

日本語学習者の語頭上げ音調の実態^(注1)

吉田夏也^{*1}・白勢彩子^{*2}

日本語学・日本文学分野

(2017年8月30日受理)

要 旨

音響音声学の手法を用いて、海外に在住する日本語学習者の音声における語頭上げ音調を観察した。その結果、母語話者と比較すると、学習者の語頭上げ音調の程度が小さいことが明らかになった。この結果は、語頭上げ音調が学習者の日本語発話の自然性や了解度と関連している可能性を示唆している。

キーワード：語頭上げ音調，学習効果，個人差

1. 目的

日本語の共通語アクセントでは、原則として単語の1モーラ目と2モーラ目のピッチ（高さ）が異なる^(注2)。しかし、日本語学習者の発音では、このピッチの違い（これ以降、簡便のためにこの違いを「IR=initial rising」、具体的なピッチの差をIR値と呼ぶ）が実現されないことが多いとされる^[1]。

IRは、単語や句の開始部を特徴づける音調なので、これを欠くと日本語の自然性や了解度などに影響を与えたと考えられる。本稿では、異なる母語を持つ海外在住の日本語学習者の日本語音声进行分析して、IRがどのように実現されているかを観察した。

2. 日本語学習者のIR値

音声収録協力者は、日本滞在経験のない初級から初中級の日本語学習者で、ベトナム語、スペイン語、イタリア語、ドイツ語、スワヒリ語を母語とする5つのグループである。話者数は、各言語それぞれ3名ずつ、計15名である^(注3)。実験対象として日本語（非

外来語）の特殊拍を含まない3モーラ語9語と4モーラ語12語の計21語を選んだ（資料参照）。いずれの語も共通語では非頭高型アクセントを持つ^(注4)。これらの単語をひらがな表記でPCディスプレイ上に提示し（パワーポイント使用）、「○○です」「(○○)」部分がターゲットの単語」というキャリアセンテンスに入れて、学習者に読み上げてもらった。単語を切り替えるタイミングは、話者の自由とした。録音は比較的静かな部屋でICレコーダを使用して行われた。

この読み上げ音声のターゲット語の語頭の第1モーラと第2モーラのピッチの差を計測してIR値とした。この際に、第2モーラの値から第1モーラの値を引いていることから、上記の非頭高型アクセント語の場合には、IR値はプラスの値をとるはずである。ピッチは波形とスペクトログラムを使用して、子音の影響を受けにくい母音定常部を計測した^(注5)。この際、アクセント（核）の位置は無視して、IR値だけを計測したので、結果として共通語アクセントと同じではないアクセント実現も分析対象としている。

また、比較のために同じ語を日本語母語話者3名が読んだ音声も分析対象とした。

*1 国立国語研究所（190-8561 立川市緑町10-2）

*2 東京学芸大学 日本語・日本文学研究講座 日本語学・日本文学分野（184-8501 小金井市貫井北町4-1-1）

読み上げられた音声は、総計で21(語数)×3(話者数)×5(言語数)=315個であるが、この中から言い間違いや、明らかにピッチの抽出ミスを起こしているもの(注6)を除外して、306個のデータを得た。また、日本語話者にも同様の処理を施して、58個のデータを得た。

2. 1 分析結果 (1)

ヒストグラムをみると学習者のIR値は0付近を中心にプラス方向にもマイナス方向にもひろがっており、IRが実現されているとは考えにくい。この分布だけをみれば、学習者は基本的に第1モーラと第2モーラのピッチの変化には無関心で、どちらを高く読むかはランダムに決定していると推察される。

一方、母語話者のIR値は、マイナス値を持たず、+2付近をピークに広くプラス方向に分布しており、IRが実現されていることがわかる(図1, 2参照)(注7)。

2. 2 分析結果 (2)

