

## 科学研究費助成事業 研究成果報告書

平成 30 年 6 月 15 日現在

機関番号：62618

研究種目：基盤研究(B) (一般)

研究期間：2014～2017

課題番号：26284062

研究課題名(和文) 自発音声コーパスの分析によるfilled pauseの音声学的特徴の解明

研究課題名(英文) Clarification of the phonetic characteristics of filled pauses based upon the analysis of spontaneous speech corpus

研究代表者

前川 喜久雄 (Maekawa, Kikuo)

大学共同利用機関法人人間文化研究機構国立国語研究所・音声言語研究領域・教授

研究者番号：20173693

交付決定額(研究期間全体)：(直接経費) 10,500,000円

研究成果の概要(和文)：研究計画にそって日本語・英語・中国語のフィラーを音声学的に研究した。日本語については、フィラー中の母音の声質が通常の単語中の母音とは組織的に異なることを示した。英語については、節頭のフィラーの生起確率に影響を及ぼす言語的要因の解明を進め、また、日本人英語学習者のフィラーを英語ネイティブのフィラーと比較し、学習レベルによる変化の進み方を解明した。中国語については、上記と同様の分析を行い、日本語と同一の結果を得たが、母音音質の変化の方向等には言語に依存した差異も確認した。最後に、声質の著しい変化を伴うフィラー音声から音声基本周波数(F0)情報を安定して抽出できる計算アルゴリズムを開発した。

研究成果の概要(英文)：The following research was conducted as planned in the research proposal. As for Japanese, 1) Systematic difference of voice quality was found between the vowels in filled pauses and ordinary lexical items. As for English, 2) factors affecting the probability of filled pauses in clause-initial position were extracted, and 3) differences of filled-pause usage between the English native speakers and Japanese learners of English were extracted, and analyzed with respect to proficiency levels of the learners. As for Chinese, 4) the same analysis as in 1) above revealed basically the same results as in Japanese, but language-dependent differences were also found with respect to the change in the vowel color, for example. Lastly, 5) a new computational algorithm that can enable reliable extraction of the speech fundamental frequency (F0) information from the vowels having strongly creaky and/or breathy voice qualities.

研究分野：音声学

キーワード：フィラー 自発音声 F0抽出 日本語 英語 中国語 声質

### 1. 研究開始当初の背景

Filled pause (以下、フィラー) とは「えー」「あー」等の言い淀みのことである。フィラーは音声コミュニケーション上重要な役割を果たしていることが近年広く認識されるようになってきたが、本来無意識的に生成される要素であるために、実験的環境での分析が困難なことが原因となって、これまで定量的な研究がほとんど行われてきていない。近年整備が進んだ自発音声コーパスを利用することでフィラーの音声学的・言語学的研究を進捗させることが、言語研究のみならず、言語教育においても期待されている。

### 2. 研究の目的

本研究のもっとも中心的な目的は自発音声の分析によって、日本語等におけるフィラーの音声的特徴と生起に関わる要因を解明し、フィラーの生成メカニズムを解明することである。ここで生成メカニズムという場合、末梢系における狭義の音声生成メカニズム(フィラーの音声特徴が音声器官によってどのように生成されるか)とともに、中枢系におけるフィラー生成のプランニング(話者が発話中のどのような環境でフィラーを生成しているか)をも意味している。

また本研究の副次的目標として、日本語の分析によって明らかになったフィラーの特徴の一部を他の言語(具体的には英語と中国語)と比較することによって対照言語学的な分析を行うことを予定している。

生成メカニズムについての検討では、フィラーの生起がどの程度まで予測可能であるかを検討する。また副次的目標として、一部のフィラーについて、英語および中国語との対照研究を試みる。

### 3. 研究の方法

日本語の自発音声进行分析するためには、『日本語話し言葉コーパス』のモノローグデータを利用した。英語については、研究分担者の渡辺が独自に構築した英語モノローグコーパス(『日本語話し言葉コーパス』と比較可能な設計となっている)、および研究分担者の Rose が独自に構築した日英両語に関する学習者コーパスを利用し、さらに米国で構築された Santa Barbara Corpus も活用した。中国語に関しては台湾の中央研究院言語学研究所で開発された MCDC8 コーパスを利用した。

