

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年5月17日現在

機関番号：17102

研究種目：基盤研究（B）

研究期間：2008～2012

課題番号：20330152

研究課題名（和文） 音声の耐雑音性を生みだす聴覚特性の研究

研究課題名（英文） How the auditory system makes speech perception noise-tolerant?

研究代表者

上田 和夫 (UEDA KAZUO)

九州大学・芸術工学研究院・デザイン人間科学部門／応用知覚科学研究センター・准教授

研究者番号：80254316

研究成果の概要（和文）：

8 言語からなる多数の文音声サンプルを多変量解析の手法を用いて分析することにより、三つの因子を抽出し、音声における音の強さの変化が、言語によらず、わずか四つの周波数帯域における音の強さの変化で充分、表されることを明らかにした。これらの帯域における音の強さの情報だけを用いた合成音声を用いた実験から、帯域間のパワー変化の時間的順序が、音声知覚のうえで大きな役割を果たしていることなどが分かった。

研究成果の概要（英文）：

Power fluctuations derived from critical-band-filtered spoken sentences in eight languages were analyzed with factor analysis. The analyses consistently yielded three factors and four frequency bands for the languages examined. Perceptual experiments utilizing noise-vocoded speech, in which only power fluctuations in these frequency bands were reflected, revealed that the order of power fluctuations among the bands was one of the important cues in speech perception.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	7,100,000	2,130,000	9,230,000
2009年度	2,000,000	600,000	2,600,000
2010年度	2,000,000	600,000	2,600,000
2011年度	2,000,000	600,000	2,600,000
2012年度	2,400,000	720,000	3,120,000
総計	15,500,000	4,650,000	20,150,000

研究分野：社会科学

科研費の分科・細目：心理学・実験心理学

キーワード：感覚・知覚，音声知覚，明瞭性，パワー変化，因子分析，音素，異聴表，伝達情報量

### 1. 研究開始当初の背景

音声は雑音や歪みに対して極めて頑健であると言われる。たとえば、断続 (Miller, 1950)、フィルタリング (Rosenzweig and Postman, 1957) といった歪みに強く、母語話者では、母音に比べて弱いエネルギーしか持たない子音の弁別正答率が、音声の強さをはるかに上回るような強い雑音が付加された条

件であっても、なかなか低下しないことが確かめられている (Ueda, Akahane-Yamada, Komaki, and Adachi, 2007)。極端な例としては、「これは音声である」という情報が与えられただけで、三つのフォルマント周波数の軌跡を正弦波でなぞった「合成音声」を聴きとることができる (Remez, Rubin, Pisoni, and Carrell, 1981)。音声をいくつかの帯域に分け

てそれぞれの帯域の包絡情報を取りだし、雑音をこの包絡情報で変調した、雑音駆動劣化音声は、もとの音声のスペクトル情報を、非常に大雑把な形でしか含んでいないが、やはり音声を聞き取ることができる (Shannon et al., 1995)。これらは、われわれの聴覚が、いかに乏しい手がかりから音声を聞き取る能力を持っているかということを示した例であると考えられる。

しかし、このような聴覚の復元能力がうまく働かず、音声を十分に聞き取れない場合も存在する。(1) 加齢による聴力低下が生じた場合、(2) 強いバブル雑音や警告音などの妨害音が、信号との分離に役立つような方向手がかりなしに与えられた場合、などがそれにあたる。後者の具体例としては、混雑した列車の駅構内で、発車ベルとアナウンスとが同時に放送された場合などを考えることができる。このような状況において、音声の耐雑音性を向上させる方法が見つかったとすれば、その方法は聴覚の特性をうまく生かしたものであるはずであるし、そのことによって聴覚の特性を解明することにもつながると考えた。

(1) に関しては、まず、基礎的なデータによる裏付けが必要な段階であった。すなわち、加齢により高周波に対する聴力から低下が目立ちはじめると言われているが、日本人の高齢者に対して大規模な聴力測定を行った例はなく、日本人高齢者の聴力特性を規格化するには至っていなかった。また、既存のデータからは、8 kHz 以上の高周波において、高齢者と非高齢者の差が顕著に表れるのではないかと予想されるが、日本において大規模な測定でこの予想を裏付け、規格化の基礎となるデータを得た例はなかった。すでにわれわれの研究グループでは、音声のどの周波数帯域が母音と子音の知覚や音節端点の検出に大きな影響を持つかということ明らかにしつつあったので (Ueda and Nakajima, 2007; Nakajima and Ueda, 2007), さらにこのようなデータがあれば、加齢による聴力低下を補い、音声の明瞭性を復活させる一般的な方法について考える基礎となると考えた。

