

日英語パラレル対話コーパスにおける名詞句の連鎖パターンの語用論的考察

吉田悦子

三重大学人文学部

tantan@human.mie-u.ac.jp

1. はじめに

談話の中で話題を担う談話要素がどのような指示表現によって言及され、情報伝達をおこなうための指示体系を構築しているのかについての実証的な研究は着実に発展してきた(Prince1981, Gundel et al. 1993)。しかし、異言語間および対話的談話のジャンルにおいては未だ説明されていない現象が多い。本稿では、指示表現の選択と分布に関して談話の展開に応じた話題要素の推移が談話の整合性の形成にどうかかわっているのかについて、語用論的視点から説明を与えることを試みる。日英語のパラレル対話コーパスを利用して、話題要素の導入から定着するプロセス、次の話題要素へと推移するパターンの考察によると、日英語の指示表現の形式的違いにもかかわらず、談話の展開の仕方には類似性が見られる。さらにセンタリングモデルを応用して複数の遷移パターンを比較対照しつつ、使用される指示表現の対応関係を考察すると、談話の展開に応じて日英語では名詞句を中心とした連鎖パターンが談話の整合性に貢献している可能性がある。談話の展開に応じた名詞句の役割にはどんな普遍的な性質があるのかについても示唆したい。

2. 研究方法

2.1 対話データについて

対話データには地図課題対話による日本語と英語のパラレルコーパス各8対話を使用した。地図課題は二人の実験参加者により共同で達成される課題である。二人の実験参加者に課される課題は、相手の地図が見えないように向かい合い、お互いに会話を交わしながら、情報提供者の地図上の経路を情報追隨者の地図上に再現することである。なお、情報提供者の地図には経路以外に出発地点と目標地点、そしていくつかの目標物とその名称がえがかれている。一方、情報追隨者の地図には、出発地点と目標物およびその名称だけが描かれており、目標地点と経路は描かれていない。二つの地図は完全に同一ではなく、その違いについては課題の遂行過程で解決すべき問題の一つとされる。(アンダーラインは著者。詳

細については吉田 2002 を参照)

2.2 分析方法

センタリング理論(Grosz et al. 1995, Walker et al. 1998)は、談話単位内部での局所焦点の移り変わりに注目して談話の整合性を予測しようとする談話モデルである。この遷移(transition)パターンを利用して日英語の指示表現の分布を分析してみる。分析の内容は:話題の焦点(Cb)と遷移パターンの分布(3節)、Cbの遷移パターンにおける指示表現の分析(4節)、遷移パターンの連続の傾向性および指示表現パターンの分析(5節)の順序でおこなう。

3. 話題の焦点(Cb)と遷移パターンの分布

まず、地図課題対話における談話要素のうち話題の焦点(center: Cb)とよばれる要素について分析する。現在の発話におけるCbの一つ前の発話(Cf)と一つ後の発話(Cp)の状態について調べ、局所的な焦点についての遷移パターンの分布を示す。英語のCbおよびno Cb(CbでないNULL要素)の総数572に対して、日本語のCbおよびno Cbの総数は634であった。このCbの分布を4つの遷移(Continue: CON, Retain: RET, Smooth Shift: SMOOTH, Rough Shift: ROUGH)およびno Cbを加えてその分布をあらわしたのが次の図1である。