各言語でのIRの平均値の結果をグラフに示す(ヒゲは標準偏差)。日本語話者のIR値の大きさと比較して、学習者のIR値が小さいことがわかる。また、学習者のなかではベトナム語話者の値が高いこと、スペイン語話者の値が低いこと、ドイツ語話者の標準偏差が大きいこともわかる。このことからドイツ語話者については、話者差が存在する可能性がある(注8)(図3参照)。

日本語も含む6言語の間でこれらのIR値について、

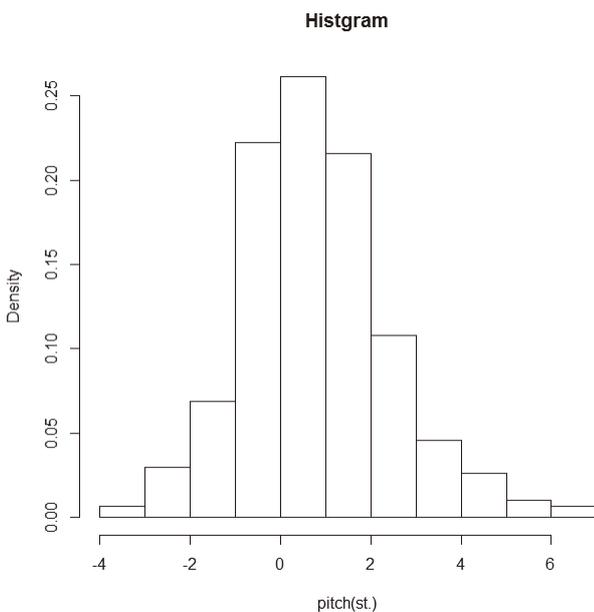


図1 学習者の度数分布

分散分析(不等分散)を行うと各言語間のIR平均値には0.1%水準で有意差がみられたが、学習者でもっともIR値が高かったベトナム語を除く4言語(スペイン語, イタリア語, ドイツ語, スワヒリ語)の話者間では有意差が見られなかった($P=0.1494$)。また、日本語話者とベトナム語話者の間で、 t 検定を行うと0.01%水準で有意差があったことから、母語話者と学習者のIR値には違いがあることと学習者のなかではベトナム語話者のIR値が有意に高い(ただし、母語話

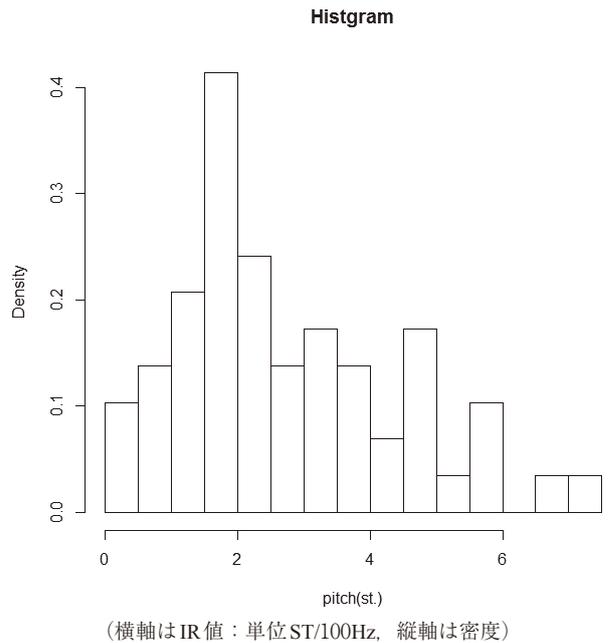
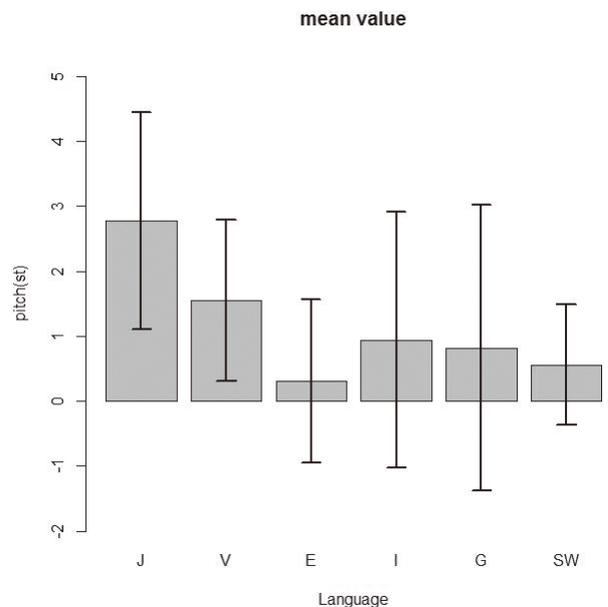


