

音声の基本周波数と抑揚が知覚年齢に及ぼす効果

The effects of fundamental frequency and its standard deviation on vocal age perception

堀 紗矢香, 田中 章浩
Sayaka Hori, Akihiro Tanaka

東京女子大学 現代教養学部心理学専攻
Department of Psychology, Tokyo Woman's Christian University

s.kitsch.bell@gmail.com, akihiro@lab.twcu.ac.jp

Abstract

This study examined the effects of fundamental frequency and its standard deviation on vocal age perception. Results of the preliminary experiment showed the variability of perceived age. Acoustic analyses revealed the importance of mean fundamental frequency and its standard deviation when judging the age of the vocal expressions. Perception experiment confirmed the validity of the acoustic manipulation according to the regression equation. These results suggest that vocal age can be perceived by a few simple parameters and that speakers can intentionally alter their own vocal age by changing their own vocal pitch and intonation.

Keywords: Age perception, Voice, Communication

1. はじめに

日常生活において、たとえば電話で初対面の相手とやり取りするときなど、音声のみを手がかりとして年齢を推定する場面がある。このとき、推定された情報は常に正しいとは限らない。わたしたちは相手の声からどの程度正確に年齢を推定できるのだろうか。そして、相手の声のどのような音響的特徴から年齢を推定しているのだろうか。また、そうした音響的特徴を変化させると、知覚される年齢にも変化が生じるのだろうか。

わたしたちは初対面の人と話すときに相手の年齢を推定し、それに応じて会話の内容や言葉遣いを変える。このように、年齢推定はその後のコミュニケーションを方向づける重要な役割を果たしているといえる。年齢推定の研究は主に顔を用いて進められてきた。たとえば Pittenger & Shaw [1]は、大学生が顔写真のみから年齢をかなり正確に推定できることを示している。

日常生活では相手の声だけではなく、顔など他

の情報も利用して相手の年齢を推定していると考えられる。このとき、マガーク効果[2]や腹話術効果[3]に代表されるように、視覚と聴覚から情報が与えられる場合、聴覚情報よりも視覚情報のほうが優先されることが多い。しかし、コミュニケーション場面で顔と声から得られる情報が競合する場合、声からの影響を強く受けることもある。山田・箱田・中村・湯田 [4]は、「活動性」「社会的望ましさ」という2次元のパーソナリティ因子が、顔と声で不一致な属性(高-低)である場合において、聴覚的刺激(声)が人物の印象に強い影響を与えることを示した。年齢推定においても、顔と声の年齢が不一致であるときには顔の情報のほうが重視されるものの、音声による影響も存在することが示されている[5]。

音声による年齢推定の研究として、粕谷らの研究がある[6,7]。彼らの研究では、青年女性(19~31歳)および老年女性(60~76歳)の2群間で音声の音響特徴量(基本周波数、ジッタ、シマ、喉頭雑音)を比較した結果、加齢に伴う平均基本周波数の下降が女性に特徴的であることを示した。彼らの研究は実年齢の変化に伴う音声の変化を検討したものである。しかし、同じ実年齢の者の間でも、音声から推測される年齢が異なることもあるように思われる。そのため本研究ではまず予備調査において、同年齢のさまざまな人物の音声を用いて、それらの知覚年齢にばらつきがあるかどうかを検討する。

次に、知覚される年齢に違いが見られる場合、音声のどのような特徴が知覚年齢を決定する要因

となっているのかを音響分析によって検討する。本研究では発話者が自ら変化させることのできる、音声の平均基本周波数、抑揚、持続時間の3変数に着目し、これらの音響特徴量と知覚年齢の関係について検討する。

さらに知覚年齢に寄与する音響特徴量を実験的に操作することで、同一人物の音声に対する知覚年齢が変化するかどうかを知覚実験によって検討する。

2. 予備調査

予備調査の目的は、音声による知覚年齢のばらつきについて検討することであった。実年齢がほぼ同程度である多数の人物の音声を用い、その音声から知覚される年齢を評定させた。刺激の発話者は女子大学生49名(20-22歳、平均21.0歳)であり、発話内容は中立な台詞「よろしくおねがいします」であった。発話者と面識のない大学生男女53名(18~45歳、平均19.3歳)に、それぞれの音声のみから人物の年齢を推定させた。知覚年齢の分布を図1に示す。知覚年齢の平均21.8歳(SD=1.3)であり、最高で24.1歳、最低で18.9歳であった。したがって、同年齢層の音声であっても、知覚される年齢にはある程度の幅があることが確認できた。

