

# 日本語母語話者による英語音節/ra/・/la/の知覚における音節前への母音付与についての考察\*

◎渡丸嘉菜子（上智大・理工），中村太一（福井大・教育地域），  
荒井隆行（上智大・理工）

## 1 はじめに

### 1.1 背景

外国語音声の知覚については長年研究がされている<sup>[1-5]</sup>。特に，外国語を知覚する際，母語の音声カテゴリーに当てはめて(誤)知覚されることは度々指摘されている<sup>[2-4]</sup>。例えば，英語の/r/と/l/は日本語母語話者には日本語の/r/として知覚される<sup>[1-3]</sup>。一方で，外国語音声の配列が，母語の音素配列規則(phonotactics)に合致しないような場合には，母語の音素配列規則に従うよう，(無意識に)知覚が影響を受けることが指摘されている<sup>[4]</sup>。例えば，日本語では特殊モーラの場合(つまり/N/や/Q/)を除いて，子音の連続(consonant cluster)が許容されていない。そのため，子音連続を聴いた際，日本語母語話者は存在しない母音/u/を知覚的に挿入してしまう<sup>[5]</sup>。具体的には，[ebzo]という音声聴いた際，実際には存在しない母音/u/を知覚的に挿入し，あたかも[ebuzo]であるかのように知覚することが報告されている。

日本語母語話者が英語の音声聴いた場合，上記2つの現象は度々共起する。ところが，/r/と/l/の知覚の際，通常の知覚的な母音挿入と様相が異なる現象が観察されている。それは，日本語母語話者が英語の/r/もしくは/l/を含む音節/CV/を知覚した際，実際には存在しないはずの母音/u/を音節前に付与し，/uCV/と知覚するという現象である。従来，知覚レベルでの母音挿入とは子音連続を防ぐ目的で行われる操作であるため，そもそも子音連続のないCV/という構造を持った音節の前に母音を付与するという現象は，従来の理論で保証されるものではない。にもかかわらず，英語の/r/や/l/の知覚の際，/u/を付与して(つまり/ura/や/ula/として)知覚するという観察は度々報告されており，その観察が正しいことを前提とした議論も度々報告さ

れている<sup>[4]</sup>。

### 1.2 研究目的

しかしながら，我々が知る限りにおいて，上記のような報告のすべてが観察的報告にとどまっており，音声的には存在しないはずの母音を音節の前に付与するという現象を実証した報告はない。そこで本研究では，英語の/r/もしくは/l/を含む音節の前に母音/u/を付与する現象の事実関係を確認することを目的に実験を行った。

## 2 実験

### 2.1 ねらい

上述の母音付与に関する観察は，当該子音部の調音の持続時間の長さに基づいた説明が可能かもしれない。/r/と/l/は口腔内で完全なる閉鎖が起こっていないため，それぞれの子音を持続的に発話することが可能である。さらに，それらの子音は聞こえ度(sonority)が高い。従って，/r/もしくは/l/の前に/u/を付与するという知覚は，子音部の調音の持続時間の長さを「母音」として誤知覚している可能性が考えられる。本研究では，そのような可能性を検証するため，/r/と/l/を含む単音節(/Ca/)とそれぞれの子音(/C/)の遷移開始前の冒頭部を引き延ばした刺激音を人工的に作成した。もし，知覚的に母音を付与する事実があるならば，聴取者は子音部の長さに関わらず子音の前に母音があると判断するはずである。本研究では，知覚的な母音付与と子音冒頭部の引き延ばしとの関係を調査するため，子音の遷移開始前の冒頭部分の長さを，50 ms, 100 ms, 150 ms, 200 ms の4段階に設定した。

### 2.2 刺激音

刺激音は Klatt のフォルマント合成器<sup>[6, 7]</sup>を用いて行った。作成方法は次の通りである。まず，先行研究<sup>[8]</sup>を参考に男性英語話者の母音/a/([a])

\* Perceptual vowel insertion to the pre-syllable positions of English /ra/ and /la/ syllables by native speakers of Japanese, by TOMARU, Kanako (Sophia University), NAKAMURA, Taichi (Univ. of Fukui) and ARAI, Takayuki (Sophia University).