図2 母語話者の度数分布



(縦軸はIR値：単位ST/100Hz, 横軸は言語で記号はJ→日本語, V→ベトナム語, E→スペイン語, I→イタリア語, G→ドイツ語, SW→スワヒリ語の各母語話者を示す)

図3 IR値の平均値と標準偏差

者とは依然として差がある)ことが明らかになった。

3. 知覚実験

次に、学習者の発話した低いIR値を持つ語に対して、日本語母語話者がIRを知覚するのかを明らかにするために実験を行った。

3.1 刺激の選択

上記の21語の中で、ベトナム語^(注9)とスペイン語(各言語とも話者3名)においてIR値の平均値が、最も大きい単語(「やすみ」と最も小さい単語(「ひらがな」)、および中間的なIR値をとる語(「ともだち」)、そして日本語話者の発話でIR値が最大であった語(「おかね」)の4語を選択して聴取実験の対象とした(図4, 表1参照)。

3.2 実験

IRの聞き取りを行ったのは、言語学の知識を持ち、東京式アクセント地域の出身者である日本語話者4名

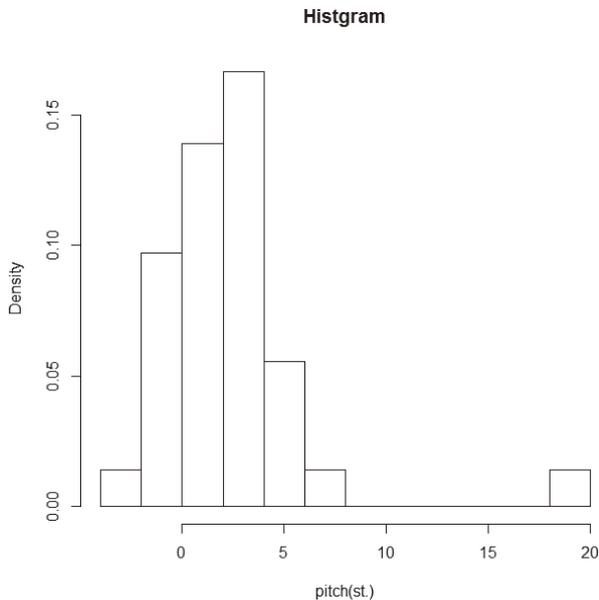


図4 刺激の度数分布

表1 各刺激のIR値

単語	平均	範囲
ひらがな (最小)	0.3	-2.2 ~ 3.0
ともだち (中間)	0.8	-1.6 ~ 3.3
やすみ (最大)	3.1	0.9 ~ 7.1
おかね	4.5	0.0 ~ 18.6

(単位ST/100Hz)

(大学生, 大学院生, 教員, 研究員)である^(注10)。刺激はPsychoPy^[2]を使用して、ノートパソコンからヘッドホンを通じてアトランダムに提示し、IRを感じるか、感じないかを強制二択によって、キーボード上から答えてもらった。単語は約1秒の間隔で二回提示されて、聞き直しは、できなかった。回答するタイミングは被験者の自由とした。実験は2回行われ、上記の語を3人の話者が読み、それを4名に聞いてもらったので、1単語あたり $3 \times 4 \times 2 = 24$ tokenということになる。

3.3 結果 (1)