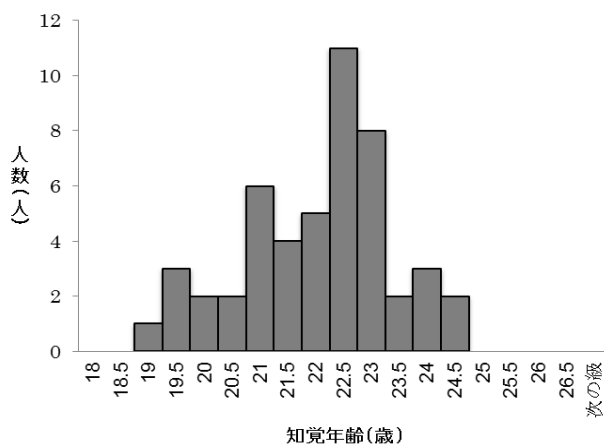


図1 知覚年齢の分布

3. 音響分析

音響分析の目的は、音声のどのような特徴が知覚年齢を決定する要因となっているのかを検討することであった。

予備調査で使用した49名分の発話音声を、目的変数を知覚年齢(perceptual age; 以下PA)、説明変数を平均基本周波数(以下 F_{0mean})、抑揚(以下 F_{0SD})、持続時間の3変数として重回帰分析をおこなった。その結果、 F_{0mean} が低く、 F_{0SD} が大きい音声ほど知覚年齢が高いこと、 F_{0mean} が高く、 F_{0SD} が小さい音声ほど知覚年齢が低くなることがわかった。音声の持続時間は知覚年齢には寄与していなかった。 F_{0mean} と F_{0SD} の2変数を説明変数とした回帰式を式(1)に示す。

$$PA = 26.740 + (-0.025) \times F_{0mean} + 0.115 \times F_{0SD} \quad \dots(1)$$

4. 知覚実験

知覚実験の目的は、知覚年齢に寄与する音響特徴量を実験的に操作することで、同一人物の音声に対する知覚年齢が変化するかどうかを検討することであった。

予備調査で推定された年齢の高低をもとに、声年下、声年相応、声年上の3群を設定し、各群5刺激を選出した。各群の音声の実年齢、知覚年齢、 F_{0mean} 、 F_{0SD} 、持続時間の平均値を表1に示した。これらの音声の F_{0mean} と F_{0SD} を、式(1)をもとに操作し、予備実験のPAより2歳下、1歳下、年相応、1歳上、2歳上に聞こえると予測される音声を作成し、実験で使用した。

表1 各年齢群の音声特徴量

	実年齢(歳)	PA(歳)	F_{0mean} (Hz)	F_{0SD} (Hz)	持続時間(ms)
声年上群	21.0	23.0	249.3	13.7	1260.4
声年相応群	21.0	21.1	260.5	12.2	1298.6
声年下群	21.4	19.6	281.1	11.5	1275.4

実験計画 独立変数は元音声の年齢 3 (年下声, 年相応声, 年上声) × 加工年齢 5 (2 歳下, 1 歳下, 元音声, 1 歳上, 2 歳上) の 2 要因配置であり, 被験者内要因計画とした。従属変数は知覚年齢であった。

実験参加者 参加者は東京女子大学の学生 16 名 (18-22 歳) であった。音声から刺激の人物の特定ができた者はいなかった。

手続き 音声刺激をランダムに 1 発話ずつスピーカーから呈示した。呈示される人物の年齢を音声のみから推定させ, 1 歳刻みに回答させた。

結果 各条件における知覚年齢の結果を図 2 に示す。いずれの年齢群においても, 知覚年齢は 2 歳上, 1 歳上, 年相応, 1 歳下, 2 歳下の順となっており, 加工年齢と知覚年齢に対応が見られる傾向にあった。元音声 3 (年上・年相応・年下) × 加工年齢 5 (2 歳下, 1 歳下, 元音声, 1 歳上, 2 歳上) の 2 要因分散分析をおこなった結果, 年齢群 ($F(2,30)=81.67, p<.001$), 加工年齢 ($F(4,60)=107.65, p<.001$) の主効果, および交互作用 ($F(8,120)=5.97, p<.001$) が有意であった。

