

## 日本語母語話者と非日本語母語話者の音声の知覚的区別

## Perceptual discrimination of Japanese utterances of native and non-native Japanese speakers

山川 仁子\*, 天野 成昭\*\*

Kimiko YAMAKAWA, Shigeaki AMANO

## 要 旨

高い日本語能力を持つ非日本語母語話者が発声した日本語音声と、日本語母語話者が発声した日本語音声とを日本語母語話者が聞き分けられるか否かを調査する目的で、両者の音声を刺激として聴取実験を行った。その結果、非日本語母語話者の音声に音素の違いや言い淀みがないにもかかわらず日本語母語話者と非日本語母語話者を95.4%という非常に高い確率で区別することができた。しかし聴取実験で刺激として用いた音声について母音部・子音部の時間長の標準偏差 ( $\Delta All$ ) と変動係数 (VarcoALL) の二つの変数を組み合わせて母語の判別分析を行ったところ、判別率は57.5%となり、日本語母語話者と非日本語母語話者を区別することができなかった。これらの結果は、 $\Delta All$  や VarcoALL などの時間長にかかわる音響特徴ではなく、それ以外の音響特徴の差によって生じる自然性の低下に、日本語母語話者が敏感に反応している可能性が高いことを示唆している。

キーワード：日本語母語話者 非日本語母語話者 音声知覚 自然性

## 1. はじめに

非日本語母語話者が発声した日本語音声は、その内容が明確に聞き取れたとしても、日本語母語話者が発声した日本語音声よりも不自然に聞こえる場合が多い。音声には、伝達内容だけでなく感情的かつ社会的な要素も含むため (Stevick, 1978), 不自然な発声をする非日本語母語話者は社会的地位や教育レベルが劣るとみなされる傾向があり、そのような傾向は良好な人間関係を築く上でのマイナス要因となる (Thomas, 1983)。したがって自然な発声は、人間関係も含めた円滑なコミュニケーションをはかるうえで重要な要因であると言え、非日本語母語話者自身も自然な日本語の発声の必要性を強く感じている (日本語教育学会, 1991)。しかし、非日本語母語話者にとって日本語の自然な発声の習得は容易ではない。

音声の自然性には、時間長、強度、基本周波数 ( $F_0$ ), スペクトルなどの音響特徴が関係しているといわれている (Ramus et al., 1999; Kato et al., 2002; Niebuhr et al., 2009; Loukina et al., 2011 など)。しかし、日本語音声の自然性を表す音響的特徴や変数の詳細は、まだ明らかにはされていない。

一方、天野ら (2012) によれば、日本語母語話者と非日本語母語話者が発声した日本語音声におけるリズムの違いを表す特徴は、母音部・子音部の時間長の標準偏差 ( $\Delta All$ ) と変動係数 (VarcoALL) の二つの変数の組み合わせで表される可能性が示唆されている。

そこで本研究では、上級レベルの非日本語母語話者が発声した音声と、日本語母語話者が発声した音声を

\* 愛知淑徳大学人間情報学部 jin@asu.aasa.ac.jp

\*\* 愛知淑徳大学人間情報学部

刺激として、両者の聞き分けが可能であるか否かを調べる聴取実験を行った。さらに、非日本語母語話者と日本語母語話者が発声した音声を対象に、二つの時間的音響特徴 ( $\Delta$  All, VarcoALL) を用いて、母語の判別分析を行い、非日本語母語話者と日本語母語話者の音声を区別できるか否かを検証した。

## 2. 聴取実験

### 2.1. 実験参加者

愛知淑徳大学人間情報学部の学生 18 名（男性 4 名，女性 14 名）が実験に参加した。参加者の平均年齢は 20.2 歳（Max = 21 歳，Min = 20 歳，SD = 0.43 歳）であった。今回実験に参加した 18 名はすべて、普段、非日本語母語話者と接する機会はなかった。

### 2.2. 音声刺激

表 1 に示す 50 単語をキャリア文「これは○○だとおもいます」に当てはめ、それを日本語母語話者 6 名および非日本語母語話者（中国語母語話者）7 人の合計 13 名が発声した音声を刺激として用いた。総刺激数は 650（50 単語 × 13 名）であった。この音声刺激のセットから、各話者について 2 刺激をランダムに選び、練習用の刺激とした。練習用の刺激数は 26（2 単語 × 13 名）であった。なお、刺激の音声を発声した非日本語母語話者（中国語母語話者）7 名全員が、日本語能力試験 JPLT（Japanese Proficiency Level Test）の最高レベルである N1 に合格している。これは彼らの日本語能力が非常に高いことを示している。また、彼らの日本語学習時間は平均 1997 時間（Max = 2700 時間，Min = 1800 時間，SD = 357.9 時間）であった。