表2 各母語話者の発話中「IRあり」と判断した発話の比率

	日本語	スペイン語	ベトナム語
IRあり	95.5	43.8	71.9

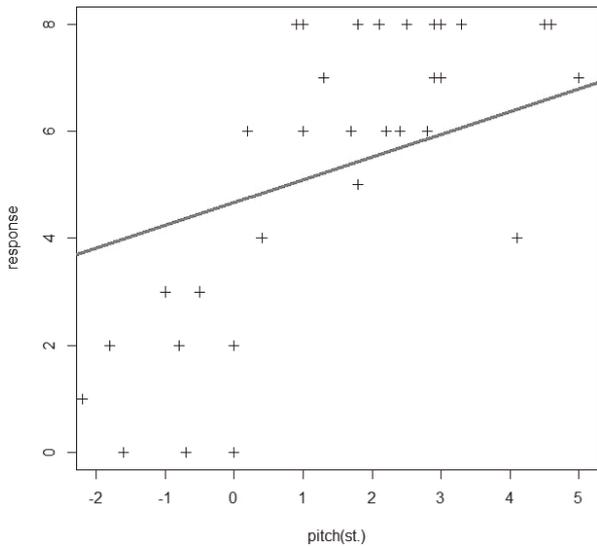
(単位 %)

表2に各母語話者の発話に対して、「IRを感じる」と答えた比率を示す。母語話者のほとんどの発話にはIRが知覚されている。これと比較すると、IR値の低いスペイン語話者の発話ではIRの知覚が可能な発話は、半数以下であり、比較的IR値の高いベトナム語話者の発話でも70%程度で、この比率に関して検定を行うと母語話者の発話とは0.001%水準で有意差があった。

次に刺激のIR値とその刺激に対する「IRあり」という反応との散布図を示す。IR値が増加すると、反応が増加している傾向が明らかである。

そこで、目的変数を「IRありと答えた人数」(response)、説明変数を「IRの値」(pitch)として、相関分析をおこなったところ、二つの変数の間には0.1%水準で有意な相関があり、相関係数は0.55であった。このことからIRの知覚はIR値とやや強い正の相関を持っていることがわかる(図5参照)。

日本語話者に対して、人工語を使い1モーラ目と2モーラ目の高さの違いの範疇境界(どちらが高いと感じるか)を求めた実験では、 $-0.2 \sim 0$ STが境界であるという報告がある^{[3](注11)}。今回の結果でもIR値が0以下の刺激を上昇と聞き取る被験者(つまり「IRあり」と答えた被験者)は半数以下であり、先行研究の結果とおおむね一致する。



(横軸はIR値, 縦軸はIRありと答えた人数)

図5 散布図と相関直線

4. 学習効果

以上では、学習者のIRを観察したが、ここまでの結果からはそもそも学習者は、IRが聞き取れるのだろうか、そして実現できるのだろうかという疑問が生じる。これを確かめるために、上述の学習者の自力による読み上げに引き続いて、同一話者に対して日本語話者（日本語教員）が、同じ単語リストでモデル発話を聞かせた直後に、発音を模倣するという課題を課した。この課題をTask 2と呼び、先ほどの自力での発話をTask 1と呼ぶことにする。学習者全員のIR値の平均値をTask毎に日本語話者の結果とあわせて示す。なお、Task 2においてもTask 1と同様の処理をして、得られたデータ数は、294個であった。

表3 学習者のタスク別のIR値

	日本語	Task1	Task2
平均値	2.77	0.84	1.53

(単位はST/100Hz)