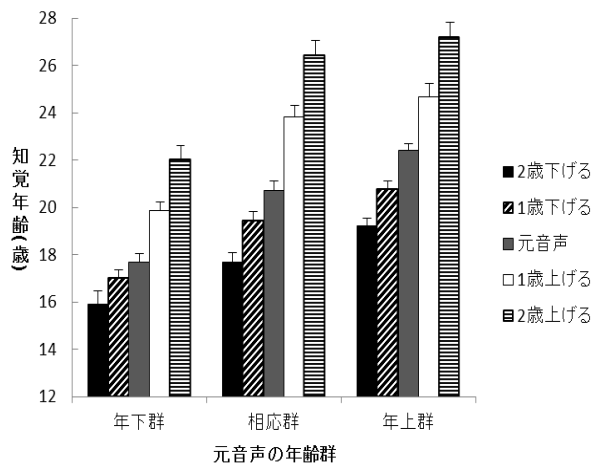


図 2 各群の平均知覚年齢

5. 総合考察

予備調査の結果から, 同年齢でも年下から年上までさまざまな年齢に知覚される声の存在が示された。音響分析の結果, 基本周波数と抑揚が年齢知覚に関与していることが明らかとなった。実際

にこれらの特徴量を操作すると, F_{0mean} を低く, F_{0SD} を大きくするほど実際に知覚される年齢が高くなることが示された。

知覚実験の結果, すべての年齢群において, 年齢を下げる操作より年齢を上げる操作のほうが, 知覚される年齢に大きく影響していた。このような非対称性が示されたことには, モデルの年齢層が関係している可能性がある。今後, モデルの年齢層や性別を変え, 結果の一般性について検討することが望まれる。しかし, 本研究では 20 歳から 22 歳という非常に狭い年齢層のモデルを対象としているにもかかわらず, 2 種類のシンプルな音響特徴量を操作することで知覚年齢を変化させることができた点は興味深い。

顔からの年齢推定の研究では, 顔のさまざまな物理的特徴 (たとえば鼻の長さや肌の肌理など) と知覚年齢の関係が明らかにされてきた。しかし, これらの特徴は意図的に変化させることは難しい。一方で本研究では, 意図的に変化させることが困難な音声の音質 (ジッタ, シマ, 喉頭雑音) の操作は一切おこなわず, 発声者自身による制御が可能である平均基本周波数, 抑揚, 持続時間という比較的単純な変数に着目した。そして, そのなかでも平均基本周波数と抑揚という 2 種類の変数だけで, 年齢の予測と操作が可能であることを示した。この結果は, 話し手が自分自身の知覚年齢を望むとおりに変化させられる可能性を示唆している。このように, 発声者が実際のコミュニケーション場に生かすことができるという点で, 本研究は実用的な示唆をもつといえる。

6. 参考文献

- [1] Pittenger, J. B. & Shaw, R.E. (1975). Perception of relative and absolute age in facial photographs. *Perception and Psychophysics*, 18, 137-143.
- [2] McGurk, H., & MacDonald. (1976). Hearing lips and seeing voices, *Nature*, 264(5588), 746-748.
- [3] Thurlow, W. R., & Jack, C. E. (1973). Certain

determinants of the “ventriloquism” effect,
Perceptual and Motor Skills, 36, 1171-1184.

- [4] 山田奈津子・箱田裕司・中村知靖・湯田恵美子 “顔と声の印象形成における逆マーカー効果の検証” 電子情報通信学会技術研究報告. *HCS*, 99(563), pp.29-33, 2000年1月
- [5] 杉本沙織・平塚真美・高木幸子・田中章浩 “顔年齢及び声年齢の推定” 日本認知科学会・知覚と行動モデリング研究会, 2014年3月
- [6] 粕谷英樹・森大毅・吉田肇・海老原敏 "加齢による声の音響的变化に関する通時的研究" 日本音響学会講演論文集, pp.215-216, 2010年9月
- [7] 粕谷英樹・森大毅・吉田肇・海老原敏 "声の年齢変化に関する一考察" 日本音響学会講演論文集, pp.271-272, 2011年3月