

科学研究費助成事業 研究成果報告書

令和 2 年 7 月 15 日現在

機関番号：37404

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K02705

研究課題名(和文)日本語音声の自然性を予測する定量的モデルの構築

研究課題名(英文)Prediction model for naturalness of spoken Japanese

研究代表者

山川 仁子(Yamakawa, Kimiko)

尚絅大学・文化言語学部・准教授

研究者番号：80455196

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：非日本語母語話者が話す日本語音声は不自然に聞こえる場合が多い。この不自然さの原因を探るために、非日本語母語話者の日本語音声における音素の時間長・基本周波数・強度等の音響解析を行い、日本語母語話者の日本語音声と比較した。その結果、自然性に関わる音響特徴として時間長と基本周波数が大きな役割を果たしていることが明らかになった。そこで時間長と基本周波数を系統的に変化させた音声を用いて知覚実験を実施し、これらの音響特徴が自然性に与える影響の大きさを特定した。さらに、重回帰分析を用いて、これらの音響特徴から日本語音声の自然性を予測する定量的モデルを構築した。

研究成果の学術的意義や社会的意義

本研究の成果は、これまで不明だった日本語音声の自然性に関する特性・原理を科学的に解明し、総合的に理解するための第一歩として学術的価値が高い。また、その知見に基づいた教育方法や音声教育ソフトウェア等の開発により、非日本語母語話者に対するIT機器による合理的・効果的な日本語音声教育の実現に寄与が期待できる。この日本語音声教育の実現は、国内外の非日本語母語話者との音声コミュニケーションの円滑化をはかり、文化・政治・経済等における日本と世界各国との連携強化のための基盤として重要な役割を果たすとと言える。

研究成果の概要(英文)：Japanese spoken by non-native Japanese speakers is often perceived as "unnatural Japanese" by native Japanese speakers. To explore a cause of the unnaturalness, phoneme duration, fundamental frequency, and intensity were analyzed in Japanese speech by non-native speakers. It was revealed that unnaturalness is caused mostly by deviation of phoneme duration and fundamental frequency from those in Japanese speech by naive Japanese speakers. Their deviation effects on naturalness were measured with a rating scale method for a stimulus continuum of phoneme duration and fundamental frequency. From A quantitative predicting model of naturalness of Japanese speech was developed using a multiple regression equation with the phoneme duration and fundamental frequency.

研究分野：音声科学

キーワード：日本語音声 自然性 相対時間長 相対基本周波数 相対強度 日本語学習者 自然性予測の定量的モデル

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

1. 研究開始当初の背景

国外で日本語を学習する非日本語母語話者は約 398 万人に及び、しかもその数は年 20 万人のペースで増加傾向にある[1]。また、アジアの多くの国の中等教育機関において日本語が第 2 外国語として採用され[1]、さらにベトナムの初等教育では 2016 年 9 月に日本語が第 1 外国語として採用された[2]。これらの事実は国外において日本語教育の需要が高く、かつその充実が求められていることを示している。その中でも特に重要度が高まっているのは、日本語の音声教育である。なぜならば音声には、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[3]、不自然な発音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、良好な人間関係を築く上でのマイナス要因となる[4]からである。この点を非日本語母語話者自身も認識し、自然な日本語の発音の必要性を強く感じている[5]。しかし、非日本語母語話者にとって日本語の自然な発音の習得は容易ではなく、日本語教師による丁寧な発音指導が必要である。ところが、国外の日本語教師数は僅か 6 万人に不足しており[1]、その数が絶対的に不足している。また自然性に関する科学的知見が乏しく、不明点も多いため、自然性を高めるための科学的・合理的な教育方法も確立されていない。

この状況を改善するためには、非日本語母語話者における不自然な発音の特性と問題点を科学的に解明し、その知見を日本語音声教育に反映させていくと同時に、工学的応用を見据えた戦略的研究を推進し、早期にIT機器による日本語音声教育の実現を目指す必要がある[6]。

以上の背景を踏まえ、山川らは非日本語母語話者が不得意とする日本語音声の破擦音や特殊モーラについて、音声知覚・生成の両面から一連の科学研究を行ってきた[7,8]。その研究を進める中で、日本語母語話者は僅かな音響特徴の差によって生じる自然性の低下に敏感であり、非日本語母語話者の音声に音素やアクセントの違いが無い場合であっても自然性が低いと判断するケースが多いことが明らかになった[9]。

