

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 8 日現在

機関番号：32620

研究種目：挑戦的萌芽研究

研究期間：2011～2013

課題番号：23652095

研究課題名(和文)日本語の統語構造は発話のタイミング制御を規定するか

研究課題名(英文)Is timing control of Japanese speech rhythm constrained by syntactic structure?

研究代表者

東 淳一(Azuma, Junichi)

順天堂大学・医学部・教授

研究者番号：90202621

交付決定額(研究期間全体)：(直接経費) 1,600,000円、(間接経費) 480,000円

研究成果の概要(和文)：日本語の発話を音声分析し、文節読みだけの発話長を基準に文発話での各文節の伸長度を分析した。その結果、深い統語境界直前の文節は相対的に長く発話される傾向にあることがわかった。ただし引伸しが顕著な文節の直前では、統語構造に関係なく文節長が相対的に短くなる傾向があり、統語構造とは関係なく文節長の相対的な長さが「短長短長」のパタンをなす文もあった。クラシック音楽演奏の音響分析からも(小)楽節内での小節長について同様な現象が明らかになっており、日本語発話、音楽演奏の両者に見られるこの種のテンポの揺れ現象のメカニズムは不明ではあるものの、人間の基本的なタイミング制御の特質に根ざす可能性がある。

研究成果の概要(英文)：Short Japanese utterances were analyzed acoustically to investigate the temporal fluctuation of syntactic phrases. With reference to the average duration of the phrase-only utterances, duration of each phrase within the sentence utterance was analyzed. Though Japanese is said to have a mora-timed rhythm with little fluctuation of mora duration, duration of a sentence-final phrase or a phrase just before a deep syntactic boundary was relatively longer. The phrase just before a prolonged phrase was relatively shorter.

Duration of a musical bar was also acoustically analyzed. Pieces of classical music with quick tempo were investigated. The results also showed the tendency of "Final-lengthening" at the end of a musical phrase, which normally consists of four bars. However, the duration of a bar just before the prolonged one was often found significantly short. Within the general four-bar phrase structure, a "Short-Long-Short-Long" pattern was often observed.

研究分野：人文学・言語学・言語学

科研費の分科・細目：3001

キーワード：統語構造 Final-lengthening モーラリズム リズムの揺れ 音楽演奏 フレーズング

1. 研究開始当初の背景

本研究は、簡単な音声産出実験を通じて、日本語の統語構造とタイミング制御の関係について考察を試みることからスタートした。日本語の統語構造と韻律的特徴については、過去にさまざまな研究が行われているが、特にその中でも日本語の統語構造と基本周波数(以下F0)に関するものは大変多い。ところが、日本語発話のタイミング制御、あるいは日本語のリズムに関する研究については、言語リズムの類型化に関するものが多いようだ。(Grabe・Low, 2002; Ramus, 2002; White・Mattys, 2007)。

日本語の統語構造とタイミング制御の関係に特化した研究はあまり多くないが、Campbell (1997)の研究は大変貴重である。この研究では大量の音声コーパスをラベリングし、日本語発話の音素片の長さを分析することでセグメント固有の継続長を求め、現実の発話のさまざまな環境で音素長がどのように伸縮するのかが分析された。この研究では、深い統語境界や文末近傍ではセグメント長が相対的に大となることがはっきりと示された。ところが、この後に Grabe・Low (2002)の Pairwise Variability Index の概念を用いた言語リズムの類型化に関する研究が一世を風靡し、日本語のタイミング制御と統語構造との関連に関する研究は残念ながら途切れてしまったようである。

日本語の統語構造とタイミング制御の関連に関する研究がその後発展しなかった理由だが、まず日本語には強固なモーラリズムがあり、セグメントの時間的伸縮は極めて少ないと頭から信じられていることがある。英語のようないわゆる stress-timed の言語の場合、同一語彙が異なる統語環境でかなり自由に異なる継続長を与えられる。このため、統語構造とタイミング制御の関係を調査・研究することはさほど困難ではなく、Lehiste (1977)や Donovan・Darwin (1979)など著名な音声学者がこの問題に組み多く業績を残した。対する日本語の場合、仮名に対応するモーラの長さそのものが語義の区別に関与するため、これらの長さを自由に変動させることは音声言語理解を妨げると想定される。しかしながら、このことが日本語において統語構造など言語の上位構造の影響を受けてモーラ長が変動する可能性を完全に排除するわけではないだろうというのが本研究のスタンスである。

