

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 8 日現在

機関番号：33921

研究種目：挑戦的萌芽研究

研究期間：2012～2014

課題番号：24652087

研究課題名(和文)言葉の鎖をターゲットとした音声の知覚・生成過程の研究

研究課題名(英文)Research for speech perception and production to explore the speech chain

研究代表者

天野 成昭 (Amano, Shigeaki)

愛知淑徳大学・人間情報学部・教授

研究者番号：90396119

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：音声の知覚と生成の処理過程と対応関係の解明に繋がる知見を得ることを目的とし、音声コミュニケーションにおける話し手から聞き手に至る過程を記述した概念モデル「言葉の鎖」を研究の枠組みに据え、音声生成実験・調音運動計測・聴覚神経回路モデルシミュレーション・音声知覚実験を連携させて研究を進めた。その結果、中枢レベルの言語過程においては、表現形式も含めて知覚と生成の情報が極めて良い対応関係にあること、および末梢レベルの生理過程においては、知覚と生成の対応が変動的であり、情報が多重かつ分散した形式で表現されている可能性が示された。

研究成果の概要(英文)：To investigate the processes of speech perception and production and their relationship, research was conducted with perception and production experiments, physiological observations of speech organs, and neural network simulations, which were based on the model of "speech chain". It was revealed that representation of phoneme information was homologous between speech perception and production in a central processing level. However, it is suggested that the representations of phoneme information both in speech perception and production were variable and distributed in peripheral processing level.

研究分野：言語心理学

キーワード：音声 知覚 生成 運動 シミュレーション 音韻

1. 研究開始当初の背景

音声の知覚と生成の処理過程および両者の対応関係に関する研究はこれまでに数多くなされてきた(例えば, Amano & Hirata, 2010; Nagao & de Jong, 2007; Pind, 1995; Yamakawa et al., 2009)。その中において最も有名なのは「言葉の鎖」(図 1)である(cf. Denes & Pinson, 1993)。「言葉の鎖」は音声コミュニケーションにおける話し手から聞き手に至る一連の処理を過程別に記述した概念モデルである。このモデルは音声の知覚と生成の繋がりを明確にし、その全体像を示したことに大きな意義がある。しかし、このモデルは定性的概念モデルに留まり、各過程の特性や処理内容、およびその相互の繋がりについて具体性に欠ける点が問題であった。

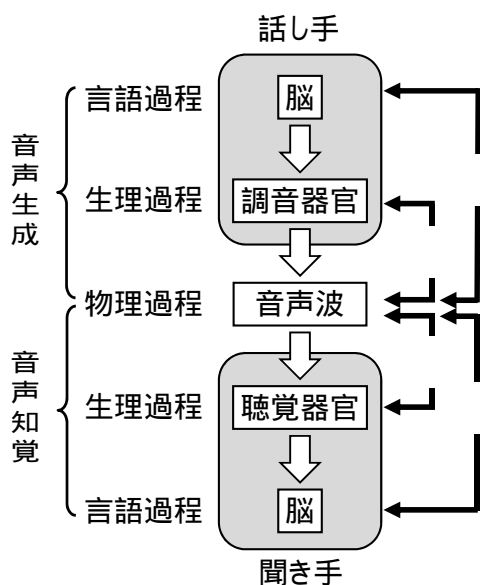


図 1. 言葉の鎖と本研究の対象(~)

2. 研究の目的

音韻の範疇境界と範疇内での変動性を研究の軸とし、音声生成実験・調音運動計測・聴覚神経回路モデルシミュレーション・音声知覚実験を連携して実施することによって、音声生成から音声知覚に至る一連の処理過程を繋いだ「言葉の鎖」の各段階の特性とその相互関係を具体的に明らかにし、音声知覚と音声生成の解明に繋がる知見を得ることを目的とした。

3. 研究の方法

音声生成の言語過程(図 1 の), 音声生成の生理過程(図 1 の), 音声知覚の生理過程(図 1 の), 音声知覚の言語過程(図 1 の) の 4 つの研究項目を立てた。これらの全研究項目における共通の研究対象として、破擦音(ツマ)-摩擦音(スマ), 促音(モット)-非促音(モト)などの対立が明確な音韻対の音声を用いた。その理由は、共通する音声を用いれば、異なるアプローチによる研究項目であって

も、音韻の範疇境界と範疇内での変動性の解析を統一的に実施することが可能となり、かつ研究項目間で得られた結果を容易に比較することが可能となると考えたからである。

(1) 音声生成の言語過程の研究項目では、音声生成実験を実施し、対象とする音韻対を含む単語を埋め込んだ短文を、様々な速度で発声させ録音した。その録音音声の音響特徴を解析し、音韻の生成範疇境界を求めるとともに範疇内での変動性を解析した。

(2) 音声生成の生理過程の研究項目では、上記の音声生成実験において、声帯の運動を、筋電位等の電気的方法およびマイクロファイバースコープ等の光学的方法によって測定し、発声時の生理データを得た。そのデータを基に、対象とする音韻対における調音運動の相異を解析した。