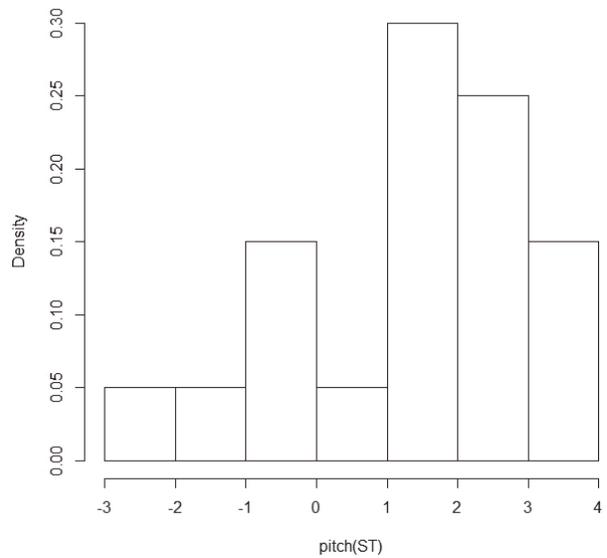
Task 2では、Task 1と比較して大幅にIR値が上昇している。Task 1とTask 2の結果は、0.001%水準で有意差があった。このことは、学習者はIRを聞き取り、再現すること能力を持っていることを示唆する。また、日本語話者とTask 2の比較では、依然として0.001%水準で有意差があった。これは、このTask 2で達成された学習者がIRを再現する能力は、限定的で日本語話者のレベルには達していないことを意味している。

以下では、学習効果の個人差をみるために、Task 1

とTask 2における各単語のIR値の差 (Task 2-Task 1) を計算した結果のヒストグラムを示す。図6はスペイン語話者のひとり (No.3) の結果であるが、負の値もあるもののTask 1からTask 2でIR値が上昇して (つまり、正の値を持つ)、+1以上の差を持っている単語が大部分である。この話者に関しては、IRの再現がある程度うまくいっていると考えてもよいだろう。

一方で、同じスペイン語の別の話者 (No.1) では、正の値は小さく、+1以下がほとんどで、-2以下の大きな負の値も多い (図7参照)。この話者ではIRの

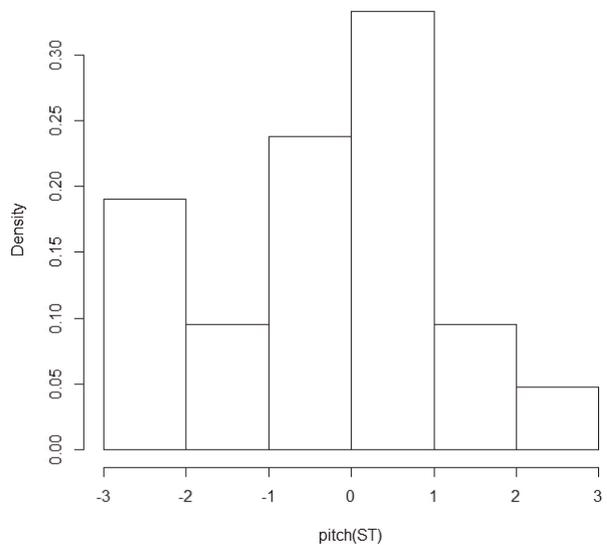
Histogram(No.3)



(横軸はIR値:単位ST/100Hz, 縦軸は密度)

図6 度数分布 (スペイン語 No.3)

Histogram(No.1)



(横軸はIR値:単位ST/100Hz, 縦軸は密度)

図7 度数分布 (スペイン語 No.1)

再現(あるいは聞き取り)が難しかったと考えられる。

このように学習効果に関しては個人差も大きいことがわかった。

5. まとめ

海外在住の日本語学習者が発話した日本語音声の中の、単語の1モーラ目と2モーラ目の高さの違い(IR値)を観測した。その結果、日本語母語話者と比較して、学習者のIR値が小さいことが明らかになった。また、IR値は、スペイン語、イタリア語、ドイツ語、スワヒリ語の話者では、同程度に小さかったが、ベトナム語話者は比較的大きかった。

このIR値の違いは、発音の自然性にも関与していると思われる。上記の知覚実験において、被験者に「日本語話者が発音した音声か」という判断を求めたところ、スペイン語話者の発話の2.1%だけが「日本語話者である」と判断されていたのに対して、ベトナム語話者の発話は25%が日本語話者と判断されていた(日本語母語話者の発話は90.9%)^(注12)。発話の自然性には単音の発音も大きく関与すると考えられるので、断言はできないが、この結果からは、ベトナム語話者のIR値の大きさが発話の自然性を高めた可能性も示唆できる。