表 1. 後続母音/a/定常部のフォルマント値 (Hz)

<b>F1</b>	730
<b>F2</b>	1090
<b>F3</b>	2440

表 2. 遷移開始点のフォルマント値(Hz)

/r/	<b>F1</b>	410
	<b>F2</b>	970
	<b>F3</b>	1408
/l/	<b>F1</b>	410
	<b>F2</b>	1098
	<b>F3</b>	2474

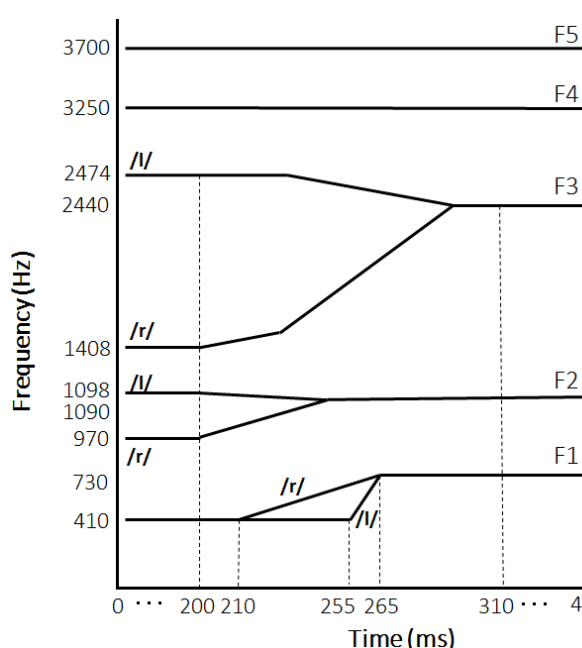


図 1. /r/および/l/のフォルマント遷移

の第 1, 2, 3 フォルマント(F1, F2, F3)を刺激音の後続母音定常部として採用した(表 1)。その母音定常部を到達点としてフォルマント遷移を付与する<sup>9)</sup>ことで音節/ra/および/la/を作成した。遷移開始点のフォルマント値は表 2, 遷移の形状は図 1 に示す通りである。

その上で, 子音冒頭部引き延ばしの持続時間を 50 ms 区切りに 0 ms から 200 ms まで変化させ, /r/と/l/を含む刺激音それぞれについて 4 条件を用意した。各条件で引き延ばされた子音冒頭部の前に, 振幅の立ち上がり区間として 10 ms 設けているが, その区間は子音の引き延ばし区間として認めない。

### 2.3 方法および参加者

本実験では, 作成した英語の/r/を含む刺激音

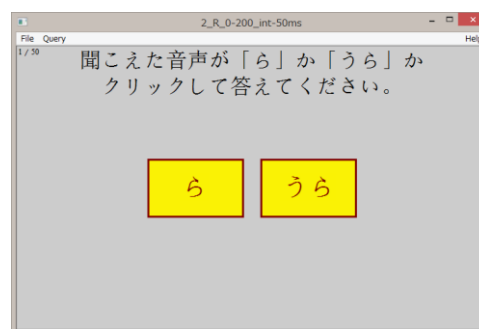


図 2. 回答画面の例

と/l/を含む刺激音について 2 者強制選択(2AFC)方式の識別実験を行った。選択肢は, 音節前に母音/u/を含む「うら」と含まない「ら」の 2 つである。タスクは, /r/を含む刺激音と/l/を含む刺激音, それぞれ別々に行った。参加者の半数は /r/を含む刺激音の識別をはじめに行い, 半数は /l/を含む刺激音の識別をはじめに行ったが, 参加者には刺激音提示の順番に関する情報は与えなかった。さらに, 参加者は「人工的な音声を聴く」とだけ伝えられ, 英語音声であることは事前に伝えなかった。それぞれのタスクにおいて刺激音はランダムに再生された。各条件の刺激音につき 10 回の繰り返しがあり, 実験参加者は合計で 100 刺激を聴取した(2 子音×5 条件×10 回繰り返し)。実験の前に, タスクに慣れてもらうための練習を行ったが, 練習の結果に対するフィードバックは与えていない。刺激音の提示は Praat ソフトウェア<sup>10)</sup>を用いて行った。

実験には, 福井大学に通う大学生 10 名(男性 7 名, 女性 3 名)が参加した(20~24 歳, 平均 21.9 歳)。参加者は, 静かな部屋でヘッドフォンを装着し, モニター画面上のボタンをクリックすることで回答を入力した。モニターには, 日本語表記で「ら」および「うら」の選択肢が表示され, 参加者はどちらが聴こえたかその選択肢をクリックすることで回答した。回答画面の例は図 2 に示す。

音量は, 参加者の聞きやすいレベルに参加者が自ら調整した。音量の調整は練習中に行うものとし, 実験中は音量を変えないよう指示した。

## 2.4 結果

### 2.4.1. /r/を含む刺激音

図 3 は, /r/を含む刺激音に対して音節前に母音を知覚した割合(「うら」と答えた割合)の平均を条件毎に示している。分散分析の結果, 主効果が見られた ( $F(1, 498)=221, p < 0.001$ )。Bonferroni 補正を用いた多重比較の結果, 子音