フィラーの音声特徴の分析では、通常の音声分析で実施されることの多い、持続時間、音声基本周波数(F0)情報、スペクトル特徴に加えて、声帯の振動様式の差にかかわる発声様式(phonation type)の特徴や、音声信号の時間的ゆらぎ(shimmer, jitter)や自己相関値、信号ノイズ比など、言葉の意味の対立には直接関係しないが、非言語的メッセージの伝達には関与していると考えられる音声特

徴、すなわち声質(voice quality)の特徴を積極的に分析する。具体的には『日本語話し言葉コーパス』コアから抽出した約 20 万回の母音について、上述の音声特徴を含む 15 種類の音声特徴を計算したデータベースを構築し、それを分析した。これと同一仕様のデータベースを中国語コーパスからも構築して対照研究に活用した。

なおフィラー中の母音では、声帯の振動様式がきしみ声(creaky voice)や息漏れ声(breathy voice)となるものがあり、F0 抽出が困難になることが知られている。そのため、通常とは異なる発声様式の音声から安定して F0 情報を抽出することのできる計算アルゴリズムの開発も上記の研究と並行して実施した。

### 4. 研究成果

まず声質については、日本語の場合、フィラー中の母音と通常の単語に含まれる母音、例えば「えー」と「英語」の語頭に位置する長母音の[e:]の音声特徴を比較すると、両者間に明瞭な差が存在することを確認した。差が最も顕著であったのは持続時間、インテンシティ、F0 の三者であったが、他にスペクトル傾斜に関わる H1\*-A1\*、F0 の時間ゆらぎ、第 1 フォルマント周波数(F1)などにも統計的に有意な差が確認できた。図に母音/a/における H1\*-A1\*の分布の差を示す。箱ひげ図のうち左半分が男性、右半分が女性であり、白い箱は通常の単語、網掛けされた箱はフィラーに含まれる/a/を示している。男女ともフィラー中の母音の方がスペクトルの傾斜が大きいことがわかる。これは聴覚的にはフィラーの方がかすれた声で発声されている可能性を示唆する結果である。

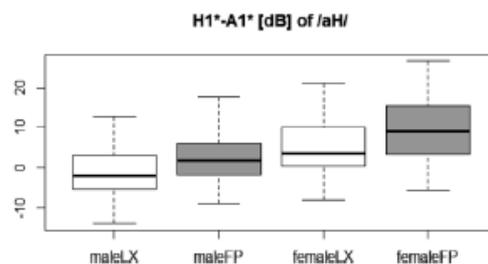


図1. 通常単語とフィラー中の/a/における H1\*-A1\*の分布

こうした音声特徴の差がどの程度までフィラーと特徴づけるかを確認するために、ランダムフォレスト法を用いた交差検証を実施したところ、全ての音響特徴量を利用した場合で 0.86 以上、持続時間情報を除外した場合でも 0.78 以上の F 値が得られた。この結果は、音声特徴量だけによって、与えられた母音がフィラー中のものか通常の単語中のものかを高い確率で予測できることを意味しており、フィラーと通常の単語とでは狭義の音声生成メカニズムに組織的な差が存

在することを強く示唆している。

これと同一の分析を中国語にも実施したところ、結果は基本的には日本語と同一であり、フィラーと通常単語中の母音には組織的な音響上の差が存在していることが確認できた。しかし中国語と日本語では、顕著な差を示す音響特徴は必ずしも同一ではなく、フィラーを通常単語から区別する音声特徴のあり方には言語差も存在することを示唆する結果であった。

なお、中国語についての研究協力者である台湾中央研究院のTsengは本研究を契機に中国語のフィラーの新しい分類法を提案するに至った。本研究の副次的成果として中国語学にも貢献できたことになる。

次に英語モノログにおけるフィラーの生起については、節(clause)を領域として観察した場合は、英語日本語ともに節を構成する単語数が増加するにつれて、節頭フィラーの生起率も上昇することが確認できた。一方、統語境界の強さの影響に関しては、英語では節頭よりも文頭の方が生起率が高いことが判明した。これは従来日本語について報告されている結果とは一致しない結果であり、発話プランニングの在り方に言語差が存在する可能性を示唆する発見となった。

