

単母音の個人性知覚における各種音響特徴量の寄与*

北村達也, 齋藤毅 (ATR 認知情報科学研究所)

1 はじめに

音声の個人性知覚では, 声帯音源や声道共鳴などの生得的な身体特徴に起因する情報が重要な役割を果たしている. しかし, これらの情報は発話内容や話者の状態によって変化することが知られている. それにもかかわらず, 人間は話者を識別することができる. この話者内変動に対する個人性知覚の頑健性を生み出すメカニズムを明らかにできれば, 様々な技術的応用が可能である. そこで, 本研究では, 各種音響特徴量の変動が個人性知覚に与える影響を聴取実験によって調査する. 未知話者 10 名の単母音/a/を対象にして, 基本周波数と振幅の時間パターン, 平均基本周波数, 声帯音源の周波数特性, 音声スペクトルの高周波数成分の個人性知覚に対する寄与をサーストンの一対比較法により調査する.

2 実験条件

2.1 刺激音

刺激音は成人男性 10 名が孤立発声した母音/a/をもとに作成した. 収録の際には, 基本周波数の個人差が実験に与える影響を抑えるため, 話者に 700 ms の調波複合音を聞かせ, そのピッチに音声の基本周波数を合わせて持続発声するよう依頼した. 発話継続時間もこの調波複合音に合わせるよう指示した. 発声練習の後, 調波複合音を聞かずに発声した音声から刺激音を作成した. 調波複合音の基本周波数は 123 Hz または 124 Hz であった.

音声は, マイクロフォン (audio-technica AT4041) を用いて, 標本化周波数 48 kHz, 量子化 24 bit にてソリッドステートレコーダ (marantz PMD671) に収録した. その後, パーソナルコンピュータ (PC) で標本化周波数 16 kHz にダウンサンプリングし, 量子化 16 bit に変換した. 収録音声のうち 3 トークンを実験に用いた.

刺激音は以下の 8 種類を用いた. 刺激音 B, C_* , D_* , E は STRAIGHT 分析合成系 [1] を用いて合成した.

刺激音 A 最大振幅を正規化した音声.

刺激音 B 有声区間の基本周波数を一定にし, 振幅パターンを正規化した合成音声.

刺激音 $C_{0.9}$ 刺激音 B の処理に加え, 基本周波数を 0.9 倍した合成音声.

刺激音 $C_{1.1}$ 刺激音 B の処理に加え, 基本周波数を 1.1 倍した合成音声.

刺激音 $D_{1.0}$ 刺激音 B の処理に加え, STRAIGHT スペクトルから求めたケプストラムの 35 次以上を 0 にした合成音声.

刺激音 $D_{0.9}$ 刺激音 $D_{1.0}$ の処理に加え, STRAIGHT スペクトルから求めたケプストラムの 1 次を 0.9 倍した合成音声.

刺激音 $D_{1.1}$ 刺激音 $D_{1.0}$ の処理に加え, STRAIGHT スペクトルから求めたケプストラムの 1 次を 1.1 倍した合成音声.

刺激音 E 刺激音 B の処理に加え, 対数 STRAIGHT スペクトルの 2.6 kHz 以上をその回帰直線で置換した分析合成音声.

ここで, STRAIGHT スペクトルとは STRAIGHT 分析によって得られるスペクトルのことである.

刺激音 B は, 基本周波数と振幅の時間変動パターンの影響を調べるためのものである. まず, 有声区間の各フレームの基本周波数を平均基本周波数で固定し音声を合成した. 次に, 有声区間の振幅を 15 ms ごとに正規化し, 最後に有声区間の始端と終端 100 ms を \cos^2 で重み付けした.

刺激音 $C_{0.9}$ および $C_{1.1}$ は, 平均基本周波数の影響を調べるためのものである. 有声区間の各フレームの基本周波数を平均基本周波数のそれぞれ 0.9 倍, 1.1 倍にした上で音声を合成した.

刺激音 $D_{1.0}$, $D_{0.9}$, $D_{1.1}$ は, 声帯音源の周波数特性の影響を調べるためのものである. これらは, 対数 STRAIGHT スペクトルをフーリエ変換して求めたケプストラム (STRAIGHT ケプストラム) に操作を加えて合成した. 刺激音 $D_{1.0}$ は, STRAIGHT ケプストラムの 35 次以上を 0 にした後, 逆フーリエ変換して得られる STRAIGHT スペクトルを用いて合成した. この処理は, 準同型分析の考えにもとづき高次ケプストラムに現れる声帯音源の情報を取り除くことを意図している. 刺激音 $D_{0.9}$ および $D_{1.1}$ は, さらに STRAIGHT ケプストラムの 1 次をそれぞれ 0.9 倍, 1.1 倍することによって, 声帯音源の周波数特性の傾斜に変化を加えた音声である.