そこで山川らは、6 言語(中国語・韓国語・タイ語・ベトナム語・仏語・英語)の非日本語母語話者の日本語音声を用いて、その自然性に寄与する音響特徴を解析する研究を行った[科研費基盤 C, 代表者 山川仁子, 2014~2016]。その結果、無声化母音や特殊モーラなどの特殊ユニットを含まない日本語音声では、自然性に対して母音の強度はあまり寄与せず、母音の時間長や基本周波数(F0)が強く寄与していることが明らかになった。ただし、この研究では特殊ユニットを含まない音声を用いたため、特殊ユニットを含む音声全体への一般化ができないという難点があった。しかも非日本語母語話者が、この無声化母音や特殊モーラなどの特殊ユニットの発音を不得意としている点を考慮すると、特殊ユニットを含む音声では自然性の低下がより強く表れる可能性が考えられた。

2. 研究の目的

本研究では上記の問題点を踏まえ、無声化母音や特殊モーラなどの特殊ユニットを含む日本語音声の時間長・基本周波数(F0)・強度に関する詳細かつ網羅的な音響解析と知覚実験を実施し、自然性に関する科学的知見を得るとともに、音響特徴から日本語音声の自然性を予測する定量的モデルを構築することを目的とした。具体的には以下の 3 項目を研究目標とした。

- (1) 日本語音声の自然性に関わる音響特徴の特定と相互作用の解明
無声化母音や特殊モーラなどの特殊ユニットを含む日本語音声の音響解析と自然性の評価実験を組み合わせ、自然性に関わる音響特徴の特定と音響特徴間の相関解析を行う。
- (2) 日本語音声の自然性に関わる音響特徴の許容範囲の特定
(1)で特定された音響特徴を系統的に変化させた刺激連続体を作成し、それを用いた自然性の知覚実験を実施することにより、自然性に関わる音響特徴の許容範囲を特定する。
- (3) 日本語音声の自然性を音響特徴から予測する定量的モデルの構築
(1)および(2)で得られた科学的知見を基に、音響特徴を独立変数、自然性を従属変数とする重回帰分析・判別分析等の多変量解析を行い、自然性の定量的予測モデルの構築を行う。

3. 研究の方法

上記の研究目的に沿って、非日本語母語話者の発声した音声の音響解析および知覚実験を以下の方法で実施した。

(1)音響解析

日本語音声の自然性に関わる音響特徴を特定するために、下記の方法で音響解析を行った。

①解析対象

Yamakawa, Amano & Kondo[11]が構築した非日本語母語話者の日本語音声データベースから、2 または 3 モーラ長の 0 型アクセントの 27 単語をキャリア文「これは○○だと思います」に埋め込んだ音声抽出した。抽出した音声の発声者はフランス語、韓国語、中国語(台湾)、タイ語、ベトナム語の非日本語母語話者各 10 名、および日本語母語話者 10 名であった。抽出した音声を日本語母語話者 22 名(男性 2 名、女性 20 名、平均年齢 27.9 歳)に聴取させ、(1)発声の正誤判断と(2)音声の自然性の 5

段階評定を行わせた。聴取の結果、発声が「正しい」と判断された場合の音声と、その音声の自然性評価値の平均を解析に用いた。

②解析方法

解析対象の各音声について、キャリア文を含む全ての母音の時間長、強度、基本周波数(Fo)、および母音中心時刻(VCT)を計測し、各測定値の全体に対する相対値を求め音響特徴量とした。次に、日本語母語話者の音声において自然性評価値の平均を求め、その平均より自然性評価値が高い日本語母語話者の音声を特定した。特定した音声の母音毎に各音響特徴量の平均を求め、日本語母語話者における基準値とした。非日本語母語話者の音声の各母音における各音響特徴量とこの基準値とのRMSE(Root Mean Squared Error)を求めた。これらをそれぞれ(1)相対時間長距離、(2)相対強度距離、(3)相対Fo距離、(4)相対VCT距離と呼ぶことにする。この4種の距離を対数変換した値を独立変数とし、自然性評価値を従属変数として強制投入法による重回帰分析を行った。

(2)知覚実験

日本語音声の自然性に関わる音響特徴の許容範囲の特定するため、音響解析で特定された音響特徴を用いて知覚実験を実施した。

①音声刺激

音声分析合成ソフトウェア Praat を用いて、非日本語母語話者の音声の基本周波数 Fo および音素の時間長を 0%, 25%, 50%, 75%, 100% の幅で変化させ、自然性の高い音声から自然性の低い音声へと変化する刺激連続体を作成した。