2. 研究の目的

このような状況を踏まえ、本研究では日本語の統語構造とタイミング制御との関連を詳しく見てみることにする。特に、文末や文中の深い統語境界においてモーラ長の引き伸ばし、つまりいわゆる Final Lengthening (以下、FL)に類似した現象があるのではな

いか、FLに類似の現象以外に何らかの特徴的なタイミング制御の特徴がないかという点について調査した。言語の普遍性を考える場合、mora-timedな言語であるとはいえ日本語においても韻律的特徴の1つである言語リズムが統語構造とパラレルに変動するのではないか、文末や文中の深い統語境界においてモーラ長の引き伸ばし、つまりいわゆるFLに類似した現象が観察されるのではないかと考えた。このことを検証するのが本研究の目的である。

3. 研究の方法

(1) 音声データ収集と前処理

本研究では、たとえば「カマタが重ねたカルテを借りたよ」のように、文節頭が必ず「か」で始まる特殊なセンテンスをさまざまな方言をもつ日本人話者にポーズなしでそれぞれ10回発話してもらい、発話を WAV ファイル形式で収録した。

収録された音声データは WaveSurfer を用いて音声分析し、「カルテを」、「借りたよ」などの文節部分の持続時間を測定した。各文節の最初の/k/の破裂にともなうバースト部分から次の文節の最初の/k/のバースト部分直前までをそれぞれの文節の持続時間とした。必要に応じてスペクトログラムを併用して測定した。文末では音声波形がバックグラウンドノイズと同等の振幅まで減衰するまでに時間がかかるため、実際よりも文節長を長く見積もってしまう可能性がある。このため音声波形とともにパワー表示とスペクトログラムも使用し文節末の有声音の終わりを判定した。

録音後に、さらに「カワノが」、「重ねた」、「カルテを」など、いわゆる文節のみを「N回目(短いポーズ) ____か」というキャリアフレーズ条件下で、テスト文の時と同様な発話速度で10回(話者によっては15回)発話してもらいデジタル録音した。これは基準となる文節長を得るために行われたものである。もともとフルセンテンス発話においても、文末を除き各文節の直後には「か」の音声が生起するので、文節のみの単独発話でも同じ環境を保った。文節単独発話においても音声波形の形状で最初の/k/の破裂が認められる部分から末尾の「か」の破裂が認められる部分の直前までを確認しその間を文節の長さとした。

(2) 分析メソッドについて

今回文節のみの発話を収録したのは、フルセンテンス発話での各文節の伸縮状況を調査する上で、基準となる文節長を得るためである。これら文節のみの発話については各文節の10回分または15回分の発話の長さを測定しているが、それぞれの文節の平均値によ

り当該話者がその文節を発話する場合の標準的な長さとして仮定した。

ここでもしもフルセンテンス発話のみのデータから各文節長を論じるとどのような不具合が生じるのであろうか。「カメラを」という文節の長さの平均値が 500 ms、「借りたよ」の文節の長さの平均値が 550 ms であったとしよう。このことで「借りたよ」の文節長は「カメラを」と比較して継続長が引き伸ばされたといえるだろうか。実はそれぞれの音節片には固有値があり、またそれらが配置される環境によってもその音素片長は影響を受ける。そうするとこれらの数値はたとえばそれぞれの文節が単独で発音されても、あるいは「借りたよ、カメラを。」という語順の文として発話されたとしても、ひょっとすると固有値のために変わらないかもしれない。結局はそれぞれのセンテンス発話での各文節の長さは、何らかの文節の固有値のようなものと比較・対照の上で分析されなければならないことになる。しかも、できれば文節ではなくむしろモーラや音素片単位の分析ができればベストである。また、このような基準となる固有値はできればフルセンテンス発話者と同一人物のものであればより基準としてふさわしいものとなる。ところが残念なことに各モーラあるいは各音素片の長さの固有値を正確に話者ごとに求めることは不可能に近い。特に有声音の連鎖を切り分けるのは全く不可能である。

しかし今回の分析方法であれば、文節単位の分析しかできないものの、文節に相当する音声の切り出しがある程度正確にできるため、フルセンテンス発話の場合の文節長と文節単独発話での文節長を客観的に比較できるようになる。つまり間接的ではあるが、これで FL の有無もある程度判定可能になる。