表 1：刺激として用いた単語

ID	読み	表記	ID	読み	表記
w009	かこ	過去	w172	しょう	小
w018	かっこ	括弧	w180	しょう	使用
w057	きた	来た	w190	じゅう	十
w059	きった	切った	w191	じゅうしょ	住所
w060	きって	切手	w192	じゅうしょう	重症
w062	きて	来て	w195	じゅう	自由
w064	きねん	記念	w206	すし	寿司
w069	きょう	今日	w207	すじ	筋
w072	きょう	器用	w212	すみ	隅
w073	きんえん	禁煙	w214	せいかい	正解
w078	くき	茎	w217	せかい	世界
w079	くぎ	釘	w227	せんえん	千円
w080	くし	串	w229	せんねん	千年
w081	くじ	九時	w238	たんい	単位
w082	くろ	黒	w239	たんに	単に
w083	くろう	苦勞	w245	ちかい	誓い
w102	こな	粉	w246	ちがい	違い
w104	こな	こな	w253	つき	月
w110	ごご	午後	w254	つぎ	次
w111	ごごう	五号	w265	でわ	では
w113	さか	坂	w266	でんわ	電話
w115	さっか	作家	w268	とおし	通し
w132	しって	知って	w271	とし	年
w134	して	して	w337	りゅう	竜
w170	しゅみ	趣味	w338	りゅう	理由

### 2.3. 実験場所および実験装置

愛知淑徳大学の聴覚情報実験室において音声刺激の聴取実験を行った。ノート型パーソナルコンピュータ (Dynabook SS RX 2 PARX 2 T 9 LLJ, TOSHIBA) に A/D コンバータ (Sound Blaster X-Fi SB 1090, Creative) とヘッドホン (MDR-Z 900 HD, SONY) を接続し、ノート型パーソナルコンピュータに保存した音声刺激を両耳に呈示するよう設定した。

### 2.4. 手続き

コンピュータに保存した音声刺激を、ヘッドホンから参加者の両耳に呈示した。音声刺激の音量は聞きやすい大きさ (Windows 音量ミキサーのスピーカー設定値 60) とし、参加者毎に全刺激をランダム順に呈示した。呈示された音声刺激が、日本語母語話者と非日本語母語話者のどちらによって発声された音声であるかを参加者に判定させ、その判定結果を、コンピュータの画面上に表示した「日本語母語話者」と「非日本語母語話者」の反応ボタンのどちらか一方をマウスでクリックすることによって入力させた。なお、参加者の半数には「日本語母語話者」の反応ボタンを右に表示し、残りの半数には「日本語母語話者」の反応ボタンを左に表示して、ボタンの表示位置のカウンターバランスをとった。参加者が反応ボタンのうちの1つをクリックした後に、画面上に表示した「決定」のボタンをクリックさせ、次の試行に移行した。

参加者1人につき上記の実験装置1台を割り当て、個人ペースで実験を行った。初めに26試行の練習を行った。練習では、13試行毎に約1分の休憩を入れた。その後に、650試行の本試行を行った。本試行では、165試行毎に約3分の休憩を入れた。練習を含めた全実験時間は約1時間であった。

### 2.5. 結果

日本語母語話者の音声刺激が日本語母語話者と判定された回数を全刺激呈示回数で割り、日本語母語話者と判定された率を求めた (図1)。同様にして非日本語母語話者の音声为非日本語母語話者として判定された率を求めた (図2)。図1, 2から日本語母語話者の音声刺激はほとんどが日本語母語話者と判定され、非日本語母語話者の音声刺激はほとんどが非日本語母語話者と判定されていることが見て取れる。日本語母語話者の音声が「日本語母語話者である」と正しく判定された率の平均は95.4%であった。一方、非日本語母語話者の音声「非日本語母語話者である」と判定された率の平均は95.4%であった。両者を併せた全体での判定率の平均は95.4%であった。

