

ゲーム理論に基づく中心化理論の再定式化と日英コーパスを用いた統計的検証

白松 俊[†] 駒谷和範[†] 尾形 哲也[†]
橋田 浩一^{††} 奥乃 博[†]

中心化理論 (Grosz et al. 1995) は注意状態・照応・結束性を扱った談話モデルであり広く研究されているが、従来の中心化理論はヒューリスティックかつ反証不能な定義に基づいており、定量的なモデルではなかった。われわれ (2005) は Hasida (1996) の意味ゲームを導入して中心化理論を定量的に再定式化し、日本語のコーパスで統計的に検証したが、再定式化したモデルにおいても参照表現の効用の定義は十分に定量的ではないという問題点があった。本稿ではこの問題点を解決するため、発話者と受話者にとっての参照表現の認知的コストを自己情報量で定義し、それに基づいた効用の定義の妥当性を示した。また、日本語コーパスだけでなく英語コーパスにおいても統計的に検証し、提案モデルの有効性を示した。さらに、日英コーパスにおける代名詞化の傾向を定量的に比較した。

キーワード: 中心化理論, ゲーム理論, 意味ゲーム, 照応, 結束性, 認知的コスト

1 はじめに

中心化理論 (centering theory) (Grosz et al. 1995) は局所的な談話構造のモデルであり、注意の中心、照応、局所的結束性を扱った理論である。局所的な談話の流れの自然さの評価や、流れを考慮した発話の生成に有用な理論であるが、従来の中心化理論は定量的でないヒューリスティックな定義に基づいていたため、実言語データを用いた定量的な検証やパラメータチューニングは困難であった。また、中心化理論が扱う代名詞化ルールと中心遷移ルールの背後にある基本原理も明示的に示されてはいなかった。われわれ (白松他 2005) は、Hasida (1996) のアプローチを継承してゲーム理論に基づく定量的な再定式化を行い、その妥当性を日本語コーパスの上で示した。その結果、従来の中心化理論よりも定量的かつ一般的原理に基づいたモデルとなった。しかし、われわれの定式化にも依然としてヒューリスティックな要素が残っていたため、本稿ではより定量的なモデルへと改良すべく再検討する。更に、英語コーパスの上でもわれわれの定式化の妥当性を示すとともに、日英コーパスの定量的な比較を行う。

2 従来の提案モデルの問題点

われわれが提案してきた中心化理論のゲーム理論的再定式化 (白松他 2005) を簡単に説明し、従来までの問題点を述べる。

中心化理論とは (顕現性と結束性) 中心化理論 (Grosz et al. 1995) で重要な役割を果たしている概念に、顕現性 (salience) と結束性 (cohesion) がある。

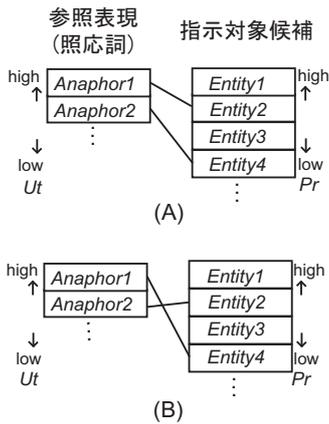
- 顕現性: 指示対象 (実体) e の顕現性とは、その文脈の当該発話で e に注がれている注意の量、つまり話題の中心らしさのこと。
- 結束性: 直前の発話からの語彙的な繋がりの強さのこと。

中心化理論では、ヒューリスティックな文法役割の順序 (Subject > Object > Indirect Object > ...) に基づいて顕現性が定義され、顕現性に基づいて話題の中心を推定する。直前発話の話題の中心は代名詞化され易く (Rule 1)、直前発話から引き続いて中心が保持され易い (Rule 2) という結束性に関するルールがある。

