

## 英語発話学習に向けた音素連結・脱落・同化パターンの分析と検出

## Analysis and Detection of Liaison, Elision and Assimilation in Patterns for English Utterance Learning

平田 里佳<sup>\*1</sup>, 入部 百合絵<sup>\*1</sup>  
Satoka HIRATA<sup>\*1</sup>, Yurie IRIBE<sup>\*1</sup>  
<sup>\*1</sup> 愛知県立大学

<sup>\*1</sup>Aichi Prefectural University  
Email: is141064@cis.aichi-pu.ac.jp

あらまし: 従来の発音教材では, 単語と単語をつなげて発音するときに生じるリンキングを重要視していない。しかし, リンキングは英語らしい発音や英語の聞き取りにおいて重要である。そこで本研究では, 日本語母語話者による英語発話の音声認識精度の向上に貢献するため, 英語の発音のうち特にリンキングに着目し, リンキングに対する英語母語話者および日本語母語話者のパターンを分析し, リンキング箇所に対する分析と音声認識を試みた。

キーワード: 音声認識, 英語発音, HMM, 音声分析, リンキング

## 1. はじめに

英語発音学習のための教材として, 自学自習を目的とした CALL (Computer Assisted Language Learning) など様々なものが存在する。しかし, これらの教材では孤立単語の発音学習が中心であるため, 文章のような複数の単語を発音する際に生じるリンキング (単語連結時に生じる音素の連結, 脱落, 同化) の発音学習には対応していないことが多い。リンキングとは, 例えば *did you* のように二つの単語を発音する場合に, *did* の最後の音素 /d/ と *you* の最初の音素 /j/ が連結することで異なる音韻となるような現象を指す。リンキングは英語らしい発音をするためには必要不可欠であるが, アクセント型発音の日本語を母語としている日本人にとっては発音が難しく訓練が必要である。リンキングはその発音が独特であるため, リンキングを理解していないと聞き取りにおいても支障が生じる可能性がある。

従って, 本研究では英語母語話者および日本語母語話者のリンキングパターンを分析し, リンキング箇所に対する音声認識を試みる。これにより, 誤ったリンキングの発音を適切に指摘できるようにする。

## 2. リンキングパターンと日本語母語話者の発音誤りパターン

## 2.1 英語母語話者のリンキングパターン

リンキングは次の3つに分類されている<sup>(1)(2)</sup>。

- 連結 (Liaison)  
単語の切れ目に変化する (例: *put on* [put an] → [putan])
- 脱落 (Elision)  
破裂の省略や子音の脱落が起こる (例: *could be* [kəd bi:] → [kədbi:])
- 同化 (Assimilation)  
音素の変化や音の挿入が起こる (例: *did you* [did ju:] → [didʒu], *may I* [mei ai] → [meiʰai])

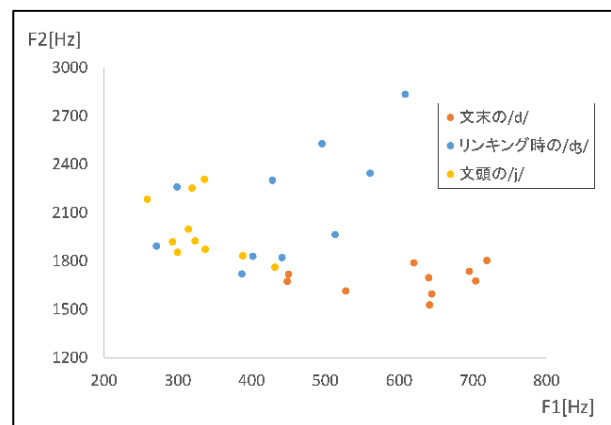


図1 リンキングによるフォルマントの変化 (/d+/j/ → /dʒ/)

本研究で調査したリンキングパターンは, 連結・脱落・同化それぞれに対して 20, 32, 24 種類である。

また, 同化の “/d+/j/ → /dʒ/” を例として, 文末の /d/ および文頭の /j/ のフォルマントと, /d/ と /j/ がリンキングにより同化したときの /dʒ/ のフォルマントを比較した。結果を図1に示す。図1より, リンキングが生じることによって /d/ と /j/ が本来のフォルマントの位置からずれていることが確認できる。

## 2.2 日本語母語話者の発音誤りパターン

日本語母語話者による主な英語の発音誤りは, 置換誤りと母音挿入誤りの大きく分けて 2 種類存在する<sup>(3)</sup>。

- 置換誤り  
発音が類似した音素同士を置き換える  
(例: /ɑ/, /ʌ/, /ə/ と日本語「ア」の混同, /l/ と /r/, /θ/ と /s/ の混同)
- 母音挿入誤り  
子音の後ろに不要な母音を挿入する  
(例: *it* [it] → [itoJ\*]) \*日本語の「オ」

