

# 英語の高頻度品詞連鎖における韻律パターン認識

村尾玲美

キーワード：高頻度品詞連鎖、句レベルの韻律、韻律パターン認識

## 1. 研究の背景と研究目的

句レベルのプロソディ情報 (phrasal prosody) は連続音声を統語構造に分節化する手がかりとなり (Gout et al., 2004)、ガーデンパス文などの統語的曖昧性を解消する役割を持つ (Millotte et al., 2007)。無意味語から成る文 (meeving gups keebed gompily) でも統語構造を為す文 (形容詞 + 名詞 + 動詞 + 副詞) のプロソディに則って発音されれば、無意味語の記憶の保持が高く (Epstein, 1961)、シャドーイングの再生率も高くなることから (Martin, 1968)、文構造に沿ったプロソディ情報は言葉の記憶や処理を容易にすると考えられる。また、Soderstrom et al. (2003) は、生後6ヶ月から9ヶ月の乳児でも名詞句や動詞句を成す構造と句を成さない構造を音韻的に識別していることを明らかにした。6ヶ月から9ヶ月の乳児は “Today, **people by the hole** seem scary.” という名詞句や “Inventive people **design telephones** at home.” という動詞句を、それぞれ “In fact, some **people # buy the whole** supply of them.” と “The director of **design # telephones** her boss.” という句を為さない発話と区別し、前者を記憶したのである (# は句の切れ目を表す)。このことから、母語話者は単語や定型表現などの具体的な語彙項目のみならず、句や節といった構造がどのような韻律を持つかを知っていると考えられる。更に、Valian & Levitt (1996) の研究では、英語母語話者に内容語と機能語の組み合わせから成る人工語を学ばせたところ、強弱アクセントによるプロソディ情報を与えて学ばせたグループの方が、プロソディ情報無しで学ばせたグループよりも、人工語の語順や単語の組み合わせのパターンをより良く習得した。つまり、英語母語話者は強弱アクセントのプロソディ情報と共に語の並びや組み合わせを覚えていると考えられる。

村尾 (2006) では、上級の日本人英語学習者も、親密性の高い文構造 (疑問詞 + be 動詞 + 代名詞 + 動詞の進行形) であれば、その構造を持つ韻律パターンを知っており、プロソディ情報のみを与えた場合でも what are you doing などの具体的な表現を認識することができること示唆された。

以上の先行研究から、プロソディ情報は統語解析や意味理解や統語構造の獲得の手助けとなり、母語話者 (および一部の学習者) は、句や節を形成する品詞の並びがどのよう

な韻律パターンを持っているか知っている可能性が示唆される。句や節がどのような韻律パターンを持っているか知っているということは、リスニングのみならず、リーディングにおいても統語構造を理解するのに役立つと考えられる (Dowhower, 1991)。そこで本研究では、学習者が句を形成する品詞連鎖を韻律的なパターンとして処理することができるかどうかを明らかにする。

## 2. 研究方法

### 2. 1 実験項目

TreeTagger によって品詞タグ付けされた英語母語話者コーパス (NICE 2.0: NS) と日本人英語学習者コーパス (NICE 2.0: NNS) から5品詞の連鎖を抽出し、① NNS・NS ともに高頻度の品詞連鎖、② NS のみ高頻度の品詞連鎖、③ NNS のみ高頻度の品詞連鎖、④ NNS・NS ともに低頻度の品詞連鎖の4条件に分類した (表1)。品詞連鎖は句読点を含まない、連続した連鎖のみを抽出した。なお、NICE 2.0 における英語母語話者コーパスの総語数は 117,296 語で、日本人英語学習者コーパスの総語数は 114,550 語である。

