

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 11 日現在

機関番号：62618

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24520494

研究課題名(和文)日英語話し言葉コーパスにおける言い淀み分類の精緻化と機能の対照分析

研究課題名(英文)Contrastive analysis of English and Japanese disfluencies using parallel spontaneous speech corpora of the two languages

研究代表者

渡辺 美知子(Watanabe, Michiko)

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・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

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：文や節の境界におけるポーズやフィラーの持続時間は、境界の深さだけでなく後続発話生成のための認知的負荷の大きさの影響を受けると考えられる。そこで、これらの境界におけるポーズやフィラーの持続時間と、境界の深さや後続節長との関係を『日本語話し言葉コーパス』を用いて調べた。ポーズ長には境界の深さや後続節長との対心が見られたが、フィラー長にはこれらの要因との関係はほとんど観察されなかった。一方、ポーズが長くなるほど後続するフィラーの出現率は上昇した。これらの結果から、文や節の境界における後続発話生成のための認知的負荷はポーズ長に反映され、ポーズ長が間接的にフィラーの出現率に影響することが示唆された。

研究成果の概要(英文)：I investigated the relationship between the use of silent and filled pauses and the speaker's cognitive load of speech production at deep syntactic boundaries. More specifically, I examined the effects of boundary depth and the subsequent clause length on the durations of silent and filled pauses at the boundaries, using "The Corpus of Spontaneous Japanese". The results revealed that silent pause duration increased with the boundary depth and the subsequent clause length, but not filled pause duration. On the other hand, filled pause rate increased with preceding silent pause durations. These results indicate that speaker's cognitive load of speech production affects silent pause durations, and that long silent pauses trigger additional filled pauses.

研究分野：音声科学

キーワード：非流暢性 言い淀み ポーズ フィラー 発話生成モデル

### 1. 研究開始当初の背景

自発発話に固有の言い誤りや言い淀みは、心理言語学分野では発話生成プロセス探求の手がかりとして早くから着目されてきた。伝統的な言語学分野では、言い淀みは発話の命題的意味に関与しないことから、研究に値しないものとして長年無視されてきた。しかし、会話分析や談話分析の進展に伴い、言い淀みがコミュニケーションにおいて果たしている積極的役割が注目されるようになってきた。

研究代表者は、これまで『日本語話し言葉コーパス (CSJ)』(国立国語研究所, 2006) を用いて、文境界や節境界におけるフィラーの分布と働きについて調べてきた。そして、文境界や節境界のフィラーの出現率が、統語的な境界の深さや後続節中の語数と関連していることを明らかにし、フィラーが後続発話のプランニングに関連した現象であることを指摘した。また、フィラーの種類を選択には、発話のタイプだけでなく、話者の性別や年齢が関係していることも見出した。さらに、既存のコーパスを用いて日本語と英語の言い淀みを比較した結果、フィラーの出現率に顕著な差があり、日本語の方が高いこと(日本語: 5.9%, 英語: 1.6%), 一方、語や語の一部の繰り返し出現率は英語の方が高いこと(日本語: 0.13%, 英語: 1.4%)などが明らかになった。しかし、これまで分析に用いてきた既存のコーパスでは、発話のタイプ、即興性、話者特性の統制がなされていないため、これらの値を単純に比較し、日本語話者の方が言い淀む頻度が高いといった結論を出すことには問題がある。そこで、既存のコーパスを用いた研究を進めると同時に、話者特性や発話場面の統制された日英語話し言葉コーパスを新たに構築し、それを利用した日英語言い淀みの対照研究を行いたいと考えた。

### 2. 研究の目的

日本語と英語の話し言葉コーパスにおける言い淀みの分布や特徴を調べることによって、両言語における発話生成プロセスの共通点と相違点を探る。そのために、既存のコーパスを分析するだけでなく、話者特性や発話場面の統制された日英語話し言葉コーパスを新たに構築し、精度の高い対照研究を目指す。本研究は日英語の話し言葉の特徴の一端を明らかにするものであり、外国語としての日本語教育や英語教育に貢献する。