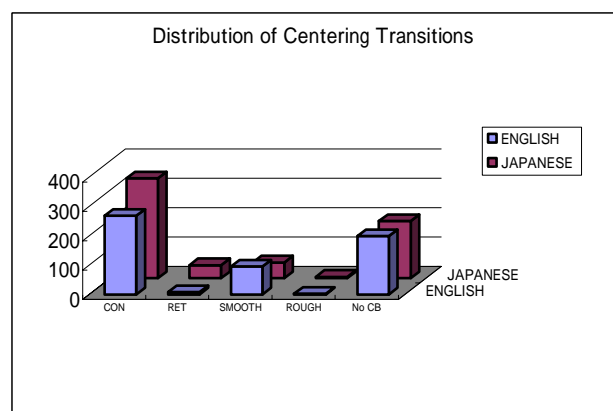


図1: センターの遷移の分布

日英語ともに CON の頻度 が最も多く、次いで No

Cb がほぼ同数でそれに次いでいる。結束性の強さに直接影響する CON の連続が日英語ともに主流であることがわかる。もっとも全体の発話のほぼ3分の1は直前の発話に現れない談話要素を含んでいることになるが、全体の談話の整合性が損なわれるわけではない。なぜなら、談話要素が全くの新出の要素か、あるいは一つ前の発話に含まれていない既出の談話要素であるかどうかは、さらに談話単位を設定して大局焦点を考察する必要があるからである (Walker 1998, 2000)。また、英語では SMOOTH が CON に次いで多く、RET はわずかであるが、日本語では SMOOTH, RET がほぼ同数である。ともに ROUGH の頻度はわずかであり、無視しても良い数値といえる。この分布を日英語それぞれ全体の分布としてまとめたのが以下の図2と図3である。

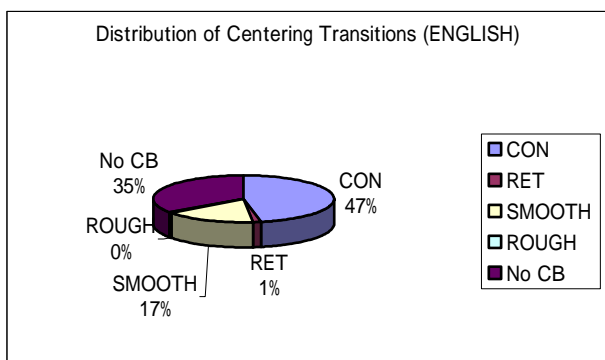


図2：遷移パターンの分布（英語）

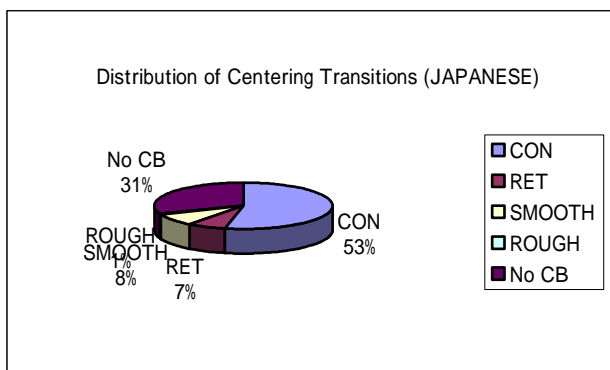


図3：遷移パターンの分布（日本語）

日英語両方ともに共通して、CON の遷移は全体のおよそ半分を占めている。このことは最も結束性に貢献すると考えられる遷移である CON によって対話の全体の整合性も支えられていることが予測される。この結果は、先行研究である Yamura-Takei(2005)や Poesio et al.(2004)とほぼ並行しているといえる。

4. Cb の遷移パターンにおける指示表現の分布

3で観察されたように、英語においては CON の遷移は全体の47%であった。各遷移の指示形式の分布について英語はゼロ代名詞、代名詞、名詞句、指示代名詞に分類してその頻度を分析してみる。