表1. 条件別の品詞連鎖<sup>1</sup>

	NNS 高頻度	NNS 低頻度
NS 高頻度	① 冠 + 名 + 前 + 冠 + 名 (263/145) 名 + 前 + 冠 + 形 + 名 (185/91) 前 + 冠 + 名 + 前 + 名 (144/95) 冠 + 形 + 名 + 前 + 名 (140/50) 形 + 名 + 前 + 冠 + 名 (120/81) 冠 + 名 + 前 + 形 + 名 (88/71)	② to 不 + 動 + 代名 + 形 + 名 (24/6) to 不 + 冠 + 名 + 前 + 名 (22/1) 前 + 動 ing + 冠 + 形 + 名 (28/3) 冠 + 形 + 接 + 形 + 名 (22/3) 前 + 冠 + 名 + to 不 + 動 (48/5)
NS 低頻度	③ 代名 + 助 + 動 + 前 + 動 (0/76) 代名 + be 動 + 形 + 前 + 動 (0/43) 前 + 動 + 前 + 冠 + 名 (0/44) 代名 + 助 + 動 + 冠 + 名 (25/102) 形 + 名 + 接 + 形 + 名 (9/44) 形 + 名 + 前 + 形 + 名 (19/51)	④ 動 + 副 + 動 + 冠 + 形 (4/4) 前 + 代名 + 名 + 前 + 動 ing (4/4) 名 + 前 + 代名 + 名 + 前 (4/4) 前 + 冠 + 名 + 形 + 前 (4/4) 冠 + 名 + 接 + 代名 + 動 (4/4) 副 + 動 + 副 + 前 + 冠 (4/4)

(NS/NNS)

①②③からは句を形成する品詞連鎖を6種類ずつ選択し、④からは句を形成しない品詞

<sup>1</sup> この表では分かりやすさを考慮して品詞タグを日本語に訳し、漢字で表示しているが、実際の品詞タグにはいくつかの下位区分がある。実際の品詞タグは Appendix 1 を参照されたい。

連鎖を6種類選択した。各条件6種類の品詞連鎖に当てはまる具体的な単語連鎖をコーパスから2種類ずつ抽出し、合計48の実験項目を用意した。具体的な単語連鎖の抽出にあたっては、短音節語のみからなる単語連鎖と多音節語を含む単語連鎖の数が各条件で同じになるように考慮した。次節で説明する、韻律パターンマッチング課題の選択肢を作成するにあたり、いくつかの項目を削除したため、条件間の項目数が変わり、最終的には①は6種類の品詞連鎖と11個の単語連鎖、②は5種類の品詞連鎖と9個の単語連鎖、③と④は6種類の品詞連鎖と10個の単語連鎖の合計40項目となった（Appendix 1参照）。実験項目の音声ファイルを作成する際には、自然なプロソディで発音を録音するため、40個の実験項目をそれぞれ文中に埋め込み、アメリカ人英語母語話者に一文ずつ音読させた。その後音声編集ソフトでターゲットの単語連鎖のみを切り取り、実験項目の音声を作成した。

## 2. 2 実験参加者と実験方法

実験参加者は英語母語話者8名、上級英語学習者10名（TOEIC平均933点）、初級英語学習者14名（TOEIC平均418点）である。実験項目の音声にはローパスフィルタをかけ（cut-off point 400Hz）、音素情報を劣化させることで、表現のプロソディのみが手がかりとなるようにした。実験では音素情報を劣化させた刺激（例：as a way of life）を音声提示した後、2種類の5単語連鎖を選択肢として視覚提示し、刺激と同じ品詞連鎖から成る表現（例：in a court of law）か異なる品詞連鎖から成る表現（例：a great deal of time）のいずれに近い音声だったかを判断させた。この韻律パターンマッチング課題を行うには、実験参加者は視覚呈示された二つの表現を頭の中で音声として再現し、その再現された2表現のプロソディのどちらが、刺激提示として聞こえてくる表現のプロソディとよりよくマッチするかを判断する必要がある。品詞連鎖の持つ内容語と機能語の強弱リズムを、学習者が頭の中で正しく再現できていれば、マッチングは可能となる。

正解の選択肢は、刺激と同じ品詞連鎖に当てはまる単語連鎖をNICEコーパスから抽出することで作成した。不正解の選択肢を作成するためには、まず刺激の品詞連鎖の強弱リズムを内容語と機能語の観点から分析した。次に、内容語と機能語が一部反転するような品詞連鎖を考え、その品詞連鎖に当てはまる具体的な単語連鎖をコーパスから抽出した。正解と不正解の選択肢は音節数を揃え、多音節語を含む場合、単語のアクセント位置も揃えるようにした。また、発音がわからないような難しい単語を含まないようにするため、単語を一部置き換えた。

正解の選択肢に刺激とは異なる単語連鎖を用いたのには二つ理由がある。一つは、刺激にローパスフィルタをかけても完全には音素情報を消し去ることはできないからである。聞こえ度の高い音素を手がかりとしてマッチングに成功するという可能性を減らす