### 3. 研究の方法

『日本語話し言葉コーパス (CSJ)』など、既存のコーパスを用いて言い淀みの分析を進めると同時に、話者特性や発話場면을統制した日英語の話し言葉コーパスを新たに構築する。それらを用いて、これまでの研究で得られている知見が、統制のとれた2言語コーパスを用いた対照研究においても当ては

まるかどうかを検証する。

### 4. 研究成果

日本語に関しては、まず、文節境界におけるフィラーの出現率について、分析を精緻化した。即ち、分析対象とした『日本語話し言葉コーパス (CSJ)』中の講演のタイプ別(模擬講演 vs. 学会講演)、話者の性別に、修飾文節の係り先までの距離(文節数)と修飾文節直後のフィラー出現率との関係を調べた。修飾文節直後のフィラー出現率は、係り先までの距離が約5までの間は、講演のタイプ、話者の性別を問わず、係り先までの距離の伸長に伴い、単調に増加した(図1参照)。この結果は、フィラーが、比較的短いスパンの言語化の認知的負荷を反映した現象であるという仮説を支持するものである。また、話者が一時に言語化しているのはほぼ5文節の長さまでであることが示唆された。さらに、同様の分析を、頻出する4種類のフィラー、「ア」「エー」「エート」「マー」について、種類別に行った。その結果、修飾文節直後のフィラーの出現率は、係り先までの距離の伸長に伴い、単調に増加しているものの、その上昇パターンは、フィラーの種類によって異なっており、フィラーの種類を選択に当たっては、話者の性別や講演のタイプが影響していることが示唆された。(渡辺・柏木, 2014)

文境界や節境界では、通常、ポーズが置かれる。ポーズの間にも、話者は後続発話のプランニングを行っているはずである。話者は、どのようなときに、黙って考えるのではなく、フィラーを発するのだろうか。この問いに答えるために、文境界と節境界におけるポーズとフィラーの持続時間を調べ、以下のような仮説を検証した。

(1) 境界の深さがポーズ長に影響する。深い境界ほどポーズは長い傾向がある(境界仮説)。

(2) 後続節の長さがポーズ長に影響する。後続節が長いほど、ポーズは長い傾向がある(複雑性仮説)。

(3) ポーズ長がフィラーの出現率(フィラー使用の有無)に影響する。ポーズが長くなるほど、フィラーの出現率は上昇する。分析対象として、詳細な言語情報の付与された、CSJ コア中の模擬講演 106 講演(男女各 53 名)を用いた。境界は、文境界、等位節境界(南(1974)による分類のタイプC)、従属副詞節境界(同、タイプB)の3種類に分けた。境界の深さを、

文境界 > タイプC > タイプB

と想定した。これら3種類の境界におけるポーズ長、ポーズ長+フィラー長を、後続節中の文節数の関数として計測した。ポーズ長の結果を図2に、ポーズ長+フィラー長の結果を図3に示す。まず、境界が深いほどポーズは長く、境界仮説は支持された。また、後続節が4文節の長さになるまでは、文節数の増加に伴いポーズ長も増加した。しかし、後続

様式 C - 19、F - 19、Z - 19 (共通)