(3) 文節長の分析方法

フルセンテンス発話において文節が相対的に引き伸ばされているのか、あるいは相対的に短くなっているのかを分析するため、二段階の分析を行った。まず、もしも仮想的に単独文節発話の各文節の平均値を並べること、分析したいと考えるセンテンスを時間的に構成するならば、各文節長が仮想センテンス長の何%を占めることになるのかを計算した。たとえば、「カマタが重ねたカルテを借りたよ」の場合、「カマタが」、「重ねた」、「カルテを」、「借りたよ」の各文節の理論的な発話長(%)はそれぞれ 25.82%、25.23%、25.09%、そして 23.86%となった。これに対し、フルセンテンス発話においても、発話ごとに各文節がそれぞれ文全体の長さの何%を占めることになるのかを計算した。ここで、後者つまりフルセンテンス発話における各文節の発話長(%)を文節単独発話の理論的な発話長(%)で割り算することで、実際のフルセンテンス発話においてそれぞれの文節長が理論的な発話長からどの程度逸

脱したのかをおおよそ知ることができる。この値を文節長レシオと名付ける。

(4) 分析上の問題点

この分析方法での問題点は、モーラではなく文節単位でしか発話リズムの伸長状況を分析できないことであるが、さらにもう1つ認識しておかねばならない問題点、あるいは分析メソッドの限界がある。それはフルセンテンス発話の最後の文節の取り扱いである。文末では音声波形がバックグラウンドノイズと同等の振幅まで減衰するまでに時間を要するため、文節長をより長く見積もってしまう可能性があり、音声波形、パワー表示、そしてスペクトル表示を併用して文節の末尾を判定した。一方、文節単独発話の場合、さらに「借りたよか」のように本来の文節に「か」を付加して発話を依頼しており、両者で文節の発話環境が異なる。文節単独発話の場合には、本来は存在していない/k/音の破裂直前まで測定しているため、閉鎖部分の長さも「借りたよ」に含まれてしまっている。このため、文節単独発話の場合のいわば基準となる「借りたよ」の長さは若干過大に見積もられている可能性がある。ただ、前述のようにフルセンテンス発話の場合には最後の「借りたよ」の文節は若干長く見積もられている可能性があるため、互いに効果がある程度相殺されているとみなした。

(5) 再び仮説について

すでに、仮説については簡単に述べたが、Lehiste (1977) などの研究者が報告している FL が普遍的であるならば、深い統語境界の前では FL が観察されるはずである。また、FL は最後のモーラだけではなく、ある程度その近傍で一定の時間的広がりをもって生起するであろうから、文中の深い統語境界の前の文節は後の文節よりも長くなると考えられる。逆に修飾・被修飾のような係り受け関係で複数文節が続く場合は、各文節の長さはほぼ同じか、または後の文節の方が長くなると推測される。当然ながら、文末の文節は FL の効果が最も強くなるため、当然長くなると考えられる。つまり、次のような構造の文について、文節長レシオは図 1 右側に図示したようになると想定される。

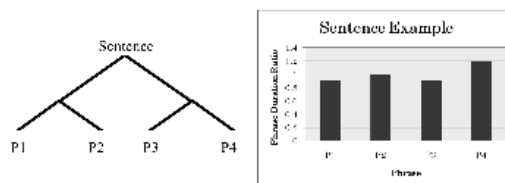


図 1: 文構造と文節長レシオ (P: 文節)

(6) 発話の分析例

それでは、次にフルセンテンス発話の文節長レシオとその平均値をいくつか図示する。

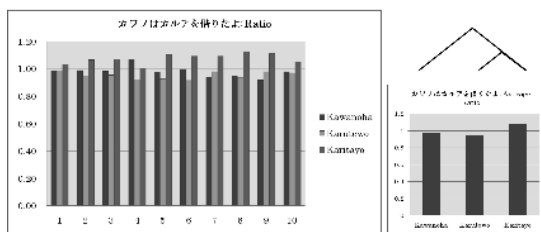


図2：「カワノはカメラを借りたよ」

まず「カワノはカメラを借りたよ」の発話では「カワノは」と「カメラを」の間の統語境界があるため、最初の「カワノは」の文節は次の「カメラを」よりも長くなっている(図2参照)。7回目や9回目の発話では例外的に最初の「カワノは」が短くなっているが、平均値を見る限り仮説どおりの値になっていると思われる。

次に「刈谷でカワノがカメラを借りたよ」では、最初に右枝分かれの構造が続くため、「刈谷で」および「カワノが」の文節がそれぞれ後続の文節と比較して長くなることが予測される。図3に示したように、実際の発話でも最後の3つの発話を除けば予測通りのパタンとなった。また、文節レシオの平均値についても予測どおり最初の2つの文節レシオが後続のものよりも長くなった。