②手続き

音声刺激を実験参加者にヘッドホンを通してランダム順に両耳呈示した。実験参加者には呈示された音声がどの程度自然に聞こえるかを、1(不自然)から5(自然)までの5段階評価尺度を用いて判断させた。

4. 研究成果

(1)主な研究成果

①日本語音声の自然性に関わる音響特徴の特定

日本語母語話者と各非日本語母語話者のデータをペアとした場合の重回帰分析における標準化偏回帰係数を図1に示す。非母語話者全体のデータを用いた場合の修正済み決定係数は0.628であり、中程度の当てはまりであった。相対時間長距離、相対Fo距離、相対VCT距離の標準化偏回帰係数はどれも0.1%水準で有意であることから、相対時間長、相対Fo、相対VCTの距離は、日本語音声の自然性を表す変数として説明力があるといえる。なお、各変数を対数変換しない場合の決定係数は0.568であり、各変数を対数変換すると決定係数が高くなるといえる。フランス語母語話者、韓国語母語話者、タイ語母語話者の修正済み決定係数は、それぞれ0.609、0.443、0.650であった。相対Fo距離の標準化偏回帰係数の絶対値が他の変数に比べて大きいことから、フランス語母語話者、韓国語母語話者、タイ語母語話者において、相対Fo距離は日本語音声の自然性を表す変数として説明力があるといえ

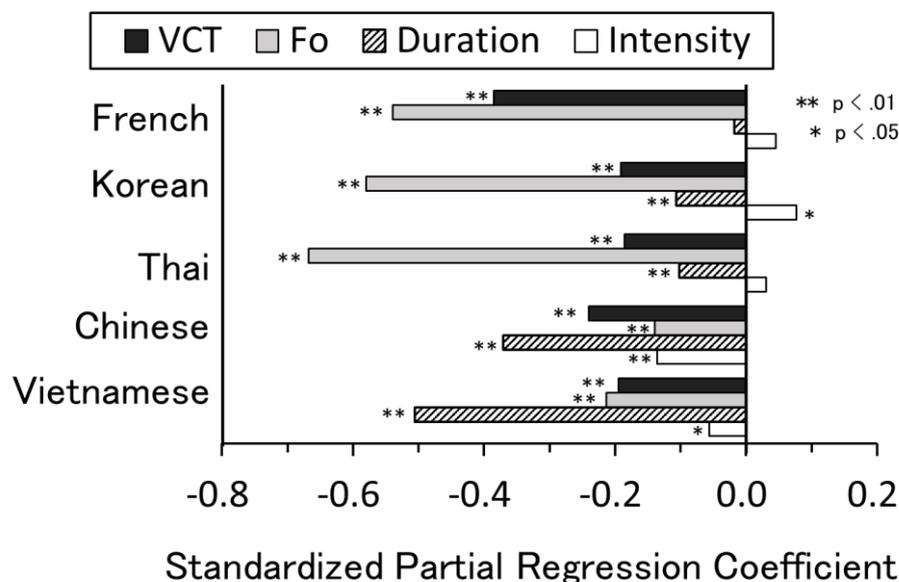


図1 言語ごとの標準化回帰係数

る。つまり、これらの非日本語母語話者では、F₀ パタンのずれが自然性低下の要因になっていると考えられる。中国語母語話者、ベトナム語母語話者の修正済み決定係数は、それぞれ 0.466、0.772 であった。相対時間長距離の標準化偏回帰係数の絶対値が他の変数に比べて大きいことから、中国語母語話者とベトナム語母語話者において、相対時間長距離は日本語音声の自然性を表す変数として説明力があるといえる。すなわち、中国語母語話者、ベトナム語母語話者は、母音の長さのずれが自然性低下の要因になっていると考えられる。以上の結果より、F₀ パタンおよび母音の長さのずれが日本語の自然性低下の要因であるといえる。さらに、言語によってその要因は異なることも示唆された。

②日本語音声の自然性に関わる音響特徴の許容範囲の特定

知覚実験にて得られた自然性評価結果を図 2 に示す。図 2 の縦軸は自然性評価値を表し、横軸は音素の時間長の変化率を表す。棒グラフの各要素は基本周波数 F₀ の変化率を表す。自然性の評価値が 4 以上の刺激を「日本語として自然である」とみなすと、音素の持続時間の変化率が 0% の場合、すなわち日本語母語話者と同様の音素時間長である場合、基本周波数 F₀ の変化率 25% 以下は「日本語として自然である」と許容される。音素の時間長の変化率が 25% である場合は、基本周波数 F₀ の変化率が 0%、すなわち日本語母語話者と同様の基本周波数である場合のみが「日本語として自然である」と許容されることが示唆された。