ところで英語のフィラーには、“uh”のように日本語のフィラーと同様に母音だけから構成されるもの以外に“um”のように子音で終わるタイプのものもある。日本人学習者が発した英語音声の中のフィラーを分析すると、学習者は母音タイプのもを著しく好んで発していることが判明した。またフィラーの音響特徴の分析からは、学習者の発するフィラー中の母音はネイティブ話者の母音よりも第1フォルマント周波数が低い傾向にあることも判明した。これらの発見は、日本人学習者の英語におけるフィラー使用の指導に関する指標を与えるとともに、将来的には日本人英語における音声上の問題点を自動検出するための指標としても活用できる可能性をもっていると考えられた。

最後に、きしみ声や息漏れ声で発声された音声から安定してF0を抽出する手法の開発では、F0(音声基本周波数)の定義にまで遡って新しいアルゴリズムを検討した。その結果、音声信号の非周期性(aperiodicity)と瞬時周波数(instantaneous frequency)に対して新しい時間周波数表現を考案することによって、F0だけでなく非周期性に関する情報も同時に提供する手法を考案することができた。この手法はF0に対して複数の候補が存在することを前提としたものであり、複数の候補の確率マップを出力する。利用者は確率分布を参照することで、対象とする音声信号がどの程度安定したF0情報を含んでいるかを定量的に評価することが可能になる。

図2にきしみ声で発声された日本語フィラー「えー」の分析例を示す。このフィラーは前半が著しいきしみ発声で生成されてい

るが、後半はノーマルな発声に復帰している。図の下半分に示されている確率マップを参照すると、きしみ発声に対応する左半分では多くの候補が競合しているのに対し、ノーマル発声に対応する右半分では候補の数が限られている。これによってこの音声サンプルでは前半に非ノーマルな発声が行われているとの推測が可能である。

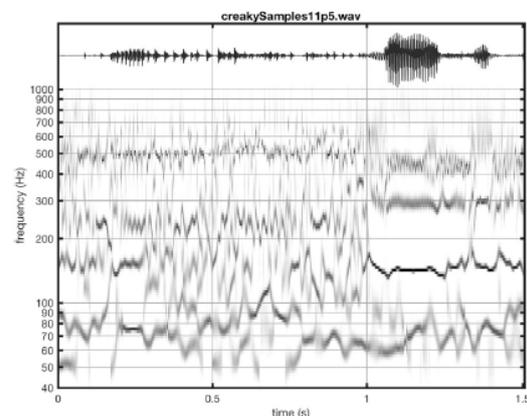


図2. 新手法によるF0確率マップの一例

## 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5件)

Kikuo Maekawa and Hiroki Mori. “Comparison of voice quality between the vowels in filled pauses and ordinary lexical items.” *Journal of Phonetics Society of Japan (Onsei Kenkyu)*, 21 (3), pp. 53-62, 2017.  
[https://doi.org/10.24467/onseikenkyu.21.3\\_53](https://doi.org/10.24467/onseikenkyu.21.3_53)

Michiko WATANABE and Yusaku KOREMATSU. “Factors Affecting Clause-Initial Filler Probability in an English Monologue Corpus.” *Journal of Phonetics Society of Japan (Onsei Kenkyu)*, 21 (3), pp. 24-32, 2017.  
[https://doi.org/10.24467/onseikenkyu.21.3\\_24](https://doi.org/10.24467/onseikenkyu.21.3_24)

Ralph Rose. “A Comparison of Form and Temporal Characteristics of Filled Pauses in L1 Japanese and L2 English.” *Journal of Phonetics Society of Japan (Onsei Kenkyu)*, 21 (3), pp. 33-40, 2017.  
[https://doi.org/10.24467/onseikenkyu.21.3\\_33](https://doi.org/10.24467/onseikenkyu.21.3_33)