提案モデル (意味ゲームに基づく再定式化) 従来の中心化理論でヒューリスティックに定義されていた顕現性と結束性を、われわれは定量的に再定式化した (白松他 2005)。そのために、意味ゲーム (meaning game) (Hasida 1996)、すなわち意味伝達過程のゲーム理論的モデルを導入した。われわれは意味ゲームに基づき、顕現性と結束性を以下のように定式化した。

[†] 京都大学 情報学研究科, Graduate School of Informatics, Kyoto University

^{††} 産業技術総合研究所 情報技術研究部門, ITRI, AIST



$$Pr(Entropy2)Ut(Anaphor1) + Pr(Entropy4)Ut(Anaphor2) > Pr(Entropy2)Ut(Anaphor2) + Pr(Entropy4)Ut(Anaphor1)$$

図 1 選好 1a, 1b: 同一述語節 (発話) の中では, 対応線が交差していない照応解 (A) の方が期待効用が高く, 好まれる

- 顕現性=参照確率 指示対象 (実体) e の顕現性を, その文脈の次の述語節 (発話) で e が参照される確率 $Pr(e)$ (参照確率) として定量的に定式化した. また, コーパスから獲得したパラメタを用いて $Pr(e)$ を測定する方法を示した.
- 結束性=期待効用 まず e を参照する表現 w の効用 $Ut(w)$ を考え, $Ut(\text{代名詞}) > Ut(\text{非代名詞})$ と仮定した. その上で, 期待効用 $\sum Pr(e)Ut(w)$ が高い照応解ほど選ばれやすいこと (図 1 / 選好 1a, 1b) と, 期待効用の高さが結束性の強さに対応する (選好 2) ことを日本語コーパス上で統計的に示した.

この再定式化はゲーム理論の期待効用最大化原理から導出された仮説であり, 従来の中心化理論よりも定量的なモデルである. しかし, 未だ以下の 2 つの課題が残されていた.

- 効用の定量的定式化 コーパス上での検証 (白松他 2005) において, $Ut(\text{代名詞}) > Ut(\text{非代名詞})$ の仮定を満たす値として $Ut(\text{代名詞}) = 2$, $Ut(\text{非代名詞}) = 1$ を用いたが, これは先験的に決めた値であり, 経験的根拠が無かった.
- 他言語での検証 日本語コーパスの上では提案モデルの妥当性を示したが, 他の言語では未検証である.

よって本稿では参照表現の効用を量的に定式化し, 更に英語コーパスでも経験的検証を行う. 効用の計測や提案モデルの検証には以下の日英コーパスを用いる.

- 日本語コーパス: 毎日新聞 1356 記事から成る照

応タグ付きコーパス (白松他 (2005) でも使用したもの)

- 英語コーパス: WSJ (Wall Street Journal) 2412 記事から成る照応タグ付きコーパス

3 参照表現の認知的コストと効用の定量的定義

$Ut(\text{代名詞}) > Ut(\text{非代名詞})$ という仮定の根拠は, 「代名詞の方が発話/筆記/聴き取り/読み取り等の表層的処理のための認知的コストが少ない」というヒューリスティクスである.

代名詞の認知的コストが少ない原因の一つとして, その表層表現に対する「慣れ」が挙げられる. 代名詞は非代名詞よりも頻出するので話者も受話者も慣れ, 認知的コストが低下する. すなわち, 参照表現 w の認知的コストは, w の出現確率 $p(w)$ が小さいほど大きくなるように定義すれば良い. 本稿ではそのような尺度として自己情報量 $I(w)$ を採用し, w の認知的コストを以下のように定義する.

$$(w \text{ の認知的コスト}) := I(w) = -\log p(w) \quad [nat]$$

(ただし, $p(w)$ は当該述語節 (発話) において参照表現 w が照応詞として出現する確率. コーパスを用いて測定する.)

表 1, 2 は, 日英コーパスのそれぞれに出現した全ての照応詞の認知的コストを測定した結果である. 表が示す通り, 日英両コーパスにおいて

$$(\text{ゼロ代名詞} | \text{空範疇}) < \text{代名詞} < \text{その他の名詞句}$$

という認知的コストの順序が得られた. この結果は自己情報量 $I(w)$ による認知的コストの定義の妥当性を示している.