(2) に関しては、研究分担者の中島らは、この時点までに連続音声の音節境界の自動検出手法を開発し、それに基づいて公共音空間における音声の子音・音節端点強調技術を開発していた (中島, 安武, 2006)。また、鉄道の駅構内などでは、複数の音信号が同時に存在するために、お互いに干渉を起し、明瞭度を下げあうということになりがちであるが、複数の音信号の相互干渉を防ぐ方法について特許出願を行っていた (中島,

上田他, 2006)。

(1) と (2) の両方が該当する状況に関しては、中島らは、高域の聴力が低下した人、あるいは周波数帯域の制限された条件でも使用できる、高周波情報を変換利用して音声の明瞭性を向上させる新しいアルゴリズムを開発した (中島, 上田, 白石, 2007)。しかし、これらのアルゴリズムは「有望なアイディア」ではあったものの、現実的な場面に適用するためには、詳細なパラメータを具体的に検討する必要がある、一般性のある形でそのようなことを行うためには、ヒトの聴覚が耐雑音性を生み出す仕組みについての基礎的な理解を深める必要があった。

## 2. 研究の目的

(1) 高齢者の一般的な聴力特性, (2) 聴覚が音節端点を検出する仕組み, (3) 聴覚が音声の耐雑音性を生み出す仕組み, について明らかにすることを目的とした。

## 3. 研究の方法

(1) 高齢者の一般的な聴力特性を調べるため、健聴者との聴力の差が出やすいと考えられる高周波聴力の測定ができる聴力測定装置を購入し、測定を行った。

(2) 子音・音節端点強調技術を実時間化し、パラメータを最適化するため、雑音の存在する環境下でパラメータを変化させ、異聴の様子から最適なパラメータを調べた。

(3) 聴覚が音節端点を検出する仕組みについて検討するため、大量の多言語音声を収録したデータベースを用いて、臨界帯域フィルターを通した音声のパワー変化について因子分析を行い、多言語音声に共通するパワー変化の因子が見られるかどうかを調べた。さらに、得られた因子の知覚上の意味を調べるため、それぞれの因子の働きを対応する周波数帯域で代表させ、パワー変化の情報を用いて雑音駆動音声を合成し、帯域間のパワー変化を入れ換えたり、混ぜ合わせたり、あるいは特定の帯域のみを欠落させたりし、そのような操作が、文中における音節の知覚にどのような影響をおよぼすのかを調べた。さらに、音素ラベルの付けられたイギリス英語の音声データベースを用いて、同様にパワー変化の因子分析を行い、各因子の因子得点と音素とがどのような対応関係にあるのかを調べた。

## 4. 研究成果

(1) 高齢者の聴力測定を行い、高周波聴力の測定結果の再現性を調べるデータを得た。

(2) 実時間音声強調方式のパラメータを最適化するためのデータが得られた。

(2) イギリス英語, 日本語, フランス語, ドイツ語, 北京語, 広東語, スペイン語の音声データベースを用いて, 自然な連続音声に共通する特徴を明らかにした。これらの音声を臨界帯域フィルターに通し, 音声のパワー変化をフィルター出力ごとに計算して, チャンネル間の相関係数行列を求め, 因子分析を行った。その結果, これら大きく異なる音声言語の間で共通する三つの因子および四つの周波数帯域が存在することが明らかになった。見いだされた周波数帯域は, 音声言語, 発話者の性別による境界周波数の違いが, たかだか 1.5 臨界帯域幅にしかならず, 極めて共通性の高いものであった。このことは, 音声言語を理解するためにヒトが用いている周波数帯域が, これらの音声言語の間で共通するものであることを示唆している。また, ヒトの聴覚は, このように大まかな情報をうまく利用し, 耐雑音性の高い音声知覚を実現していると考えられる。

(3) 音声の因子分析の成果を基に, イギリス英語の音素ラベル付き音声データベースを用いて, 音素ごとの時間的中央点における各因子の因子得点がどのように分布するのかを調べたところ, 音節の核となる母音と, それ以外の子音とがそれぞれ異なる因子の軸に沿って, きれいに分布する様子が見られた。それだけではなく, 音節の開始から終了に至る音素変化の経路上に, 従来の音声学の理論ではうまく扱うことができなかつた/s/を含めて, 母音, 半母音, 子音が音素の「ソノリティー」にしたがって, 矛盾なく分布していることが確認された。このことは聴覚における音素知覚の手がかりを特定するうえで, 極めて大きな意義を持つ。