ため、本研究では品詞連鎖は同じまま、異なる単語連鎖を用いた。二つ目の理由は、本研究の目的は品詞連鎖の韻律パターン処理能力を調査することであるため、品詞連鎖に当てはまる具体的な単語の並びは多様である必要があるからである。また、本研究では問題は二者択一とし、「わからない」という選択肢は設けなかった。これは刺激と正解の単語連鎖が異なり、頭の中で再現した音声と刺激が完全に合致することはないため、「どちらでもない」という意味で「わからない」を選択する学習者が増える可能性を排除するためである。本研究では回答への確信度は調査していないが、チャンスレベルの回答であるかどうかを見極める方法の一つである。

よく使用する品詞連鎖の韻律パターンとして処理しているのであれば、英語母語話者は③と④の品詞連鎖に比べ、①と②の品詞連鎖を正しく認識するのに対し、学習者は②と④に比べ、①と③を正しく認識するという仮説が成り立つ。

### 2. 3 分析方法

実験では正答率と音声再生回数と回答までの時間を測定したが、本稿では正解率の結果のみを報告する。分析は正解率を従属変数とし、実験グループ（母語話者・上級者・初級者）と四つの品詞連鎖条件を独立変数とする、二要因分散分析混合計画を行った。

### 3. 結果

図1は実験グループによる条件ごとの正解率を示した図である。二要因分散分析の結果、実験グループと条件の交互作用が有意であった( $F_{(6,87)} = 2.602, p < .05, \eta_p^2 = .152$ )。単純主効果の検定を行ったところ、④ NNS・NS ともに低頻度の品詞連鎖については実験グループ間に有意な差が見られなかったのに対し、その他の条件については母語話者と学習者の間に有意な差が見られた。以下に実験グループの単純主効果の結果をまとめる。

- ①： 母語話者 > 上級者 = 初級者
- ②： 母語話者 > 上級者 = 初級者
- ③： 母語話者 = 上級者 = 初級者  
母語話者 > 初級者
- ④： 母語話者 = 上級者 = 初級者

各実験グループにおける条件間の差については、母語話者は②と③が④に比べ有意に高い点であったのに対し、学習者は③が①に比べ有意に高い点であった。初級学習者は仮説に反して①よりも④の方が有意に点が高いという結果となった。以下に条件の単純

主効果の結果をまとめる。

母語話者： ②③>④  
 上級者： ③>①  
 初級者： ③④>①

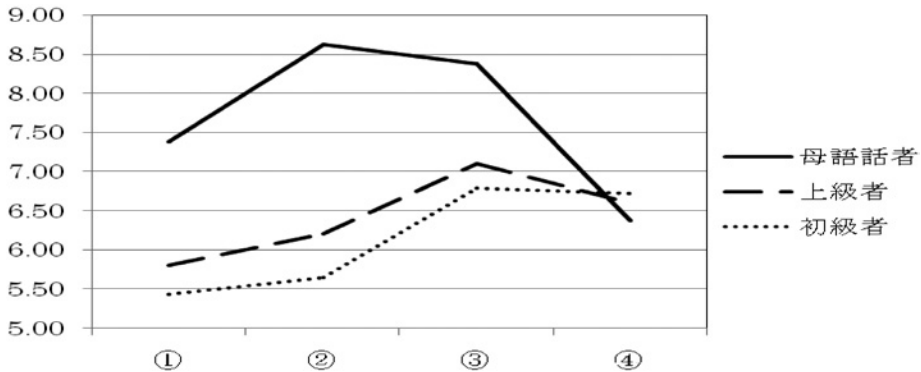


図1. 実験グループによる条件ごとの平均<sup>2</sup>

#### 4. 考察

##### 4. 1 仮説を支持する結果

実験の結果、②（NSのみ高頻度の品詞連鎖）は母語話者による認識は高く、学習者は低いという仮説を支持した。また、学習者は③（NNSのみ高頻度の品詞連鎖）の認識が高いという結果も仮説通りであった。③の認識に関しては母語話者も高いという結果を得た。これは、③の表現が、本研究で用いた母語話者のエッセイライティングのコーパスにおいては多く出現しなかっただけであり、他のジャンルの書き言葉や話し言葉では頻出する品詞連鎖であったためと考えられる。表1からも分かる通り、③には代名詞を含む表現が多く含まれていることから、学習者がエッセイライティングに一人称を多用していることが見て取れる。これらの結果から、母語話者も学習者も、よく使う品詞連鎖は韻律的なパターンとして身に付けている可能性が示唆された。