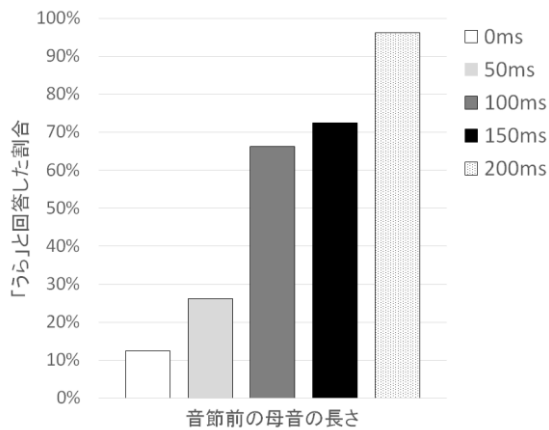


図 3. 英語の/r/の前に母音を付与する割合

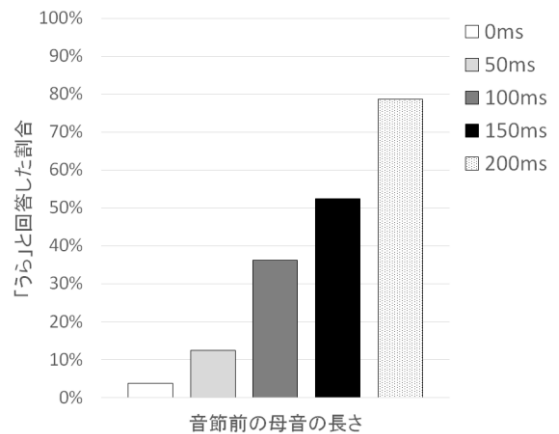


図 4. 英語の/l/の前に母音を付与する割合

部の持続時間が 0 ms と 50 ms の条件間と、100 ms と 150 ms の条件間には有意差が見られなかった。その他の条件間では有意差が見られた ( $p < 0.001$ )。ここから、/r/を含む刺激に関しては、聴取者は 0~50 ms と 100~150 ms、そして 200 ms 以上について知覚的なグループ分けをしていることが示唆される。

#### 2.4.2. /l/を含む刺激音

図 4 は、/l/を含む刺激音に対して「うら」と答えた割合の平均を条件別に示している。分散分析の結果、主効果が見られたため ( $F(1, 498)=202, p < 0.001$ )、Bonferroni 補正を用いて多重比較を行った。結果、0 ms の条件と 50 ms の条件間を除き、条件間に有意差が見られた ( $p < 0.05$ )。従って、/l/を含む音節の場合、50 ms 以上の条件下においては、子音の持続時間の長さに従って「うら」と回答する割合が増加している。しかしながら、子音の長さが 150 ms になってもなお、「うら」と回答する割合は 50%程度である。

### 3 まとめと考察

先行研究にて報告された、/r/もしくは/l/から始まる音節を聴いた際、音節前に母音を知覚的に付与するという点について、調音時間が長い子音部を母音と誤知覚している可能性が疑われた。知覚実験の結果、/r/については、100 ms 以上子音部が続いたときに「母音あり」と判断されることが分かった。一方/l/については、200 ms の子音部が続いた場合でも、「母音あり」と判断する回答が 100%になることはなかった。

#### 3.1 全体の考察

結果の差は、/r/と/l/の本質的な差を反映していると考えられる。一般的に、/r/は/l/よりも聞こえ度が高いとされる。Selkirk<sup>[11]</sup>によると、聞こ

え度は/u/ > /r/ > /l/の順で高い。この要因の 1 つに、調音の違いがある。英語の/r/は口腔内に閉鎖や、大きな空気の流れの途切れがないため、母音性が高い。一方、英語の/l/は舌先が歯茎硬口蓋付近に触れていることで、正中での閉鎖が起こり、空気の流れがより阻害される。従って、子音性が高い。さらに、フォルマント遷移の持続時間もその要因の 1 つと考えられる。/r/と/l/の遷移の、特に F1 の持続時間を比較すると(図 1 参照)、/l/の方が/r/よりも短い。これは、調音の素早い動きを反映している。このような調音の変化の本質的な違いからも、/l/の方が/r/よりも子音性が高いといえる。このことを裏付ける現象として、英語には、r 色のついた(r-colored)母音が存在する。つまり、/r/の調音的特徴を残した母音が起こりうる。そこから、/r/の遷移前の冒頭部を引き延ばした音声は/r/の特徴が残った母音として(誤)知覚されることは妥当である。