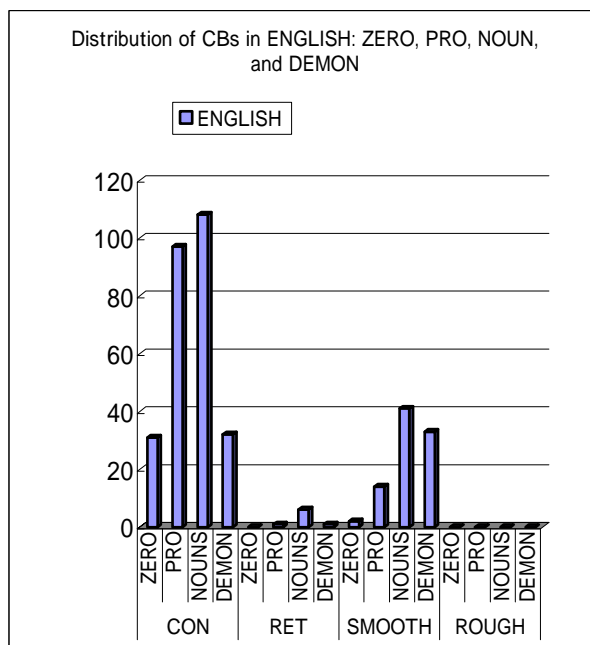


図4：各遷移における指示表現の分布（英語）

英語の CON において最も多い指示形式は名詞句で40.30%、次いで代名詞、ゼロ代名詞、指示代名詞の順である。ゼロ代名詞と代名詞を合わせると名詞句をやや上回るが、名詞句の多用は予想外の結果である。No Cbを除いてあるので、初出の名詞句はここには含まれておらず、すべて既出の名詞句である。

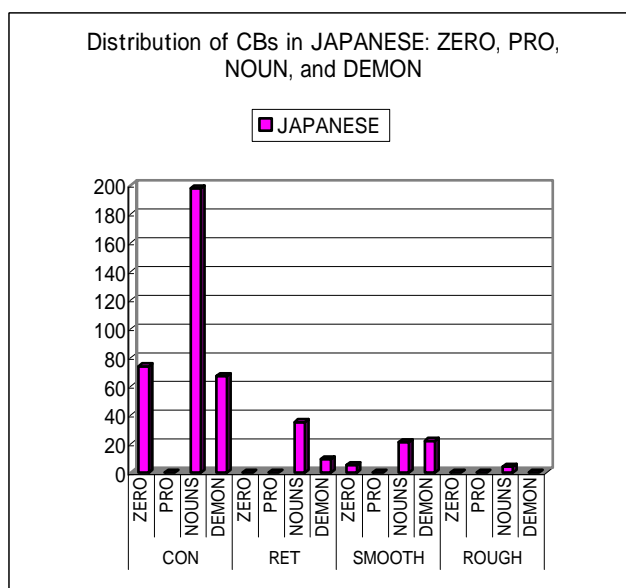


図5：各遷移における指示表現の分布（日本語）

日本語の分類も英語同様だが代名詞は生起しない。日本語の代名詞はゼロ代名詞に取って代わられるが、CON 遷移での主流となっている指示表現は名詞句で 58.41%であり、ゼロ代名詞の 21.83%のおよそ 2.68 倍となっている。英語同様、No Cb を除いてあるので、初出の名詞句はここには含まれておらず、すべて既出の名詞句であり、日本語名詞句の割合は英語のおよそ 1.5 倍である。

英語と日本語のそれぞれの Cb の頻度分布を重ねたグラフを以下に示す(図 6)。

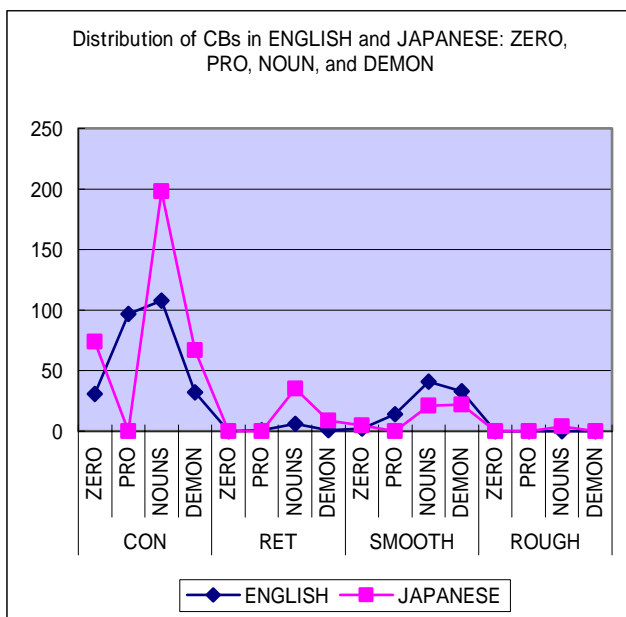


図 6：各遷移における日英語の指示表現の分布

CON 遷移に注目して、Cb の連鎖パターンを観察すると、英語では代名詞・ゼロ代名詞と名詞句とに分担されて連鎖が形成される傾向があるのに対して(例 1)、日本語ではゼロ代名詞の一時的な連鎖と名詞句の継続的な連鎖とが分化して生じている傾向が読みとれるのではないだろうか(例 2)。