刺激音 E は, 音声スペクトルの高周波数成分の影響を調べるために, 対数 STRAIGHT スペクトルの 2.6 kHz 以上をその回帰直線で置換して合成したもの

* Contribution of acoustic features of sustained vowels on perception of speaker characteristic. by KITA-MURA, Tatsuya and SAITOU, Takeshi (ATR Cognitive Information Science Laboratories)

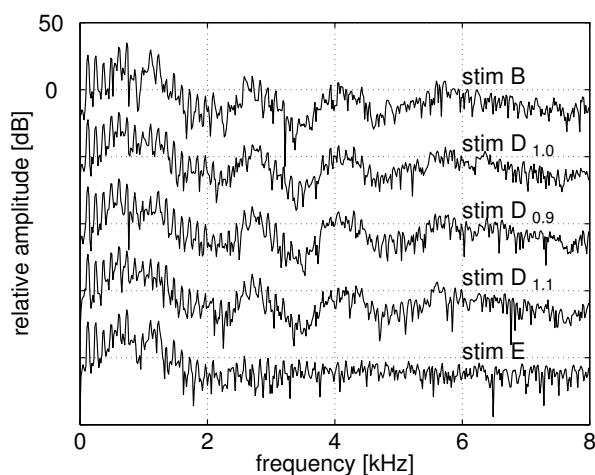


Fig. 1 FFT spectra of stimuli B, D_{1.0}, D_{0.9}, D_{1.1}, and E.

である。この処理によって、対数 STRAIGHT スペクトルの 2.6 kHz 前後に不連続が生じることがあるため、置換後の対数 STRAIGHT スペクトルにローパスフィルタをかけた。2.6 kHz 以上という周波数帯域は、話者 10 名が発声した 5 母音の第 3 ホルマントを含む帯域として設定した。

刺激音 B, D_{1.0}, D_{0.9}, D_{1.1}, E の FFT スペクトルを Fig. 1 に示す。

2.2 実験参加者

正常聴力を有する成人女性 16 名。これらの実験参加者は、刺激音の話者に会ったり、その声を聞いたりしたことがない。つまり、刺激音の話者は実験参加者にとって未知話者である。

2.3 実験方法

サーストンの一対比較法 [2][3] により実験を行った。1 名の音声から作成した 3 つの刺激音を S1, S2, S3 の順序で実験参加者に呈示した。S1 としては常に刺激音 A を呈示し、S2 と S3 にはそれ以外の刺激音を呈示した。そして、S2 と S3 の話者のどちらが S1 の話者に近いかを回答するよう実験参加者に依頼した。

3 つの刺激音は別トークンから作成したものを呈示した。話者 10 名について S1, S2, S3 の全ての組み合わせを作成し、ランダムに実験参加者に呈示した。刺激音の間隔は 500 ms である。順序効果を打ち消すために S1, S3, S2 の順序でも呈示した。全試行数は 420 である。

実験参加者はヘッドホン (Sennheiser HDA200) により刺激音を両耳受聴し、PC のディスプレイ上のボタンをマウスで選択することにより回答を行った。刺激音の聴取は 1 回のみで聴き直しは許さなかった。実験参加者は各々の聴きやすいレベルにて受聴した。

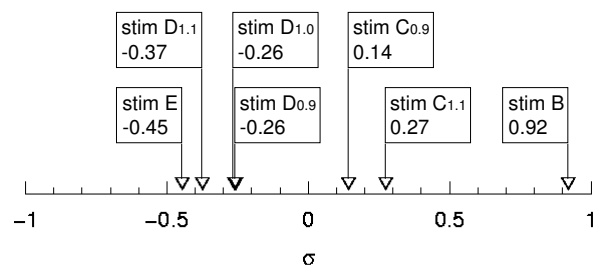


Fig. 2 Interval scale for the stimuli.

3 実験結果と考察

サーストンのケース V を適用して算出した間隔尺度を Fig. 2 に示す。なお、この結果は、サーストンの一対比較法のモデルに適合している ($\chi^2 = 0.712 < \chi^2(15, 0.05) = 24.995$)。

この結果は、刺激音 B, C_{*}, D_{*}, E の順に刺激音 A の話者に近いと判断されたことを示している。すなわち、本研究の実験条件の下では、

1. 音声スペクトルの高周波数成分
2. 声帯音源の周波数特性
3. 平均基本周波数
4. 基本周波数と振幅の時間変動パターン

の順序で個人性の近さの判断に対する影響が大きかったことを示している。このリストは、より上位の特徴量ほど音声生成系の生得的な特徴に依存し、話者が意識的に制御しにくいものであると見ることができる。従って、人間が音声の個人性を知覚する際には、音声中で変動しにくい特徴量をより重視している可能性がある。

4 おわりに

本研究では、音声の個人性知覚における各種音響特徴量の寄与について調査した。成人男性 10 名の単母音 /a/ から作成した刺激音を用いた実験の結果、本研究で対象にした特徴量のうち、音声スペクトルの高周波数成分の寄与が最も高いという結果が得られた。

謝辞 音声データの収録にご協力いただきました東京工科大学メディア学部の飯田朱美先生、石本祐一先生に感謝いたします。本研究は平成 18 年度総務省 SCOPE-R により実施したものである。

参考文献

- [1] Kawahara, Masuda-Katsuse, Cheveigne, Speech Commun., 27, 187–207 (1999).
- [2] 佐藤, 日科技連 (1985).
- [3] 天坂, 長沢, 日本規格協会 (2000).