全実験参加者の評価値データをプールし、基本周波数 F₀ および音素の時間長を要因とする 2 要因分散分析を行った結果、交互作用が有意であった。これは基本周波数 F₀ と音素の時間長が相乗的に自然性評価値に影響を及ぼすことを意味している。交互作用が有意であったので、基本周波数 F₀ と音素の時間長の各要因における単純主効果検定を行ったところ、いずれの単純主効果も 5% 水準で有意であった。したがって、基本周波数 F₀ も音素の時間長も自然性に影響を及ぼす音響特徴であるといえる。

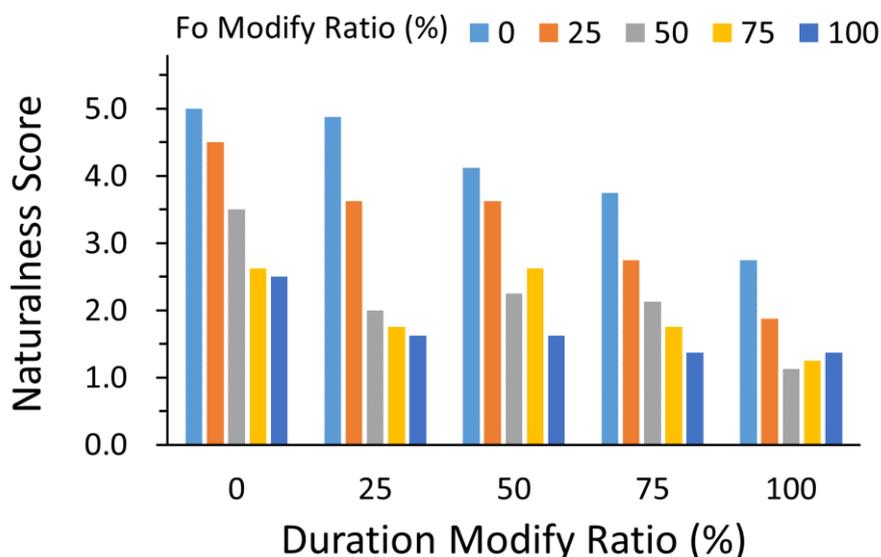


図 2 自然性評価実験結果

③日本語音声の自然性を音響特徴から予測する定量的モデルの構築

日本語音声の自然性を音響特徴から予測する定量的モデルを重回帰式によって構築した(式 1)。

$$y = 2.078 x_1 + 1.623x_2 - 332.5 \quad (1)$$

式(1)において、y は自然性評価値を表し、x₁ は基本周波数 F₀ が日本語母語話者から離れている割合、x₂ は音素の時間長が日本語母語話者から離れている割合を表す。式(1)の決定係数は 0.88 であり、あてはまりのよいモデルであることを示している。

(2) 今後の展望

本研究では、非日本語母語話者が発声した日本語音声の自然性を音響特徴から予測する定量的モデルを構築した。今後は、今回の研究で得られた予測モデルを用いて日本語学習者の発音訓練システムや発音評価システムの構築を進めていく予定である。しかし、課題もいくつか残されている。本研究では、日本語を含む 6 言語を対象としてモデルを構築した。新たな言語を追加した場合、本研究ではあ

まり強い影響が見られなかった強度や母音中心時刻 (VCT) などの音響特徴が自然性に関わってくる可能性が考えられる。今後、これらの可能性を検討するために、解析対象を広げて研究を進める予定である。