(例 1) (English: L1eq4c2)

TA109: And then you're turning up towards this grass

TB110: And going round it

TA111: Yeah round the left hand side [Ø] and then over the top [Ø]

TB112: Right

TA113: And then stop when you get to the/

TB114: Edge of it

TA115: Edge of it

TB116: Okay

TA117: And then do you have a house with on it?

Just up from the grass on the right

TB118: No

(例 2)(Japanese: da)

10F: たきがありますかさいしょとおってきた

10G: や[Ø]ないですね

10F: たきよりもさらににしがわにいってことで
すか

10G: たきよりもさらにやいやいやひがしがわでと
まっています

10F: あ ひがしがわでとまっています

10G: んあはいわかった

10F: だいたいきたにありますがたきって

10G: はい[Ø]きたにありますがだいたい

日本語では Cb である「たき」は存在への問いに答える文脈ではゼロ代名詞に、「たき」の場所についての説明の文脈では裸名詞が用いられている。話題になっている要素は同じでも談話文脈に応じて表現が使い分けられている可能性がある。

5. 遷移パターンの連続の傾向性および指示表現パターンの分析

最後に 3 でとりあげた 4 つの遷移パターンの組み合わせから連続した複数の遷移パターンを比較対照して、その中で使用される指示表現の分布を見てみる。どの遷移パターンの組み合わせが談話の展開に好んで用いられ、各遷移パターンにおいてどんな指示表現が生起しているかを観察することで、談話の整合性に関連のある指示の流れを予測することができる。¹ 対話的談話の性質上、Cb になる要素の変化は自然であり、話題の焦点は常に遷移の変化を伴う。したがって、「NULL-CON-RET-SHIFT-CON」は理想的な CENTER 移行の流れであると考えられる(竹井他 2005)。つまり、「NULL-CON」は新しく導入された要素が Cb として定着する状態、「CON-CON」は Cb の高い連続性、「CON-RET」は後続の要素が Cb となる変化を予測させ、

¹ Yamura-Takei (2005), 竹井他(2005)では 11 の遷移の連続パターンが指摘され、本稿でもそれに準拠した。ただし、NULL-SHIFT については理論上存在しないとされているが、新出の談話要素が後続発話で表現を付与されて拡張した形式で言及される場合(例:「こやのようなもの」(NULL) --> [こやのうえにちょっとはたがたっているようなあんなえ (SHIFT) [bd]])についてこの連続パターンを認めることにする。

「RET-SHIFT」はその後続の要素が新しいCbに交替した状態、「SHIFT-CON」も新たなCbの定着として妥当な変化といえる。この予測の妥当性は、以下の図7と図8のように日英語それぞれに反映されている。