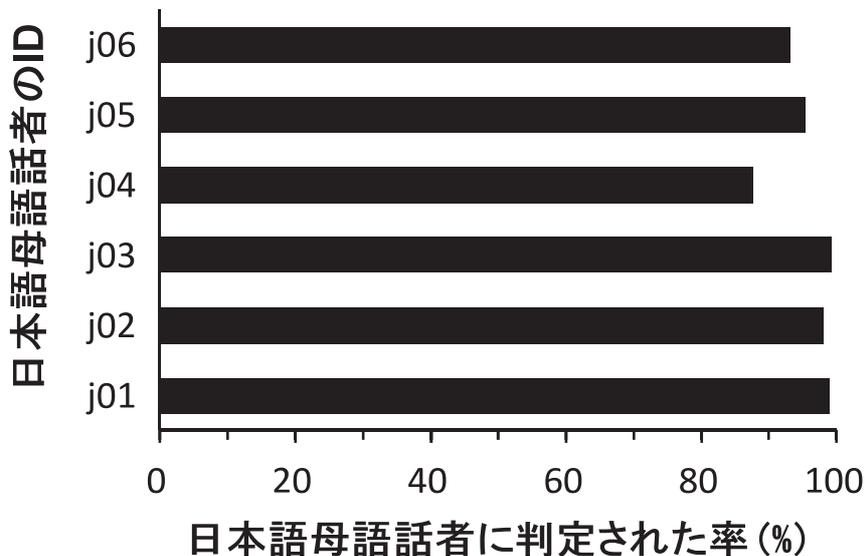


図1. 日本語母語話者の音声が日本語母語話者として判定された率

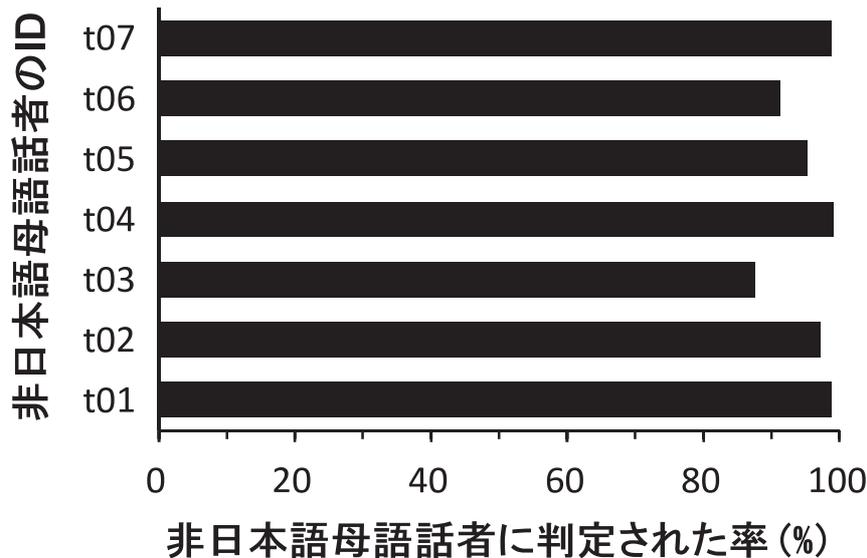


図2. 非日本語母語話者（中国語母語話者）の音声为非日本語母語話者として判定された率

### 3. 音響解析

天野ら（2012）が提案した発話音声における母音部・子音部の時間長の標準偏差（ $\Delta All$ ）と母音部・子音部の時間長の変動係数（VarcoALL）の2つの変数を組み合わせて聴取実験で使用した音声を対象に、母語の判別分析を行った。 $\Delta All$ は発話音声における時間長のばらつき、VarcoALLは発話速度を考慮した時間長のばらつきを表す。

#### 3.1. 解析対象

本研究の聴取実験で使用した日本語母語話者6名および非日本語母語話者（中国語母語話者）7名の合計13名が発声したキャリア文を含めた650単語（50単語×13名）の音声文を解析対象とした。

#### 3.2. 方法

解析対象単語を含むキャリア文全体に対し、各音韻の開始・終了時刻と音韻ラベルを付与した。ラベリングで得た時刻から、母音部・子音部の時間長の標準偏差（ $\Delta All$ ）と母音部・子音部の時間長の変動係数（VarcoALL）を求めた。なお本研究における子音部とは、ポーズなどの無音区間も含む母音以外の区間（Ramus et al., 1999）を指す。 $\Delta All$ とVarcoALLを独立変数とし、母語の種類を従属変数として、判別分析を行った。

