics, 2009, 260-262. 査読有

〔学会発表〕 (計 1 件)

- ① 村尾玲美、聞き間違い現象から探る音声言語認識メカニズム、第 35 回全国英語教育学会鳥取大会、2009 年 8 月 9 日、於鳥取大学。

〔図書〕 (計 1 件)

- ① 村尾玲美、「国際言語としての英語における韻律の役割」安武知子・小泉直・川岸貴子・今井隆夫・内田政一 (編) 『ことばとコミュニケーションのフォーラム』、開拓社、2011、197-208.

6. 研究組織

(1) 研究代表者

村尾玲美 (MURAO REMI)

名古屋大学・大学院国際言語文化研究科・准教授

研究者番号：80454122

(2) 研究分担者

なし

(3) 連携研究者

なし