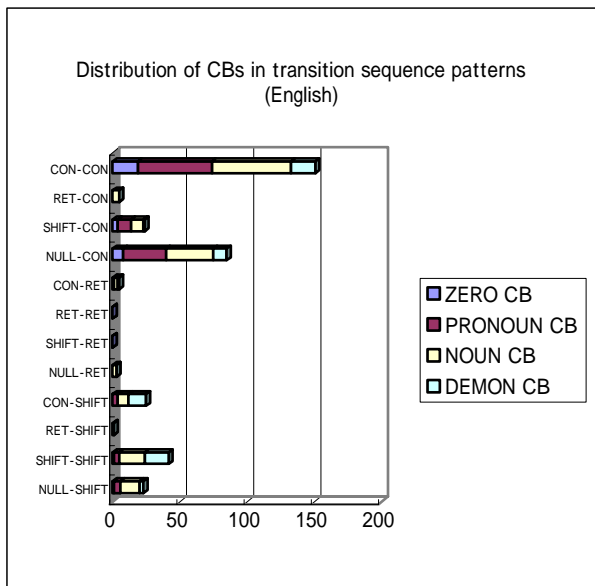


図7：遷移パターンの連続と指示表現（英語）

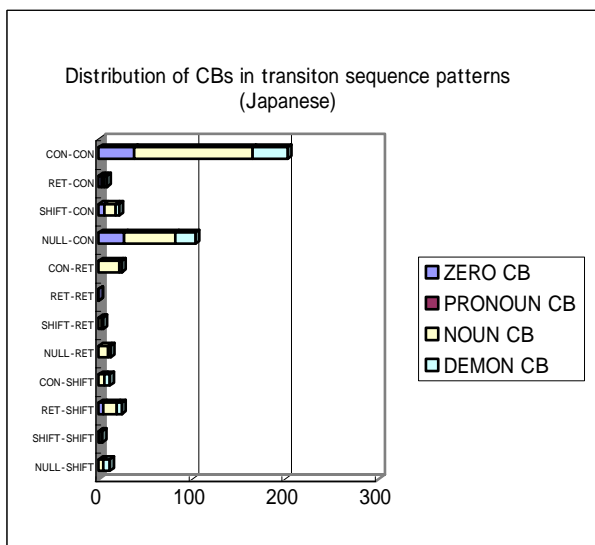


図8：遷移パターンの連続と指示表現（日本語）

CON 遷移中心のパターンに注目すると日英語の分布はよく似ている。NULL-CON、CON-CON、SHIFT-CON の好ましい遷移パターンの優勢に加え、CON-SHIFT や SHIFT-SHIFT のように推論を要するパターンも見られる。たとえば、SHIFT-SHIFT は好ましくない流れと解釈されがちであるが、実際には談話要素の言い換えや付加的表現として実現さ

れている場合もあり、整合性が著しく損なわれているわけではない。とりわけ英語データではCbとして定着するまでに複数の形式がやりとりされている。

6. まとめ

日英語の指示表現の形式的違いにもかかわらず、談話の展開の仕方には類似性があり、名詞句を中心とした連鎖パターンが談話の整合性に貢献していることが明らかになった。そして指示表現の分布状況から、談話における名詞句は遷移パターンのタイプによって異なる役割をもっていることが実証された。すなわち、NULL や SHIFT においては表現拡充的な完全名詞句として導入され、CON においては裸名詞として代名詞同様に話題の連鎖を形成する役割を担っているといえる。日英語それぞれの遷移に生起している指示詞の役割についてはまた今後の課題としたい。

参考文献

- Grosz, Barbara, Aravind Joshi, and Scott Weinstein. (1995) Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21/2, 203-225.
- Gundel, J. K., N. Hedberg and R. Zacharski (1993) "Cognitive status and the form of referring expressions in discourse." *Language*, 69, 2: 274-307.
- Poesio, Massimo, Rosemary Stevenson, Barbara Di Eugenio, and Janet Hitzeman. (2004) Centering: A parametric theory and its instantiations. *Computational Linguistics*, 30/3, 309-363.
- Prince, E. (1981) 'Toward a taxonomy of given-new information' In P. Cole (ed.), *Radical Pragmatics*. New York: Academic Press. 223-56.
- 竹井光子, 相沢輝昭, 藤原美保 (2005)「コーパスとしての教材：語学教師のための分析とツール」言語処理学会第11回年次大会発表論文集
- Walker M. A., A. K. Joshi and E. Prince (eds.) (1998) *Centering Theory in Discourse*. Oxford: Clarendon Press.
- Walker, M. A. (1998) "Centering, Anaphora Resolution, and Discourse Structure." In Walker M. A., A. K. Joshi and E. Prince (eds.). 401-436.
- Walker, M.A. (2000) 'Toward a Model of the Interaction of Centering with Global Discourse Structure' *Verbum*.
- Yamura-Takei, Mitsuko. (2005) *Theoretical, Technological and Pedagogical Approaches to Zero-Arguments in Japanese Discourse: Making the Invisible Visible*. Doctoral thesis, Hiroshima City University.
- 吉田悦子(2002)「日本語名称なし地図課題対話コーパスの概要と転記テキストの作成：報告」『人文論叢』第19号. 241-249.