##### 4. 2 仮説を支持しない結果

学習者は仮説に反し、④（NNS・NSともに低頻度の品詞連鎖）の認識が他と比べ高かった。Thorp & Trehub (1989)によると、言語獲得過程にある6ヶ月から9ヶ月の乳

<sup>2</sup> 正解は10.00点を満点として換算した。

児は、音のまとまりよりもピッチやポーズによって境界が変わる部分に敏感に反応する。本研究で用いた④の項目は句としてのまとまりを為さない品詞連鎖であるため（例：the paper, and they look）ピッチやポーズによる音韻変化がより顕著であり、マッチングの手がかりが多かったため、他と比べると認識しやすかったと考えられる。興味深いのは、④の正解率が母語話者と学習者で同じである点である。④をコントロールと考えるなら、母語話者は④よりも句を成す品詞連鎖を正しく認識したのに対し、学習者（初級者）は、品詞連鎖というまとまりとして韻律パターンを認識するよりも、音韻変化が顕著な部分に反応して認識していると考えられる。

また、学習者は仮説に反し、①（NNS・NSともに高頻度の品詞連鎖）が有意に低く、母語話者も予測より低い結果となった。①の項目を見ると、高頻度であるがゆえに品詞のバリエーションが少なく、冠詞、名詞、形容詞、前置詞の四種類しか含まれていないことが分かる。単語認知の研究では、高頻度語は音韻的に競合する候補語である“Neighborhood density”が高く、認知が難しくなると言われている（Garlock, Walley & Metsala, 2001）。品詞連鎖も同様に、高頻度であれば音韻的に競合する表現が増え、その分認識が困難になることが予想される。一つの解釈としては、高頻度品詞連鎖の韻律パターンは習得はされているものの、類似したパターンが多いために、今回のような韻律パターンマッチング課題で選択するのは困難であったと考えられる。二つ目の解釈としては、高頻度品詞連鎖の韻律パターンは習得されていないという考えである。言語習得における頻度効果については Ellis（2002）が述べているが、Peters（1985）が述べているように、こどもは言語獲得において韻律的に顕著なまとまり（prosodically salient chunks）に注目して心的辞書に保持していく。言語の連鎖が暗黙的に心的辞書に保持されるには高頻度であるだけでなく、音韻的に顕著である必要があり、①の項目はその条件に満たなかった可能性も考えられる。

#### 4. 3 全体的傾向

学習者が他の条件に比べ③を高く認識したのは仮説通りであったが、正解率は全体的に低かった。TOEIC900点を超える上級者でも高頻度品詞連鎖を持つプロソディを区別できなかったことから、品詞連鎖を韻律的なパターンとして捉える能力は言語習得において重要な要素ではないと考えられる。また、上級者と初級者に差が見られなかったことから、この能力はTOEICが測定するリーディング力やリスニング力に直接影響する能力ではないと考えられる。しかしながらまた、母語話者とは点数の差が見られ、交互作用も見られたため、TOEICで測られていない能力となんらかの関係がある可能性は否めない。

## 5. 結論

本研究では、日本人英語学習者が句を形成する品詞連鎖を韻律的なパターンとして処理することができるかどうかを調査した。品詞連鎖を高頻度と低頻度の条件に分けて韻律パターンマッチング課題を行ったところ、上級英語学習者と初級英語学習者の間には品詞連鎖条件に関わらず正解率に差が見られなかった。一方、英語母語話者との間には正解率に有意な差が見られたのみならず、品詞連鎖条件による違いが見られた。以上の結果から、①品詞連鎖の韻律パターンとして処理できるかどうかは言語習得において重要な要素ではないこと、② TOEIC で測られるリスニング力とリーディング力に影響する要素ではないこと、③母語話者との処理の仕方に違いがあること、が明らかとなった。

## 参考文献

- Cutler, A., Dahan, D., Donselaar, W. V. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40, 141-201.
- Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice*, 30 (3), 165-175.
- Echols, C. H. (1993). A perceptually-based model of children's earliest productions. *Cognition*, 46, 245-296.
- Ellis, N. (2002). Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Epstein W. (1961). The influence of syntactical structure on learning. *American Journal of Psychology*, 74, 80-85.
- Garlock, V. M., Walley, A. C., and Metsala, J. (2001). Age-of-acquisition, word frequency, and neighborhood density effects of spoken word recognition by children and adults. *Journal of Memory and Language*, 45, 468-492.
- Gout, A., Christophe, A. and Morgan, J. L. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51, 548-567.
- Jusczyk, P. W. & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and Present. *Ear & Hearing*, 23, 2-40.
- Martin, J. G. (1968). Temporal word spacing and the perception of ordinary, anomalous, and scrambled strings. *Journal of Verbal Learning and Verbal Behavior*, 7, 154-157.
- Miller, G. A. and Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 2, 217-228.
- Millotte, S., Wales, R., and Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, 22(6), 898-909.
- 村尾玲美 (2006). 「ゲーティング法を応用した英語リスニング能力の要因分析」*STEP Bulletin*, 18, 61-76.