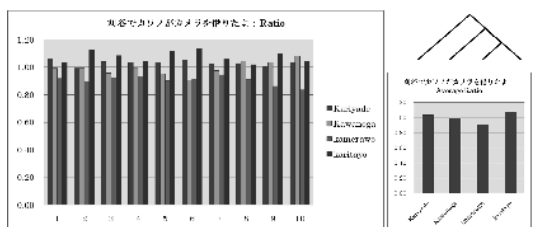


図3：「刈谷でカワノがカメラを借りたよ」

次に、「カマタが重ねたカルテを借りたよ」という発話では、発話者には、「カマタが」は「重ねた」の動詞の主語であり、文末の本動詞「借りたよ」の主語は別の人(つまり「私は」などが省略されていると仮定)であることを説明した。その文意で発話をしてもらった結果を示している。

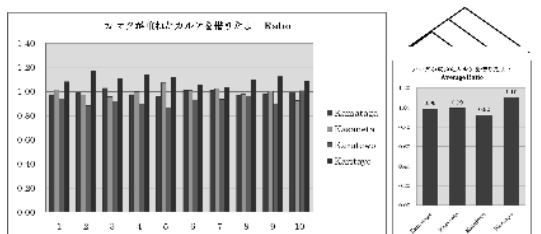


図4：「カマタが重ねたカルテを借りたよ」

このため、図4に示したように発話番号2、

3、10をのぞいては、最初の文節は後続のものよりも短くなっている。平均値についても同様で、最初の文節は後続のものよりも短いという結果になった。ところが、2つ目の文節も3つ目の文節よりも短いとか同等の長さになるであろうと予測したものの、実際は後続の文節よりも長くなった。このため、ここでは4つの文節レシオが相対的に短長短長となるパタンを示した。

以上のように、多少の例外はあるが、ほとんどの文において統語構造と文節長との関係には何らかのマッチングがあり、深い統語境界があればその前の文節長は長くなり、文末でも文節長が相対的に長くなる傾向があった。

4. 研究成果

(1) 成果概略

Campbell (1997) は大規模音声コーパス中の音素片をラベリングし、さまざまな言語環境での継続長を分析することでそれぞれの音素固有の継続長を求めた。そして各音素片の継続長の固有値を基準にして、現実の文発話での音素片の継続長がどのように伸縮するのかを解析した。その結果、英語のみならず、日本語においても統語境界直前では境界に向かって徐々に継続長が大となり、当然ながら文末においてはFL的な現象が観察されることを示している。本研究では音素片の継続長ではなく文節長を分析したため分析単位が粗くなっているものの、やはり同様の傾向が観察されている。

まず、すべてにおいて文末の文節長レシオは相対的に大きい。文中においても同様で、文頭の副詞語句や主部となる文節は後続の文節と係り受けをなさないため、文節長レシオは大きくなる。また、2つの文節が修飾・被修飾の連鎖をなす場合、修飾する方の文節長レシオは修飾される方の文節長レシオよりも若干短くなることも確認された。

(2) 例外をどう考えるか

今回の分析ではある程度統語構造と文節長の伸縮には平行性が見られたが、個々の発話を見ると、ある程度例外も観察されている。図4も典型的な例で、[[カマタが]重ねた]カルテを]において、話者は十分文意を理解しているにもかかわらず、「重ねた」の文節が相対的に「カルテを」の文節よりも長くなっている。これら2つの文節の間には深い統語境界があるわけではないが、この現象が見られており、また全般的にはどの文においても最後から2つ目の文節は必ず文節レシオが小さくなる傾向があった。これは最後の文節でのFLに対する補償効果である可能性もあるが、さらなる検証が必要である。なお、4つの文節からなる文で例外的なデータが