なお、ベトナム語学習者のIR値が、他の言語の学習者よりも高かったことが、母語の声調やアクセントの違い、つまり、日本語は自由なピッチアクセントを持ち、ベトナム語は同様にトーンを持っているが、他の言語は、原則的に強弱アクセント(ドイツ語、イタリア語、スペイン語)か固定ピッチアクセント(スワヒリ語)であることに起因しているのかについては、この実験では、明らかにはできなかった。今後の研究課題としたい。

謝辞

ご協力いただいた被験者のみなさまに感謝します。

また、本稿は、外国語発音習得研究会第6回研究集会(2016年12月23日 於 広島修道大学)、および東京音声研究会(2017年6月3日 於 日本大学文理学部)における第一著者がおこなった口頭発表に加筆・訂正を加えたものである。コメントをいただいた方々に感謝します。

注

(注1) 本報告の一部は、2013年度～16年度 文部科学省科学研究費 課題番号25284094 「海外における日本語韻律指導の実践と普及」(代表:磯村一弘)の成果である

(注2) 特殊拍が2モーラ目にある場合などを除く。なお、この音調の音韻論的な取り扱いに関しては、郡^[4]、上野^[5]の各論文を参照

(注3) ピッチのレンジを考慮して、全員女性とした

(注4) 2モーラ語も同時に収録したが、母語アクセントにかかわらず、頭高アクセントを持つことが多かったため、今回の観察対象からは除外した

(注5) Praat ver.5.4.19^[6] 使用

(注6) 抽出ミスに関しては、各話者のIR値の標準偏差の2倍を目安として、それを超える値を持つものを除外した

(注7) 統計処理とグラフ描画に関しては、R ver.3.3.2^[7]を使用した

(注8) 分散分析の結果、ドイツ語とイタリア語については、話者差がみられた

(注9) 音質が低かったため、原音声からノイズ除去を行った

(注10) いずれの被験者も日本語教育歴はない

(注11) ただし、この結果は1モーラ目が2モーラ目より高いか否かという境界を計測しているとも解釈できる

(注12) 被験者1名が日本語話者1名を既知であると報告したので、日本語話者の結果からはこの話者の反応8個を除外した

参考文献

- [1] Hsiang-Yu Lei, 'Acoustic analysis of the allotonic variation of initial rising in Tokyo-style Japanese by native speakers and learners.', Proceedings of 1st International Symposium on Applied Phonetics (ISAPh2016) held at Chubu University. 2016/3/25-28, 2016
- [2] Peirce, JW, "PsychoPy - Psychophysics software in Python.", J Neurosci Methods, 162 (1-2), 8-13, 2007
- [3] Kawai, Goh and Carlos Toshinori Ishi, 'A system for learning the pronunciation of Japanese pitch accent.', Proceedings of EUROSPEECH'99, 1999
- [4] 郡 史郎, 「東京アクセントの特徴再考 —語頭の上昇の扱いについて—」, 『国語学』, 第55巻2号, 16-31, 2004
- [5] 上野 善道, 「アクセントの体系と仕組み」, 『朝倉日本語講座』, 第3巻第4章, 61-84, 2003
- [6] Boersma, Paul and David Weenink, "Praat: doing phonetics by computer.", University of Amsterdam, 2015
- [7] R Core Team, "R: A language and environment for statistical

computing.”, R Foundation for Statistical Computing, Vienna,
Austria, URL <https://www.R-project.org/>, 2016

資料

分析対象語一覧表 (発話順)

1	2	3	4	5
ひらがな	おみやげ	おふろ	おとこ	くだもの
6	7	8	9	10
ななつ	かたかな	てがみ	かみなり	はなし
11	12	13	14	15
とりにく	たまご	こども	おかね	おとし
16	17	18	19	20
いちにち	そのまま	やすみ	ここのつ	ろくがつ
21				
ともだち				