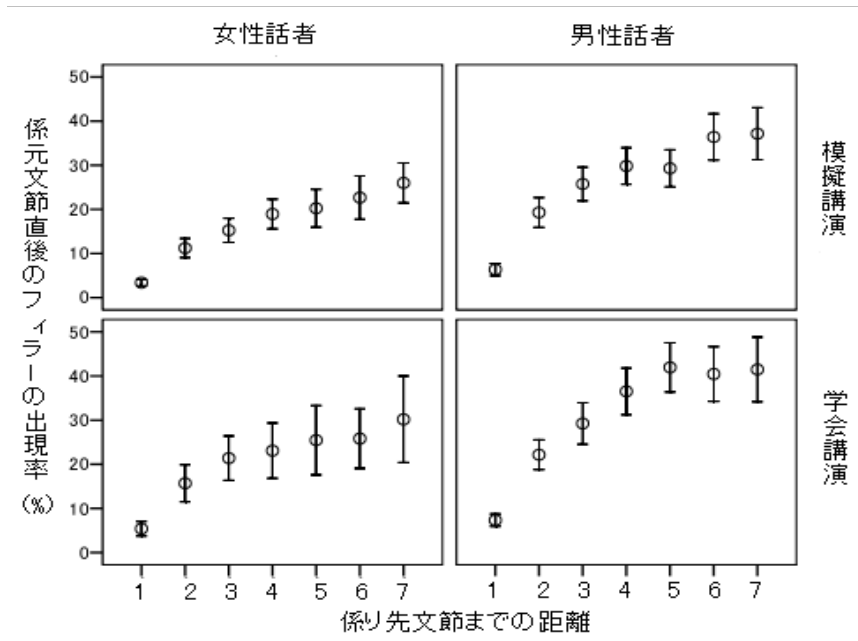


図1 係り元文節直後のフィラーの出現率 (エラーバーは標準誤差)