#### 3.2. 結果

判別分析の結果、 $\Delta All$ とVarcoALLの2変数を用いた母語の判別率は57.5%であった。これはチャンスレベルの判別率50%に非常に近い値であった。つまり、 $\Delta All$ とVarcoALLで表された時間長の音響特徴では、本研究で用いた非日本語母語話者と日本語母語話者の音声を区別することはできないということが明らかになった。

### 4. 考察

聴取実験において日本語母語話者の音声は日本語母語話者に判定される率も、非日本語母語話者が非日本語母語話者に判定される率も約95%と高いことから、非日本語母語話者が発声した日本語音声であることを、日

本語母語話者が聞き分ける精度は極めて高いと言える。一般的に考えれば、日本語能力があまり高くない非日本語母語話者ならば、日本語音声をつたつどしく発声する場合はほとんどであるので、非日本語母語話者が発声した日本語音声であることを聞き分けるのは容易だろう。一方で、本研究で用いた非日本語母語話者の音声は、音素の間違いや言い淀みのない流暢な音声であり、非日本語母語話者が発声した日本語音声であることを聞き分けるのは困難だろうと予測された。しかし、聴取実験で示されたように、日本語能力試験 JPLT (Japanese Proficiency Level Test) N1 という非常に高い日本語能力を持つ非日本語母語話者であっても、彼らが発声した日本語音声は、非日本語母語話者が発声した日本語音声であると容易に聞き分けられた。この結果は、わずかな音響特徴の差によって生じる自然性の低下に、日本語母語話者が敏感に反応している可能性が高いことを示唆している。

音声の自然性を表す変数として時間長のばらつきに注目し、母音部・子音部の時間長の標準偏差 ( $\Delta All$ ) と母音部・子音部の時間長の変動係数 (VarcoALL) を用いた母語の判別分析を行ったところ、日本語母語話者と非日本語母語話者を区別することは困難であることが示された。このことから、日本語母語話者と非日本語話者の音声を区別する音響特徴には、時間長のばらつき以外の音響的特徴が影響していることが考えられる、すなわち、抑揚・イントネーション、さらには母音のスペクトル等において、日本語母語話者と非日本語母語話者の音声の僅かな差を、日本語母語話者が敏感に感じ取っていることを意味している。

日本語能力が高い非日本語母語話者が発声した音声は、通常、その内容をほぼ正確に聞き取れるので明瞭性は高いといえる。しかし本研究の結果は、非日本語母語話者の音声の自然性において日本語母語話者の音声とは異なっていることを示唆している。今後、このような自然性の差をもたらす原因となる音響特徴の詳細を解明できれば、非日本語母語話者がより自然な日本語を発声するための教育方法等への応用が期待できるであろう。

## 謝辞

本研究は科研費 (22320081, 24652087, 25284080)、愛知淑徳大学研究助成 (共同研究 2013-2014 年度, 特定研究 2013-2014 年度)、国立国語研究所共同研究プロジェクト「コーパス日本語学の創成」の助成を受けた。We would like to thank Professor C. Lin and M. Hsu at Ming Chuan University, Professor R. Shen at Shih Hsin University, and Professor Y. Maa at Tamkang University for their assistance in utterance recording.

## 引用文献

- 天野成昭, 山川仁子, 近藤眞理子 (2012). “日本語母語話者と中国語母語話者の日本語音声を区別する時間的変数,” 日本音響学会春季講演論文集, 409-410.
- Kato, H., Tsuzaki, M., and Sagisaka, Y. (2002). “Effects of phoneme class and duration on the acceptability of temporal modifications in speech,” *J. Acoust. Soc. Am.* 111 (1), 387-400.
- Loukina, A., Kochanski, G., Rosner, B., and Keane, E. (2011). “Rhythm measures and dimensions of durational variation in speech,” *J. Acoust. Soc. Am.* 129 (5), 3258-3270.
- Niebuhr, O. (2009). “F0-Based Rhythm Effects on the perception of Local Syllable Prominence,” *Phonetica* 66, 95-112.
- 日本語教育学会 (編) (1991). 『日本語教育機関におけるコース・デザイン』, 凡人社.
- Ramus, F., Nespors, M., and Mehler, J. (1999). “Correlates of linguistic rhythm in the speech signal,” *Cognition* 73, 265-292.
- Stevick, E. (1978). “Toward a practical philosophy of pronunciation; Another view,” *TESOL Quarterly*, 12 (2), 145-150.
- Thomas, J. (1978). “Cross-cultural pragmatic failure,” *Applied Linguistics* 4, 91-112.