Hideki Kawahara. “Application of Time-frequency Representations of Aperiodicity and Instantaneous Frequency for Detailed Analysis of Filled Pauses.” *Journal of Phonetics Society of Japan (Onsei Kenkyu)*, 21 (3), pp. 63-73, 2017.  
[https://doi.org/10.24467/onseikenkyu.21.3\\_63](https://doi.org/10.24467/onseikenkyu.21.3_63)

Shu-Chuan Tseng. “Chinese Demonstratives and Their Spoken Forms in a Conversational Corpus.” *Journal of Phonetics Society of Japan (Onsei Kenkyu)*, 21 (3), pp. 41-52, 2017.  
[https://doi.org/10.24467/onseikenkyu.21.3\\_41](https://doi.org/10.24467/onseikenkyu.21.3_41)

※すべて査読有・オープンアクセス(J-STAGE)

[学会発表] (計 9 件)

Kikuo Maekawa, Ken'ya Nishikawa, and Shu-Chuan Tseng. "Phonetic characteristics of filled pauses: a preliminary comparison between Japanese and Chinese." *Proc. Disfluency in Spontaneous Speech (DiSS2017)*, Stockhol, August 2017. 査読有

前川喜久雄・森大毅「日本語フィラーの声質分析」第 334 回日本音声学学会研究例会シンポジウム「フィラーの音声学と言語学：日英中を対象に」十文字学園女子大学, 2016.12.03.

渡辺美知子「英語と日本語のインフォーマルなスピーチにおけるフィラーの頻度の比較」第 334 回日本音声学学会研究例会シンポジウム「フィラーの音声学と言語学：日英中を対象に」十文字学園女子大学, 2016.12.03.

Ralph Rose "Realtime detection of filled pauses and temporal variables for fluency development." 第 334 回日本音声学学会研究例会シンポジウム「フィラーの音声学と言語学：日英中を対象に」十文字学園女子大学, 2016.12.03.

河原英紀「基本周波数および非周期性の詳細な解析手法による CSJ コーパスの分析について」第 334 回日本音声学学会研究例会シンポジウム「フィラーの音声学と言語学：日英中を対象に」十文字学園女子大学, 2016.12.03.

Shu-Chuan Tseng. "Filling pauses in Chinese conversation" 第 334 回日本音声学学会研究例会シンポジウム「フィラーの音声学と言語学：日英中を対象に」十文字学園女子大学, 2016.12.03

Kikuo Maekawa and Hiroki Mori. "Voice-quality difference between the vowels of filled pauses and ordinary lexical items". *Proc. INTERSPEECH 2016*, pp.3171-3175, San Francisco, 2016.09.12. 査読有

Kikuo Maekawa and Hiroki Mori. "Voice-quality analysis of Japanese filled pauses: A preliminary report". *DiSS 2015 (Disfluency in Spontaneous Speech)*, Edinburgh, 2015:08. 査読有

Michiko Watanabe, Yosuke Kashiwagi, and Kikuo Maekawa. "The relationship between preceding clause type, subsequent clause length and the duration of silent and filled pauses at clause boundaries in Japanese monologues". *DiSS 2015 (Disfluency in Spontaneous Speech)*, Edinburgh, 2015:08. 査読有

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称：  
発明者：  
権利者：  
種類：

番号：  
出願年月日：  
国内外の別：

○取得状況 (計 0 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年月日：  
国内外の別：

[その他]  
ホームページ等

6. 研究組織

(1) 研究代表者

前川 喜久雄 (MAEKAWA, Kikuo)  
国立国語研究所・音声言語研究領域・教授  
研究者番号：20173693

(2) 研究分担者

森 大毅 (MORI, Hiroki)  
宇都宮大学・工学 (系) 研究科 (研究院)・  
准教授  
研究者番号：10302184

渡辺 美知子 (WATANABE, Michiko)  
国立国語研究所・コーパス開発センター・  
プロジェクト非常勤研究員  
研究者番号：60470027

Ralph Rose ( )  
早稲田大学・理工学術院・准教授  
研究者番号：30404916

河原 英紀 (KAWAHARA, Hideki)  
和歌山大学・システム工学部・名誉教授  
研究者番号：40294300

(3) 研究協力者

Shu-Chuan Tseng ( )  
中央研究院(台湾)言語学研究所・Research  
fellow