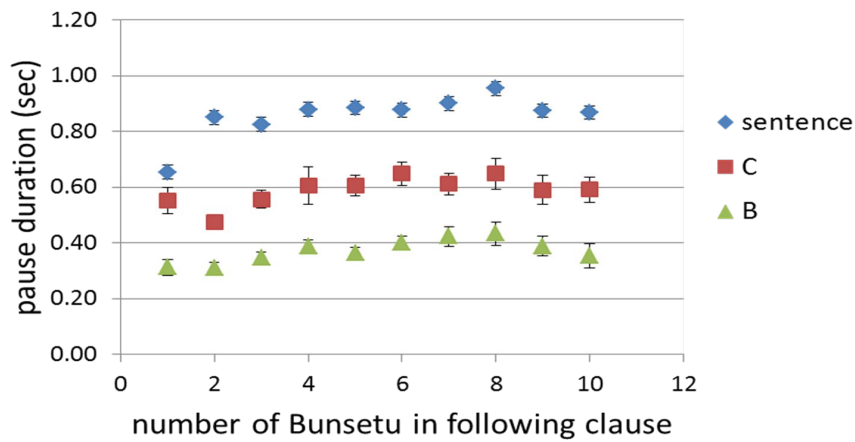


図2 3種類の境界におけるポーズ長。横軸は後続節中の文節数

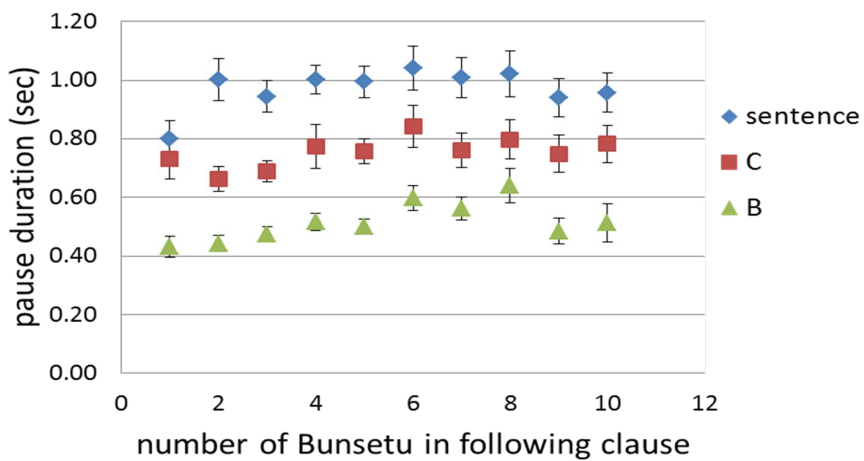


図3 3種類の境界におけるポーズ長+フィラー長。横軸は後続節中の文節数

節がそれ以上の長さになっても、ポーズ長の増加は見られなかった。したがって、複雑性仮説は後続節長が4文節以下の場合に限って支持された。ポーズ長+フィラー長も、ポーズ長だけの場合とほぼ同様のパターンを示した。これらの結果から、境界の深さや後続節長の影響を受けるのは、主としてポーズであることが示唆された。各境界におけるフィラーの出現率は、いずれも、ポーズ長の伸長に伴って上昇する傾向があった。このことから、話者は、文や節などの深い統語境界においても、沈黙している時間をある程度以下に抑えようとしていることが示唆された。

以上のように、日本語の言い淀みに関しては、文節の係り先までの距離とフィラーの出現率との関係を、フィラーの種類による特徴を含め、詳細に分析した。また、文境界、節境界における、ポーズとフィラーの持続時間（音響的特徴）の分析を行い、両者についての新たな知見を得た。

研究開始時には、日本語と英語のコーパスを新たに作成する予定であった。しかし、日本語に関してはCSJに詳細なデータがあること、時間的・経済的制約を考慮し、CSJの模擬講演と対照研究が可能な英語コーパス、“The Corpus of Oral Presentations in English (COPE)”を作成することにした。CSJの模擬講演のトピックの一つに類似した、“The most impressive event in my life”というタイトルで、10分～15分のスピーチを、大学生または大学卒以上の学歴を持つ20代の米国西部英語母語話者男女各10名から、現地で収集した。スピーチの書き起こしテキストには、詳細な言い淀みラベル、文境界、節境界情報が付与されている。研究開始当初は、主として、文境界、節境界における言い淀みの分布を分析する予定であった。しかし、日本語の分析が進むにつれ、言い淀み、特にフィラーは、節よりも小さい、句などの言語単位との関係が深いことが予測されるようになった。そこで、現在、日本語における文節の係り元と係り先の距離に相当するような、句レベルの言語的複雑さの指標を得る方法を検討している。また、ポーズを含めた、言い淀みの持続時間情報の付与を進行させている。

## 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 1 件)

渡辺美知子・柏木陽佑「後続句の複雑さが文節境界におけるフィラーの出現率に与える影響」『音声研究』18巻1号, pp.45-56. 2014年4月 査読有り

[学会発表](計 4 件)

渡辺美知子・柏木陽佑「先行節の種類や

後続節の長さが文境界・節境界のポーズ長・フィラー長に及ぼす影響」『日本音響学会 2014 年秋期研究発表会講演論文集』, pp.281-282. 2014年9月3日, 北海学園大学豊平キャンパス 査読なし

Watanabe, M. “Phrasal complexity and the occurrence of filled pauses in presentation speeches in Japanese”, Proc. Disfluency in Spontaneous Speech, DiSS 2013, pp.71-72. August, 23, 2013, KTH Royal Institute of Technology, Stockholm, Sweden. 査読有り

渡辺美知子「後続句の言語的複雑さとフィラーの使用について」第327回日本音声学研究会研究例会, 2013年6月22日, 国立国語研究所 発表要旨『音声研究』17巻2号, p.90. 査読有り

Watanabe, M. “The distance between a modifying phrase and the head is a determining factor for the rate of filled pauses in Japanese spontaneous monologues,” 14th Australasian International Conference on Speech Science and Technology, December, 4, 2012, Macquarie University, Sydney, Australia. 査読有り

[図書](計 2 件)

渡辺美知子「2章 音声学・音韻論」(pp.19-42) 近藤安月子・小森和子(編)『研究社 日本語教育事典』研究社 2012年7月

Watanabe, M., and Rose, R. “Pausology and Hesitation Phenomena in Second Language Acquisition.” In Robinson, P. (ed.), The Routledge Encyclopedia of Second Language Acquisition, New York: Routledge, pp.480-483. July, 2012.

## 6. 研究組織

(1) 研究代表者

渡辺 美知子 (Watanabe Michiko)

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 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