- O'Connell, D. C., Turner, E. A. and Onuska, L. A. (1968). Intonation, grammatical structure, and contextual association in free recall. *Journal of Verbal, Learning and Verbal Behaviour*, 7, 110-116.
- Peters, A. (1985). Language segmentation: Operating principles for the perception and analysis of language. In Slobin, D. I. (ed.) *The crosslinguistic study of language acquisition*, 2, 1029-1067. Lawrence Erlbaum Associates.
- Soderstrom, M., Seidl, A., Nelson, D. G. K. and Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, 49, 249-267.
- 杉浦正利 (2011) 「言語習得研究のための学習者コーパス」藤村逸子・滝沢直宏 (編) 『言語研究の技法』ひつじ書房. pp.123-140.
- Trehub, S. E., Bull, D. and Thrope, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55, 821-830.
- Trehub, S. E. and Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology*, 43 (2), 217-229.
- Valian, V. and Levitt, A. (1996). Prosody and adults: Learning of syntactic structure. *Journal of Memory and Language*, 35, 497-516.
- Weinert, S. (1992). Deficits in acquiring language structure: the importance of using prosodic cues. *Applied Psychology*, 6, 545-571.



Appendix 1.<sup>3</sup>

## 条件別の品詞連鎖および単語連鎖

## 【① NNS・NS ともに高頻度の品詞連鎖】

NN IN DT JJ NN	alcohol on a daily basis
NN IN DT JJ NN	crew from a local station
DT NN IN DT NN	the purpose of this essay
DT NN IN DT NN	the end of the day
DT JJ NN IN NN	a basic knowledge of money
DT JJ NN IN NN	a great deal of time
JJ NN IN DT NN	new use of the language
JJ NN IN DT NN	particular position in the world
IN DT NN IN NN	as a source of entertainment
IN DT NN IN NN	as a way of life
DT NN IN JJ NN	a gun in real life

## 【② NS のみ高頻度の品詞連鎖】

IN VVG DT JJ NN	on becoming the first person
IN VVG DT JJ NN	in learning a foreign language
IN DT NN TO VV	with the ability to obtain
IN DT NN TO VV	in an effort to help
TO VV PP\$ JJ NN	to end their own life
TO VV PP\$ JJ NN	to respect their private lives
TO DT NN IN NN	to a group of people
TO DT NN IN NN	to a lack of competition
DT JJ CC JJ NN	a long and difficult process

## 【③ NNS のみ高頻度の品詞連鎖】

IN VV IN DT NN	to look for the answer
IN VV IN DT NN	to write about the education

<sup>3</sup> 公開中の NICE-PARSED は TreeTagger によって機械的に品詞タグ付けがされている。杉浦 (2011) によれば精度はおよそ 97% であるが、3% ほどの誤りも残っている。本研究では公開されているタグ付けデータをそのまま使用している。

JJ NN CC JJ NN	modern world or future world
JJ NN IN JJ NN	fourth grade to six grade
PP MD VV IN VV	we would like to disagree
PP MD VV IN VV	I would like to write
PP MD VV DT NN	we can reduce the number
PP MD VV DT NN	I should pay a lot
PP VBZ JJ IN VV	it is possible to say
PP VBZ JJ IN VV	it is interesting to speak

【④ NNS・NSともに低頻度の品詞連鎖】

VVP RB VV DT JJ	do not think the teenage
VVP RB VV DT JJ	do not start any new
IN DT NN JJ IN	in a dump full of
IN DT NN JJ IN	in a country obsessed with
IN PP\$ NNS IN VVG	for our bodies in playing
IN PP\$ NNS IN VVG	with her friends as getting
DT NN CC PP VVP	the paper and they look
DT NN CC PP VVP	the future and they interact
RB VV RP IN DT	probably look down on the
NNS IN PP\$ NN IN	mandates in their teaching of