

英語教育のための照応詞使用についての調査

谷村 緑 竹内 和広 井佐原 均
独立行政法人 情報通信研究機構 (NICT)
{mtanimura, kazuh, isahara} @nict.go.jp

はじめに

計算言語学では、新聞記事などの大量データから記述的知識をコンピュータに学習させ、言語解析・生成の精度を向上させるという手法を用いることが多い。しかし、教育の分野では、学習者の誤りが非常に複雑で多様なので、人手で詳細に誤りを記述し、それを基にモデル化する必要がある。

本稿では、焦点¹という概念を用いて整合性 (coherence) の良さを扱うことができるセンタリング理論を分析手法に用いた学習者の照応詞使用の調査結果から英語教育・学習における学習者の話し言葉コーパスのタグ体系としての検討を行う。

1. 従来の研究

照応表現を含め、結束関係を示す単語レベルの関係から文間のつながりについて記述分類を行っている先駆的研究には、Halliday and Hasan [2]²がある。彼らは、結束表現が整合性を示す指標であると捉える。しかし、結束性が低くとも整合性の高い談話が存在することから、局所的で表面的な関係だけでは整合性は十分に説明できない (Trabasso, Suh and Payton [3])。

Chafe (ed.) [4] や Givón [5] をはじめとする心理学的な処理を考慮する研究者は、産出者が用いる照応表現が、談話文脈に既に導入されている聞き手の知識を喚起し、談話のまとめりつまり整合性や聞き手の意味理解を導くと考える。しかし、

これらの従来の研究は、談話は整合性の高いものであるという前提に立っており、整合性の低い発話、言い換えると、伝わりにくい談話という問題を扱うことはできない。

2. センタリング理論

センタリング理論は、談話の処理過程において連続する二発話に見られる照応表現の表層の形式選択 (名詞/代名詞/ゼロ代名詞) と整合性について記述するモデルである (Walker, Joshi and Prince (eds.) [6]; 石崎・伝 [7])。この理論の主要概念であるセンタ (中心) は、何について (aboutness) の発話であるかをモデル化するための Grosz のいう焦点に相当する概念と考えられる (Grosz, Joshi and Weinstein [8])。特に、発話と別の発話を結び付けている対象を参照表現と関係している名詞句に限定しているのが特徴である。

センタリング理論では、整合性を、談話の順序 (sequence) を解釈する際の聞き手の推論の負担を反映するものとする。ある談話の最小単位である発話において代名詞化された対象が、その直後の発話においても文法面で上位に位置すれば、それは聞き手にとって最も予測可能な談話のつながりであるとされる。つまり、代名詞化された主語が後続の発話の主語と同一指示の場合に最も好ましい談話のつながりを生じるため、そのような談話は労力を要せずに理解できる。従って、ある中心から別の中心に遷移するよりも、より整合するといえる。

指示対象となり得る候補の探索においても、最も焦点の当たっている顕在性の高い名詞句が優先される。これは中心と結びつく代名詞に対する処

¹ Focus “refers to the influence of a listener’s memory for the linguistic form of an utterance (the actual words and the syntactic structure) on his interpretation of a subsequent utterance” [Grosz, [1]:5]

² 結束表現とは、整合性が達成されるための最も頻繁に言及される要素間の意味的な関係とされる。

理の優先性を示唆し、代名詞が中心となる対象を指示する場合の方が、中心でない対象を指示する場合よりも、より速く理解されると考えられる。

中心は、以下のように定義される。

(1) 中心の定義

前向き中心 (forward-looking centers) (Cf と略す)	一つの発話(U)内に示される候補となる全ての対象
後ろ向き中心 (backward-looking center) (Cb と略す)	現発話内の話の中心となる要素
優先中心 (preferred center) (Cp と略す)	一つの発話内に示される全ての対象(Cf)の中で最上位に位置する要素

(Grosz, Joshi and Weinstein 1995:208)

談話単位中の各発話 U に示される全ての対象を、前向き中心 (forward-looking centers) Cf (U) とよぶ。現発話内の話の中心となる要素が、後ろ向き中心 (backward-looking center) Cb (U) とみなされる。先頭以外の各発話 U には関心の中心である対象が一つ存在すると仮定する。Cb は先行発話で最も序列が高かった Cf で、現在話の中心になっている要素をいう。Cf の要素は Cf ランキングによって順序づけられており、Cf の中で最上位に位置する要素を優先中心 (preferred center) Cp とよぶ。Cb が現在の発話の中心を表すのに対して、Cp は次の発話の中心に関する予測を表す。