(4) 雑音駆動音声を用いた知覚実験により, これらの帯域におけるパワー変化情報のうち, 下から 2 番目の帯域の情報が, 特に帯域周波数に対する依存性が強く, 別の帯域のパワー変化と混合したり, 交換したりすることによって, 音声全体の明瞭性が大幅に低下し, ほとんど聞き取りが不可能になることを見いだすことができた。このことから, 音声のパワー変化の情報は, 帯域によって周波数依存性が異なることが明らかとなった。このことは, 音声知覚において, 他の帯域で代替可能な帯域と, そうでない帯域とが存在することを示し, 音声の帯域圧縮の方法を考えるうえで, 大きな意味を持つ。

(5) 音声のパワー変化に対して因子分析を行った研究から得られた 4 周波数帯域を用いて,

日本語 101 単音節から 20 帯域 (臨界帯域) および 4 帯域の雑音駆動音声を合成し, さらに 4 帯域の雑音駆動音声からいずれか一つの帯域を除去した音声を合成した。これらに原音声 (男女各 1 名) を呈示する条件を加えて聴取実験を行った。単音節の同定結果 (異聴表) から, 子音の有声性, 調音方法, 調音位置, および母音の種類について結果をまとめ, 伝達情報量の計算を行った。その結果, 4 帯域のうち, 最も低い周波数帯域を除去することにより, 子音の有声性, 調音方法, 調音位置のいずれにおいても顕著に伝達情報量が低下した。特に, 子音の有声性に対する影響が甚だしかった。このことは, 有声子音と無声子音の区別は, 声帯波の出現する時点と破裂の出現する時点との時間差 (有声開始時間) が手がかりとは限らず, 最も低い周波数帯域とそれ以外の帯域との間で生ずる立ち上がりの時間差が手がかりとなり得ることを示している。調音方法や調音位置の知覚に関しても, 破裂等の物理的手がかりが存在する高い周波数帯域よりも, 低い周波数帯域と他の周波数帯域との相対的な時間的關係が重要であることを示していると考えられる。下から二番目の帯域を除去することにより, 母音の種類と同定に関する伝達情報量が顕著に低下した。このことは下から二番目の周波数帯域とその上下の周波数帯域との間で生ずるパワーの時間的変化が母音の知覚にとって重要であることを示している。このような, 帯域間のパワー変化の時間的順序という, 雑音の影響を受けにくい手がかりを利用して, 音素の区別をすることが可能であることが, 音声知覚に高い耐雑音性が見られる理由の一つであろう。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 18 件)

- ① Yuko Yamashita, Yoshitaka Nakajima, Kazuo Ueda, Yohko Shimada, David Hirsh, Takeharu Seno, and Benjamin Alexander Smith. (2013). "Acoustic analyses of speech sounds and rhythms in Japanese- and English-learning infants," *Frontiers in Psychology*, 4: 57, 1-10. doi: 10.3389/fpsyg.2013.00057 (査読有)
- ② Emi Hasuo, Yoshitaka Nakajima, and Kazuo Ueda. (2011). "Does filled duration illusion occur for very short time intervals?" *Acoustical Science and Technology*, vol. 32, no. 2, 82-85. (Letter) (査読有)
- ③ Takayuki Sasaki, Yoshitaka Nakajima, Gert ten Hoopen, Edwin van Buringen, Bob

Massier, Taku Kojo, Tsuyoshi Kuroda, and Kazuo Ueda. (2010). "Time-stretching: Illusory lengthening of filled auditory durations," Attention, Perception, and Psychophysics, vol. 72, 1404-1421. (査読有)

- ④ 坂田俊文, 上野哲子, 一番ヶ瀬崇, 白石君男, 中川尚志. (2009). 無難聴性耳鳴症における高周波数領域の聴力像と耳鳴ピッチ, *Audiology Japan*, 52, 220-226. (査読有)
- ⑤ 中島祥好. (2009). 聴覚におけるリズム知覚, *月刊言語*, 38, 66-73. (査読なし)

[学会発表] (計 77 件)