【引用文献】

- [1] 国際交流基金 (2012). 『2012 年度日本語教育機関調査 結果概要』, くろしお出版.
- [2] 産経ニュース (2016). "日本語が「第 1 外国語」に ベトナムの小学校で・東南アジアで初," 産経ニュース, 3 月 2 日, <http://www.sankei.com/life/news/160301/lif1603010038-n1.html> (2016 年 10 月 9 日閲覧).
- [3] Stevick, E. (1978). "Toward a practical philosophy of pronunciation; Another view," *TESOL Quarterly*, 12(2), 145-150.
- [4] Thomas, J. (1983). "Cross-cultural pragmatic failure," *Applied Linguistics* 4, 91-112.
- [5] 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.
- [6] 外務省 (2013) "海外における日本語の普及促進に関する有識者懇談会議論の総括と政策提言," <http://www.mofa.go.jp/mofaj/files/000022908.pdf> (2016 年 10 月 2 日閲覧) .
- [7] Yamakawa, K. & Amano, S. (2015). "Discrimination of Japanese fricatives and affricates by production boundaries in time and spectral domains: A case study of a female native speaker, " *Journal of the Acoustical Science & Technology*, 36(4) , 296-301.
- [8] Amano, S. & Hirata, Y. (2015). "Perception and production of singleton and geminate stops in Japanese: Implications for the theory of acoustic invariance," *Phonetica*, 72, 43-60.
- [9] 山川仁子, 天野成昭 (2014). "日本語母語話者と非母語話者の音声の弁別," 愛知淑徳大学論集人間情報学部篇第 4 号, 15-19.
- [10] Yamakawa, K. & Amano, S. (2015). "Acoustic features of Japanese words spoken by Japanese natives and non-natives", 18th International congress of phonetic sciences (ICPhS2015),1-5, Glasgow, UK.
- [11] Yamakawa, K., Amano, S., & Kondo, M. (2011). "Development of Japanese read-word database for non-native speakers of Japanese," *Proc. of Oriental COCODA*, 65-70.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 山川仁子, 天野成	4. 巻 52
2. 論文標題 非日本語母語話者による日本語音声の自然性の原因解明に向けた音響解析	5. 発行年 2020年
3. 雑誌名 尚綱大学研究紀要 A. 人文・社会科学編	6. 最初と最後の頁 152-167
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計13件（うち招待講演 0件 / うち国際学会 6件）

1. 発表者名 山川仁子, 天野成昭
2. 発表標題 非日本語母語話者の発声した日本語の自然性に関わる変数
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2018年

1. 発表者名 天野成昭, 山川仁子
2. 発表標題 破裂音/t/, 摩擦音/s/, 破擦音/ch/の促音・非促音の生成範疇境界を表す変数
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2018年

1. 発表者名 天野成昭, 山川仁子
2. 発表標題 撥音と鼻子音の生成範疇境界を表す変数
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2019年

1. 発表者名 Kimiko Yamakawa, Shigeaki Amano
2. 発表標題 Identifying acoustic features that cause unnaturalness of non-native speakers' Japanese
3. 学会等名 International Congress of Phonetic Sciences 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Shigeaki Amano, Kimiko Yamakawa, Mariko Kondo
2. 発表標題 Logarithmic duration to predict and discriminate singleton and geminate consonants in Japanese
3. 学会等名 International Congress of Phonetic Sciences ICPhS 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Shigeaki Amano, Kimiko Yamakawa
2. 発表標題 Perception boundary of Japanese singleton and geminate stops with devoiced vowels.
3. 学会等名 International Symposium on Monolingual and Bilingual Speech 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Kimiko Yamakawa, Shigeaki Amano
2. 発表標題 Two-dimensional mapping of Japanese fricative and affricates spoken by Japanese and Korean speakers.
3. 学会等名 International Symposium on Monolingual and Bilingual Speech 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 天野成昭, 山川仁子
2. 発表標題 母音の無声化を伴う促音・非促音の知覚範疇境界を表す変数
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2017年

1. 発表者名 天野成昭, 山川仁子
2. 発表標題 中国語母語話者における日本語の促音・非促音の知覚
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2017年

1. 発表者名 天野成昭, 山川仁子
2. 発表標題 促音・非促音を含む単語における発声速度と母音無声化率
3. 学会等名 日本音響学会春季研究発表会
4. 発表年 2018年

1. 発表者名 Kimiko Yamakawa, Shigeaki Amano
2. 発表標題 Identifying acoustic features that cause unnaturalness of non-native speakers' Japanese
3. 学会等名 The 18th International Congress of Phonetic Sciences (国際学会)
4. 発表年 2019年

1. 発表者名 Shigeaki Amano, Kimiko Yamakawa, Mariko Kondo
2. 発表標題 Logarithmic duration to predict and discriminate singleton and geminate consonants in Japanese
3. 学会等名 The 18th International Congress of Phonetic Sciences (国際学会)
4. 発表年 2019年

1. 発表者名 天野成昭, 山川仁子
2. 発表標題 撥音と鼻子音の生成範疇境界を表す変数
3. 学会等名 日本音響学会季春研究発表会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	天野 成昭 (Amano Shigeaki) (90396119)	愛知淑徳大学・人間情報学部・教授 (33921)	