一般に顕在性の高いものほどトピックになりやすいと言われているが、表層情報、つまり、主格や目的格などの統語的位置に基づく序列 (Cf ランキング) が中心を決定する際に利用される。

(2) Cf ランキング

subject > indirect object > direct object > object
of preposition

また、Cb の遷移には以下の優先順位がある。

(3) CONTINUE > RETAIN > SMOOTH-SHIFT
> ROUGH-SHIFT

(4) 中心の遷移パターン

	Cb(Ui)=Cb(Ui-1) Or Cb(Ui-1)=不定	Cb(Ui)≠Cb(Ui-1)
Cb(Ui)=Cp(Ui) Cb(Ui)≠Cp(Ui)	CONTINUE RETAIN	SMOOTH-SHIFT ROUGH-SHIFT

CONTINUE は、前発話から引き継がれた中心が次の発話でも同じであることを示し、この遷移が続くと整合性が高いと判断される。RETAIN は、前発話の中心が次の発話で移動すること、SMOOTH-SHIFT は、前発話から引き継いだ現在の中心が次の発話で異なる中心に移動し、新しい中心が続くことを示す。ROUGH-SHIFT は、前発話と現在の中心が異なり次の発話でも異なる中心に移動することを示す。

3. 調査

第一筆者の谷村 [9] は、昔話「桃太郎」のストーリーテリング³における照応詞使用の調査を行い、日本人英語学習者 (初級、中級、上級 各 10 名) と英語母語話者 10 名 (以降NS) との間に、①整合性の構築、②照応表現の選択のそれぞれの観点において、差があるかどうか分析を行った。

①整合性の構築に関しては、表 1 より、初級に CONTINUE の選択が多く、逆に、上級に SMOOTH-SHIFT, ROUGH-SHIFT の選択が多いことが示された⁴。

³ ナラティブは大きく報告 (reporting) とストーリーテリング (story-telling) に二分されるが (Östman [10])、国を問わず相対的に共通した使用域で、比較的引き出し (elicit) やすい (Tannen [11]) ことから、資料収集に適切であると判断した。典型的な物語 (開始部、展開部、終了部があり、問題解決型である) である桃太郎を選んだ。これは初級にもなじみが深く、又、同じ題材を使用することで統制を取ることができると考えたためである。

⁴ 各レベルそれぞれ 10 名分を総合した観測度数を示す。

表1 レベル別の遷移パターン⁵

遷移パターン	C	RE	S	RO	遷移なし
初級	119	20	18	9	29
中級	117	38	42	23	33
上級	110	41	60	28	35
NS	274	83	97	18	70

たとえば、初級では、中心の移動が表現できず、場面を飛ばすという方略により移動を避ける例もみられた(例(5)参照)。

(5) 初級 06⁶

- 6.⁷ and and the baby was growing up, the healthy boy,
 7. and someday the boy [P] was heard [P] the [P]
 the demon,
 8. and er [P] Momotaro go to Onigashima,
 9. on the way, the Momotaro meets the dog,

これらの結果は、初級では、中心の移動があまり生じなかった、つまり、単調な談話展開しかできないこと、上級では、中心の移動が可能になった、つまり、話題の推移を意識的にコントロールした中心の移動が可能であることを意味している。つまり、これは、上級者になるほど、整合性の高い談話を築くことができなくなるということではなく、逆に、センタリング理論では学習者の整合性構築の説明ができないことを示唆している。

② 照応表現の選択の観点に関しては、CONTINUE と RETAIN の遷移において、NS では代名詞の選択率が高いのに対し、学習者では名詞の選択率が高いことが明らかになった。先の例(5)と以下の例(6)では CONTINUE の遷移が続くが、名詞表現の使用に大きな違いがみてとれる。

⁵ C, RE, S, RO,は、遷移パタンの頭文字である。

⁶ 被験者番号を示す。

⁷ 文頭の数字は発話番号、[p]は長いポーズ、,は短いポーズを示す。

(6) NS 01

22. on the way to fight the demons, Momotaro found,
 he found, oh, he met en ah a dog,
 23. Ø gave him some food,
 24. and Ø be [P] he be befriended the dog [P]
 25. then slowly his entourage grew
 26. and he also [P] met a monkey,
 27. and Ø made friends with a monkey as well [P]

また、例(7)の RETAIN の遷移でも学習者は、その遷移にあたる発話番号 15,16,17 で、新情報を文頭に持ってくるため、名詞表現を多用していることがわかる。