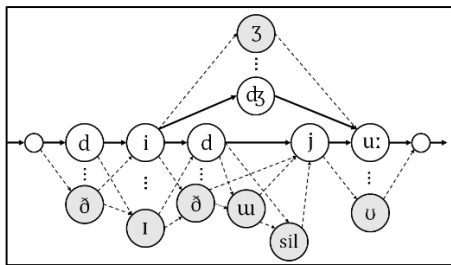


図2 誤り候補を含んだネットワーク文法の例  
「... did you ...」

表1 ネットワーク文法の有無に対する  
4best 認識結果

ネットワーク文法	正解率(%)	正解精度(%)
なし	31.3	28.6
あり	63.3	63.0

表2 連結・脱落・同化の検出精度

	連結	脱落	同化
正解率(%)	80.0	80.4	50.0

### 3. リンキングパターンを考慮したネットワーク文法の作成

本研究で構築した音声認識システムでは、入力音声から抽出した MFCC (Mel-Frequency Cepstrum Coefficients), ΔMFCC, ΔΔMFCC, パワー, Δパワー, ΔΔパワーの計 39 次元の特徴量を用いて学習した音響モデルを作成する。このとき、HMM (Hidden Markov Model) から得られる音素列に対して、音響モデルと言語モデルを用いて尤もらしい音素を選択する。一般的な発音学習における音声認識器では、音響モデルに英語母語話者による英語音声のみを学習させるが、母語である日本語に影響された英語発音を適切に認識できない可能性がある。

本研究では、音響モデルに英語母語話者による英語音声 (TIMIT) に加え、日本語母語話者による日本語音声 (JNAS) も学習させることで、日本語特有の音素にも対応できるようにした。

本研究で明らかにしたリンキングパターンを精度よく検出できるようにするために、言語モデルにはネットワーク文法を用いた。ネットワーク文法とは、図 2 のように正解音素(白色)と誤り音素(灰色)をネットワーク状につないだラティスのことである。

### 4. 評価実験

ネットワーク文法によるリンキングパターンの検出精度を確かめるため、ネットワーク文法の有無による比較実験を行った。また、ネットワーク文法を用いた場合の連結・脱落・同化に対する各々の認識精度を確認した。

#### 4.1 構築した音響モデル

使用した音声試料は以下の通りである。

- 学習データ  
TIMIT: 男性 438 名, 4,380 文  
JNAS: 男性 145 名, 7,267 文
- 評価データ  
日本人による英文読み上げ音声 UME-ERJ: 男性 48 名 90 単語列

本研究で構築した音響モデルは 5 状態 3 ループの left-to-right 型で、16 混合 monophone とした。

#### 4.2 ネットワーク文法の有無に対する認識性能の比較

ネットワーク文法の有無に対する認識結果を表 1

に示す。認識結果から、ネットワーク文法を使用することで正解率および正解精度が 32% 以上向上した。特にネットワーク文法ありの場合では正解率と正解精度の差が小さくなっていることから、認識時の挿入誤りを削減できたといえる。

しかし、ネットワーク文法を使用した場合でも正解率、正解精度共に 63% ほどであり、十分に認識ができていない。認識誤りについて調査すると、破裂音の前の破裂部分を表す cl, vcl の脱落誤りが非常に多かった。これを改善する方法としては、無音区間を音韻環境によって細分化することが挙げられる。

#### 4.3 連結・脱落・同化に対する認識性能

ネットワーク文法ありの場合において、正解率および正解精度が低い理由を明らかにするため、連結・脱落・同化に対する検出精度を算出した (表 2)。表 2 から、連結と脱落に関しては比較的高い精度で検出できているが、同化に関しては精度が低いことがわかる。同化は単語連結時に音変化を生じさせるリンキングパターンであるが、構築した音響モデルに含まれる音素モデルとは異なる音韻に変化している可能性もあり、認識を困難にしている要因と考えられる。今後はリンキングの中でも特に同化に対する認識改善が必要である。

### 5. おわりに

本研究では、従来の発音学習システムでは重要視されていないリンキング (単語連結時に生じる音素の連結・脱落・同化) に対する英語母語話者および日本語母語話者のパターンを分析し、リンキング箇所に対する音声認識を試みた。言語モデルとして誤り候補を含んだネットワーク文法を組み込むことで、正解率および正解精度が 32% 以上向上した。今後は、同化の認識手法を改善していく予定である。

#### 参考文献

- (1) 巽一朗: “英語の発音がよくなる本”, KADOKAWA, 東京 (2005)
- (2) 深澤俊昭: “英語の発音パーフェクト学習辞典”, アルク, 東京(2000)
- (3) 坪田康, 壇辻正剛, 河原達也: “日本人の誤りパターンの対判別を利用した英語発音教示システム”, 信学技報, SP2000-125 (2001)