(3) 音声知覚の生理過程の研究項目では、音声生成実験において録音した音声を、聴覚神経回路モデルへ入力し、対象とする音韻対の判別可能性をシミュレーションによって検討した。

(4) 音声知覚の言語過程研究項目では、音声生成実験で録音した音声を刺激とする音声知覚実験を実施し、対象とする音韻対の知覚特性データを得た。そのデータを解析し、知覚範疇境界を求めるとともに範疇内での変動性を解析した。

以上の 4 つの研究項目で得られたデータを総合し、音声知覚と音声生成の各処理過程の特性と内容および対応関係を検討した。

4. 研究成果

(1) 音声生成の言語過程

音声の音響特徴を解析した結果、破擦音-摩擦音の判別には、摩擦部の立ち上がり時間長と残部の時間長の 2 変数が有効であり、これらの変数により破擦音-摩擦音の生成判別境界が表されることが分かった(Yamakawa & Amano, 2012)。また、促音-非促音の判別には、閉鎖に伴う無音区間の時間長とその前後の 2 モーラの時間長の 2 変数が有効であり、これらの変数により促音-非促音の生成判別境界が表されることが分かった(Yamakawa et al., 2012)。一方、対数変換した無音区間の時間長と発声速度との間には単調な直線関係があり(山川ら, 2015)、これらの変数でも促音-非促音の生成判別境界を表せることが分かった。

(2) 音声生成の生理過程

促音および非促音の無音区間において、声帯は閉じた状態から開いていく開大運動の後、開いた状態から閉じていく閉小運動が生じる。促音と非促音の発声時の生理データを解

析し、開大・閉小運動の時間を計測したところ、促音でも非促音でも、開大運動の時間は発声速度に依らずほぼ一定であること、一方閉小運動は発声速度に依存し、速度が遅いほど長くなることが明らかとなった(山川ら, 2015)。これは開大運動が声帯を開放するという受動的運動であり、閉小運動は声帯を閉鎖するという能動的運動であることを示唆している。すなわち、発声速度に応じた適切な無音区間を生成する声帯の開閉時間の制御には、能動的運動である閉小運動が強く関わっていると考えられる。

(3) 音声知覚の生理過程

聴覚神経回路モデルによるシミュレーションの結果、聴覚神経の発火パターンでは、破擦音-摩擦音や促音-非促音などの音韻対をあまり判別できないことが分かった(牧ら, 2013)。すなわち、知覚の言語過程において用いられる変数を、聴覚末梢レベルにおいて直接的に表現できなかった。この原因の1つとして、聴覚末梢レベルにおける時間の区別が数 ms のオーダーであるのに対し、音韻の区別は 100ms のオーダーであることが考えられる。例えば、破擦音-摩擦音の区別に有効な摩擦部の立ち上がりの時間長と残部の時間長は 50ms ~ 100ms のオーダーであり、また促音-非促音の区別に有効な閉鎖区間の時間長は 100ms ~ 500ms のオーダーである。したがって、音韻の区別には、末梢神経レベルの処理よりも上位の中枢神経レベルの処理が深く関わっている可能性が高いと考えられる。

(4) 音声知覚の言語過程

音声知覚実験を行った結果、音声生成の言語過程の研究項目で得られた変数と同一の変数が、音韻の判別および知覚範疇境界の表現に有効であることが分かった。すなわち、破擦音-摩擦音の判別には、摩擦部の立ち上がりの時間長と残部の時間長の2変数が有効であり、促音-非促音の判別には、閉鎖に伴う無音区間の時間長とその前後の2モーラの時間長の2変数が有効であった。さらに、知覚範疇境界と生成範疇境界はほぼ一致することが明らかになった(Amano & Hirata, Accepted)。ただし破擦音-摩擦音において、発声速度の影響を取り除くためには、各変数を平均モーラ長の対数で正規化することが必要であった。

(5) 総合考察

知覚と生成の2つの言語過程において、知覚範疇境界と生成範疇境界に関わる変数が同じであり、かつ両境界がほぼ一致するという結果が得られた。このことから、知覚と生成における音声情報は相同の形式で表現されていると考えられる。つまり、言葉の鎖における始点と終点の処理過程においては、表現形式も含めて知覚と生成の情報が極めて良

い対応関係にあると考えられる。

音韻知覚には末梢の聴覚神経における短い時間の処理だけでは説明できない部分があるという結果は、言葉の鎖における知覚の生理過程が、末梢および中枢神経を含めた多段階の処理過程によって構成されていることを示唆している。しかも、知覚の言語過程において用いられる変数を、聴覚末梢レベルにおいて直接的に表現できなかったという結果は、末梢レベルにおける情報が中枢レベルとは異なる表現形式をとっている可能性を示唆している。そのような表現形式の1つとして異なる時間長を持つ多段階の分散表現が考えられる。

一方、音声生成の生理過程では、無音区間のように音響特徴としては均質に見える場合であっても、それを生成する調音運動が質的に異なる受動的調音運動と能動的調音運動によって構成されているという結果が得られた。これは、言葉の鎖における生成の生理過程において、1つの音響特徴が2つ以上の分散した調音情報として表現される場合があることを意味している。