(7) 初級 04

14. Momotaro gave gave it [P] KIBIDANGO
 15. a dog [P] was take Momotaro
 16. a monkey was taken Momotaro, too
 17. KIJI a bird taken Momotaro, too

一方、NS は、例(8)の RETAIN の遷移に当たる発話番号 8 と 11 で、代名詞(she, they)を使用し、焦点化している。

(8) NS04

8. she showed him the large peach [P]
 9. all of a sudden, the large peach split in half,
 10. and out came a crying little baby [P]
 11. they decided to adopt the baby,

センタリング理論の考え方からいえば、焦点化されていない対象は、名詞で再導入されるはずであるが、本調査結果から考えると、この予測は、学習者の発話には当てはまるが、NS の発話には当てはまらないことになる。

4. まとめ

調査結果より、学習者、特に初級で、CONTINUE

の遷移の連続が多くみられ、談話主題の遷移方法が十分学習しきれていないことが明らかになった。英語教育の観点としては、学習者への教示が必要であると考え。また、初級に比べて、上級、NSに、CONTINUE 以外の遷移パターンが多くみられたことから、CONTINUE の連続・不連続で整合性の良さを説明するというセンタリング理論の課題が明らかになった。自然な談話主題の遷移を説明するためには、更に、定量的なデータ検証を行い、理論を精緻化していく必要がある。この他にも、センタリング理論は、対象が照応詞によって具現される中心に限定されていること、前向きの焦点が統語情報だけで決定されること、二文間の遷移のみしかモデル化されていないことなどの課題を抱えている。談話主題と参照との関係の精緻化、談話主題確立のメカニズムの調査も必要である。

5. 今後の課題

本調査では、限定された学習者の遷移パターンしか見ることができなかった。また、独話を扱ったため、対話での中心の遷移について明らかにならなかった。今後、スピーキングテストに基づいてレベル分けされている NICT JLE コーパス (旧 SST 学習者コーパス) を利用して、タグ体系を設計し、その妥当性を量的に検討していきたい。既に文法の誤りが付与されているので、言語運用の特質 (適切/不適切) のモデル化は、高次タグの決定、付与を可能とし、談話タグ体系の開発の一助となると考える。

また前向きの中心を再考し、新たな方向性を探りたい。中心の優先性に影響を与える要素は、統語情報だけではなく、主節・従属節の述語、アスペクト辞、時制なども深くかかわっており、それらの特徴を決定する必要がある。そのためには語彙知識との連携が不可欠であろう。

参考文献

- [1] Grosz, B. J. 1977. *The representation and use of focus in dialogue understanding*. Technical Report 151, Artificial Intelligence Center, SRI International, Menlo Park, California, July.
- [2] Halliday, M. A. K. and R. Hasan 1976. *Cohesion in English*. London: Longman.
- [3] Trabasso, T., Suh, S. and P. Payton 1995. 'Explanatory coherence in understanding and talking about events.' In M. A. Gernsbacher and T. Givón (eds.), *Coherence in Spontaneous Text*, 169-214. Amsterdam: John Benjamins Publishing Company.
- [4] Chafe, W. L. (ed.) 1980. *The Pear Stories*. Norwood, N. J.: Ablex Publishing Corporation.
- [5] Givón, T. 1983. *Topic Continuity in Discourse*. Amsterdam: John Benjamins Publishing Company.
- [6] Walker, M. A., Joshi, A. K. and E. F. Prince 1998. *Centering Theory in Discourse*. Oxford: Clarendon Press.
- [7] 石崎雅人・伝康晴 2001. 『言語と計算－3 談話と対話』 東京大学出版会
- [8] Grosz, B. J., Joshi, A. K. and S. Weinstein 1995. 'Centering: A framework for modeling the local coherence of discourse.' *Computational Linguistics* 21 (2), 203-225.
- [9] 谷村緑 2003. 「英語学習者のストーリーテリングにおける指示表現の実証研究：センタリング理論を分析の手がかりとして」『JACET 関西紀要』第7号
- [10] Östman, J. 1997. 'Coherence through understanding through discourse patterns; focus on news report.' In W. Bublitz, U. Lenk, and E. Ventola (eds.), *Coherence in Spoken and Written Discourse*, 77-100. Amsterdam : John Benjamins Publishing Company.
- [11] Tannen, D. 1982. 'Oral and literate strategies in spoken and written narratives.' *Language* 58 (1), 1-21.