さらに、/r/を含む刺激音に対して 100 ms 以上の条件で「うら」と答える割合が急増した理由については、以下のように説明できる。音節内に出現する母音は、英語、日本語に関わらず、約 100~150 ms の長さで発話される<sup>[12,13]</sup>。そのため、本実験で用いた刺激音を英語、日本語のどちらとして聞いていたとしても、/r/を含む刺激音において 100~150 ms の母音部を持つ刺激音から急激に「母音あり」の回答が増えたことは妥当である。またここから、先行研究で日本語話者は/r/を含む音節の前に/u/を付与するという報告があったのは、そこで用いられた刺激音の/r/の前に少なくとも 50 ms 以上、母音と解釈できる成分が存在していたのではないかと推測できる。

一方、/l/は子音性が高いという点で、子音の冒頭部を引き延ばした音声は母音として知覚されるには不自然だったと考えられる。そして、持続時間が長い/l/は、/l/を調音する際の声帯振動として知覚され、結果「母音あり」という判断が少なかった可能性がある。/l/を含む音声の知覚傾向については、考察の余地がある

### 3.2 今後の課題

本研究の実験参加者は、授業等で日常的に英語を使用しているため、英語音声には親しみがあつたと思われる。しかしながら、英語能力および海外滞在歴にばらつきがあるため、本実験結果と英語能力との関係を議論するには不十分である。参加者の英語能力は、刺激音を日本語と英語のどちらとして知覚するのかにも大きく影響すると考えられるため、英語能力の統制は今後の検討課題である。

さらに、今後刺激音を何語として知覚するかについて議論をする際、実験での選択肢の提示言語についても考慮に入れる必要がある。本調査は2AFC法を用いて行い、選択肢の表記として日本語の平仮名表記を用いた(図2)。従って、表記から、聴取者は刺激音を日本語であろうと推測した上で判断を行っている可能性が考えられる。そのような点に配慮した上で、今後、刺激音を英語、あるいは日本語として知覚した場合の傾向の差異についても明らかにする必要がある。

## 4 結論

先行研究で報告されていた、日本語母語話者が英語の/r/もしくは/l/を頭子音として含む音節の前に知覚的に母音/u/を付与するという現象について、子音の持続時間が長い、もしくは長いと知覚された場合、それを母音と誤知覚した可能性が考えられた。そこで本研究では、問題の子音を人工的に作成し、子音部を4段階に引き延ばすことで、子音部の長さとも母音の知覚的

付与について検証した。結果、/r/と/l/では異なる母音付与の傾向が見られた。傾向の違いは、それらの子音の本質的な差異に起因すると考えられるが、詳細の検証が必要である。

### 謝辞

実験実施にご協力くださった福井大学大学院教育学研究科及び教育地域科学部、同学部言語教育コースの学生の方々に感謝申し上げます。

### 参考文献

- [1] Goto, *Neuropsychol.*, 9, 317–323, 1971.
- [2] Logan *et al.*, *J. Acoust. Soc. Am.*, 89 (2), 874–886, 1991.
- [3] Best, *Haskins Lab. Stat. Rep. of Sp. Research.*, SR-107, 1–30, 1991.
- [4] Guion *et al.*, *J. Acoust. Soc. Am.*, 107 (5), 2711–2724, 2000.
- [5] Dupoux *et al.*, *JEP: Human Percept. Perform.*, 25 (6), 1568–1578, 1999.
- [6] Klatt, *J. Acoust. Soc. Am.*, 67 (3), 971–994, 1980.
- [7] Klatt & Klatt, *J. Acoust. Soc. Am.*, 87 (2), 820–857, 1990.
- [8] Peterson & Barney, *J. Acoust. Soc. Am.*, 24 (2), 175–184, 1952.
- [9] Tomaru & Arai, *Acoust. Sci. Tech.*, 35 (3), 166–169, 2014.
- [10] Boersma & Weenink, Praat: doing phonetics by computer [Computer program]. Version 6.0.07, retrieved 30 November 2015 from <http://www.praat.org/>
- [11] Selkirk, *On the major class features and syllable theory*, MIT Press, 1984.
- [12] Minagawa *et al.*, *15<sup>th</sup> ICPHS*, 2127–2130, 2003.
- [13] Umeda *et al.*, *J. Acoust. Soc. Am.*, 58 (2), 434–445, 1975.