これらを考え合わせると、言葉の鎖の言語過程(中枢レベル)では、変動性の少ない記号の情報表現がなされ、一方、言葉の鎖の生理過程(末梢レベル)では変動性の大きい分散の情報表現がなされている可能性が示唆される。このような性質は従来の研究でも指摘されてきた。しかし、それをデータに基づいた定量的な観測事実によって証拠づけた点は本研究の意義の1つといえる。

以上のように、言葉の鎖をターゲットとした一連の研究を行い、その各段階における処理過程の特性とその相互関係の一部を具体的に示すことができた。ただし、音声知覚と音声生成の各処理過程とその相互関係には、未解明の部分がまだ多く残されており、今後、さらなる研究が必要である。本研究は、そのような将来の研究に対し、切り口の1つを具体的に示すことができたといえよう。

<引用文献>

- Amano, S. & Hirata, Y. (2010) Perception and production boundaries between single and geminate stops in Japanese, *The Journal of the Acoustical Society of America* 128, 2049-2058.
- Amano, S. & Hirata, Y. (Accepted) Perception and production of singleton and geminate stops in Japanese: Implications for the theory of acoustic invariance, *Phonetica*.
- Denes, P. B. & Pinson, E. N. (1993) *The speech chain: The physics and biology of spoken language*, Second edition, Worth

- Publishers.
- 牧勝弘・山川仁子・榊原健一・天野成昭 (2013) 聴神経発火の計算機シミュレーション: 摩擦音[s]と破擦音[ts]の神経表現, 日本音響学会聴覚研究会資料, (2013年3月), 山梨大学 J-2013-22 43(2), 121-126.
- Nagao, K. & de Jong, K. (2007) Perceptual rate normalization in naturally produced rate-varied speech, The Journal of the Acoustical Society of America 121, 2882-2898.
- Pind, J. (1995) Speaking rate, voice-onset time, and quantity: The search for higher-order invariants for two Icelandic speech cues, Perception & Psychophysics 57, 291-304.
- Yamakawa, K. & Amano, S. (2012) Acoustic features to discriminate affricates and fricatives in Japanese speech corpora, Proceedings of the International Conference on Speech Database and Assessments, December 9-12, Macau, 181-184.
- Yamakawa, K., Amano, S., & Itahashi, S. (2009) Production boundary between fricative and affricate in Japanese and Korean speakers, Proc. INTERSPEECH 2009, 468-471.
- Yamakawa, K., Hirata, Y., & Amano, S. (2012) Discriminant variables for Japanese single and geminate stops in spoken newspaper articles, Abstracts of XXX International Congress of Psychology, July 22-27, Cape Town, 136.
- 山川仁子・榊原健一・今川博・後藤多嘉緒・山内彰人・牧勝弘・天野成昭 (2015) 発声速度に依存した促音の音響的および生理学的特徴, 日本音響学会聴覚研究会資料, (2015年3月) 北海道医療大学 45(2), 67-72.

5. 主な発表論文等

[雑誌論文](計1件)

- 牧勝弘・山川仁子・榊原健一・天野成昭, (2013) 聴神経発火の計算機シミュレーション 摩擦音[s]と破擦音[ts]の神経表現 日本音響学会聴覚研究会資料 Vol.43, 121-126. (査読無).

[学会発表](計4件)

山川仁子・榊原健一・今川博・後藤多嘉緒・山内彰人・牧勝弘・天野成昭, 発声速度に依存した促音の音響的および生理学的特徴, 日本音響学会聴覚研究会, 2015年3月2日, 北海道医療大学(北海道・札幌市)

天野成昭・山川仁子, 破裂音/t/および摩擦音/s/における促音・非促音の生成範疇境界を表す変数, 日本音響学会秋季研究発表会, 2014年9月3-5日, 北海学園大学(北海道・札幌市)

榊原健一・天野成昭・山川仁子・牧勝弘・

今川博, 語頭の摩擦音/s/と破擦音/ts/における喉頭調節への話速の影響, 日本音響学会春季研究発表会, 2014年3月10-12日, 日本大学(東京都・千代田区)

牧勝弘・山川仁子・榊原健一・天野成昭, 聴覚モデルを用いた摩擦音[s]と破擦音[ts]の知覚シミュレーション, 日本音響学会春季研究発表会, 2013年3月13-15日, 東京工科大学(東京都・八王子市)

6. 研究組織

(1) 研究代表者

天野成昭 (AMANO Shigeaki)
愛知淑徳大学・人間情報学部・教授
研究者番号: 90396119

(2) 研究分担者

牧勝弘 (MAKI Katsuhiko)
愛知淑徳大学・人間情報学部・准教授
研究者番号: 50447033

榊原健一 (SAKAKIBARA Kenichi)
北海道医療大学・心理科学部・准教授
研究者番号: 80396168

山川仁子 (YAMAKAWA Kimiko)
愛知淑徳大学・人間情報学部・助教
研究者番号: 80455196