あった場合、文節長について相対的に「短長短長」といったグルーピングが行われているように思われるケースもあった。

(3) 音楽の場合

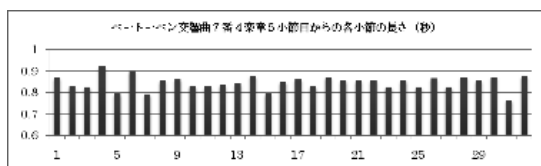


図5：音楽演奏でのテンポの揺れの例

実は、「短長短長」のグルーピングが行われる、または短い単位があってその後FLが配置されるという傾向はクラシック音楽の演奏においても観察される。たとえば、上の図5はベートーベンの交響曲7番（Roy Goodman 指揮 Hanover Band）第4楽章（2/4拍子）の最初のテーマが始まる部分、つまり5小節目からの各小節の持続時間をWaveSurferにより測定した結果を図示したものである。5小節目を1として、4小節ごとにまとまりをなす楽節の最初の小節数を付記してある。「短長短長」のグルーピングの存在、短い小節の後にFLが来るといったテンポの揺れ現象は大変多く観察されるが、基調となる弦楽器の動きが16分音符で、しかも大変速いパッセージを奏する場合においても楽節を作る4小節ごとにこのパターンが観察されることは大変興味深い。この速いパッセージを奏する場合、小節ごとに意識してテンポを揺らすことは奏者がプロであったとしても実際のところ不可能である。音楽など言語以外の人間のリズム的なパフォーマンスにおいても、フレーズ末（文末）に相当する部分においてFLがあり、反面その直前の単位では逆に継続長が短くなるという現象が頻繁にあるとすれば、例外も含め今回の文発話での現象と同じタイミング制御メカニズムが働いている可能性があり、興味深い。

(4) 結語

以上、小規模な発話実験からではあるが、深い統語境界がある場合に境界直前の文節が相対的に長く発話され、文末の文節は相対的に最も長く発話されることが確認された。また、最後から2つ目の文節は最も短く発話される傾向があった。時に、統語構造に関係なく、4つの文節からなる文で文節長が「短長短長」となる傾向も一部観察された。一連の分析から当初の仮説はほぼ検証されたといえるが、これらの現象は音楽演奏時のテンポの揺れ等、人間の一般的なリズムカルなパフォーマンスにおける普遍的なタイミング制御機構研究も参照しつつさらに精査されねばならないであろう。また、今後は何

らかの形でモーラレベル、音素レベルでの検討も必要であり、より多くの音声資料、多彩な方言話者の発話資料による分析が求められる。

参考文献

Campbell, N. (1997). 「プラグマティック・イントネーション：韻律情報の機能的役割」 音声文法研究会編『文法と音声』東京：くろしお出版、55-74.

Donovan, A., & Darwin, C. J. (1977). The perceived rhythm of speech, Proceedings of the Ninth International Congress of Phonetic Sciences, 1, 268-274.

Grabe, E. and E. L. Low (2002). Durational variability in speech and the rhythm class hypothesis, Gussenhoven, C. and N. Warner (Eds.) Laboratory phonology 7, Berlin: Mouton de Gruyter, 515-546.

Lehiste I. (1977). Isochrony reconsidered, Journal of Phonetics 5, 253-263.

Ramus, F. (2002). Acoustic correlates of linguistic rhythm: Perspectives, Proceedings of speech prosody, Aix-en-Provence, 115-120.

White, L. and S. Mattys (2007). Calibrating rhythm: First language and second language studies” Journal of Phonetics 35-4, 201-522.

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)

東淳一、日本語リズムの揺れと音楽演奏テンポの揺れ 人のリズム活動に潜む機序を求めて、外国語教育メディア学会(LET)関西支部メソドロジー研究部会 2012 年度報告論集、査読なし、2012、1-13、http://www.mizumot.com/method/2012-01_Azuma

〔学会発表〕(計5件)

東淳一、統語構造は言語リズムを規定するか：英語および日本語の考察、2014 年度大学英語教育学会(JACET)関西支部春季大会、大阪薬科大学、2014.6.14.

Azuma, Junichi, Temporal Fluctuation of Japanese Speech and Music Performance: Pursuit of General Principle Underlying the Control of Human Rhythmic Performance, 11th International Conference on New Directions in the Humanities, Eötvös Loránd University, Budapest, Hungary,

2013.6.19.

東淳一：日本語の発話タイミングの揺れに関する考察，近畿音声言語研究会月例研究会，西宮市大学交流センター，2013.2.2.

東淳一：日本語リズムの揺れと音楽演奏テンポの揺れ—人のリズム活動に潜む機序を求めて，外国語教育メディア学会関西支部メソドロジー研究会／言語テスト・第二言語習得合同勉強会，流通科学大学東京オフィス、2012.11.10.

Azuma, Junichi, Timing Control of Japanese Speech and Temporal Fluctuation of Musical Performance, 13th International Conference on Language and Social Psychology, WTC Leeuwarden, Leeuwarden, 2012.6.22.

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

6. 研究組織

(1) 研究代表者

東 淳一 (AZUMA, Junichi)

順天堂大学・医学部・教授

研究者番号：90202621

(2) 研究分担者

なし

(3) 連携研究者

なし