- ① Kazuo Ueda, Yoshitaka Nakajima, Kiyoto Noguchi, and Yuichi Satsukawa (2012-10-20). "Perceptual roles of different frequency bands in Japanese syllable identification," *Fechner Day 2012, the 28th Annual Meeting of the International Society for Psychophysics*, 62-67.
- ② Yoshitaka Nakajima, Kazuo Ueda, Shota Fujimaru, Hirotohi Motomura, and Yuki Ohsaka (2012-10-20). "Acoustical correlate of phonological sonority in British English," *Fechner Day 2012, the 28th Annual Meeting of the International Society for Psychophysics*, 56-61.
- ③ 上田和夫. (2012-03-25). "人は音声の何を聴いているのか," 京都大学文学部心理学教室読書会, 講演会 (招待講演)
- ④ 古川歩, 中島祥好, 上田和夫, 安武達朗, 藤丸翔太, 本村博利. (2012-03-13). "準実時間子音強調システムの実用化に向けた研究," 日本音響学会春季研究発表会 1-10-6, 講演論文集, pp. 515-516.
- ⑤ 堂本香那, 上田和夫, 中島祥好, Wolfgang Ellermeier, Florian Kattner. (2012-03-13). "無関連な雑音駆動音声の妨害効果," 日本音響学会春季研究発表会 1-10-7, 講演論文集, pp. 517-518.
- ⑥ Kazuo Ueda, Yoshitaka Nakajima, and Yuichi Satsukawa. (2010-10-19). "Effects of frequency-band elimination on syllable identification of Japanese noise-vocoded speech: Analysis of confusion matrices," *Fechner Day 2010, Padova, Italy, Proceedings of the 26th Annual Meeting of the International Society for Psychophysics*, 39-44.
- ⑦ Kazuo Ueda, Tomoya Araki, and Yoshitaka Nakajima. (2009-10-26). "The effect of amplitude envelope coherence across frequency bands on the quality of noise-vocoded speech," *EURONOISE 2009*,

Edinburgh, Scotland, *Acta Acustica united with Acustica*, 95, Suppl. 1, S107 (invited paper).

- ⑧ Kazuo Ueda and Yoshitaka Nakajima. (2008-08-01). "A critical-band- filter analysis of Japanese speech sentences," *Proceedings of the 24th Annual Meeting of the International Society for Psychophysics*, Toronto, Canada, 249-252.
- ⑨ Kazuo Ueda and Yoshitaka Nakajima. (2008-07-01). "Factor analyses of critical-band-filtered speech of British English and Japanese," *Acoustics'08 Paris*, Paris, France, *J. Acoust. Soc. Am.*, 123, 3163 (invited paper).

[図書] (計 1 件)

- ① 白石君男. (2010). オープンフィッティングとハウリング制御: 軽度難聴者の補聴器装用のために, 細井裕司編集, ENTONI 高齢者の補聴: 実地診療に役立つ最新の知識, 全日本病院出版会, 8 ページ.

[産業財産権]

○取得状況 (計 4 件)

名称: 放送装置  
発明者: 中島祥好, 上田和夫, 河原一彦, 安村真奈  
権利者: 国立大学法人九州大学  
種類: 特許  
番号: 特許第 5124768 号  
取得年月日: 24 年 11 月 9 日  
国内外の別: 国内

名称: 音声強調処理装置  
発明者: 中島祥好, 上田和夫, 白石君男  
権利者: 国立大学法人九州大学  
種類: 特許  
番号: 特許第 5046233 号  
取得年月日: 24 年 7 月 27 日  
国内外の別: 国内

名称: 音声信号強調装置  
発明者: 中島祥好, 上田和夫  
権利者: 国立大学法人九州大学  
種類: 特許  
番号: 特許第 5115818 号  
取得年月日: 24 年 10 月 26 日  
国内外の別: 国内

名称: 子音加工装置, 音声情報伝達装置及び子音加工方法  
発明者: 中島祥好, 安武達朗  
権利者: 国立大学法人九州大学

種類：特許  
番号：特許第 4876245 号  
取得年月日：23 年 12 月 9 日  
国内外の別：国内

〔その他〕

ホームページ等

<http://hyoka.ofc.kyushu-u.ac.jp/search/details/KO02356/research.html#21407338793554225991>

## 6. 研究組織

### (1)研究代表者

上田 和夫 (UEDA KAZUO)

九州大学・大学院芸術工学研究院・准教授  
研究者番号：80254316

### (2)研究分担者

中島 祥好 (NAKAJIMA YOSHITAKA)

九州大学・大学院芸術工学研究院・教授  
研究者番号：90127267

白石 君男 (SHIRAIISHI KIMIO)

九州大学・大学院芸術工学研究院・教授  
研究者番号：90187518