つぎに, 参照表現 w の効用 $Ut(w)$ を, 上記の認知的コスト定義に基づき以下のように定義する.

$$Ut(w) := Ut_0 - I(w) \quad [nat]$$

(ただし, Ut_0 はコスト 0 における効用の基準値)

効用 $Ut(w) > 0$ を保証するためには, コスト 0 における基準値 Ut_0 を $\max_w I(w) = -\log\left(\frac{1}{\text{コーパスの述語節数}}\right)$ よりも大きな値に設定する必要がある. 効用 $Ut(w)$ が正の値になるように基準値 Ut_0 を設定すべき理由は, 次節 選好 2 の検証にて述べる.

4 提案モデルの日英コーパスを用いた統計的検証

われわれが再定式化した選好 1a, 1b, 2(白松他 2005) を日英コーパスを用いて検証する．指示対象 e の参照確率 $Pr(e)$ は，白松他 (2005) と同じ 3 素性¹を用いた多重ロジスティック回帰で計測する．参照表現 w の効用 $Ut(w)$ は，基準値 $Ut_0 = 12$ から認知的コスト $I(w)$ を減算した値を用いる．基準値 12 を設定した理由は，本節 選好 2 の検証にて述べる．

表 1 認知的コストの低い参照表現と種別毎の平均 (毎日新聞コーパス)

参照表現 w	出現確率 $p(w)$	認知的コスト $I(w)$ [nat]
(ゼロ代名詞)	2.111×10^{-1}	1.556
私	3.682×10^{-3}	5.604
その	2.846×10^{-3}	5.862
これ	2.135×10^{-3}	6.149
この	1.356×10^{-3}	6.604
日本	1.299×10^{-3}	6.646
者	1.299×10^{-3}	6.646
それ	1.220×10^{-3}	6.709
大統領	1.062×10^{-3}	6.848
米国	9.713×10^{-4}	6.937
⋮	⋮	⋮
w の種別	種別毎 $p(w)$ 平均	認知的コスト
ゼロ代名詞	2.111×10^{-1}	1.556
代名詞	1.725×10^{-3}	6.362
それ以外	1.631×10^{-5}	8.721

表 2 認知的コストの低い参照表現と種別毎の平均 (WSJ コーパス)

参照表現 w	出現確率 $p(w)$	認知的コスト $I(w)$ [nat]
(空範疇)	2.547×10^{-1}	1.368
it	4.232×10^{-2}	3.162
he	3.049×10^{-2}	3.490
they	1.850×10^{-2}	3.990
company	1.652×10^{-2}	4.103
we	1.112×10^{-2}	4.499
I	1.020×10^{-2}	4.585
U.S.	8.342×10^{-3}	4.786
you	6.357×10^{-3}	5.058
that	5.503×10^{-3}	5.202
⋮	⋮	⋮
w の種別	種別毎 $p(w)$ 平均	認知的コスト
空範疇	2.457×10^{-1}	1.368
代名詞	3.257×10^{-2}	3.836
それ以外	1.317×10^{-3}	6.632

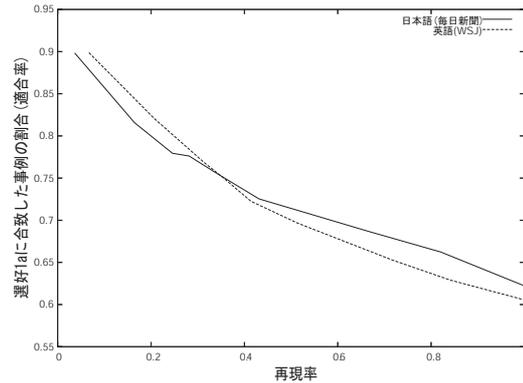


図 2 選好 1a の recall-precision 曲線

選好 1a : 「1 つの述語節が複数の照応詞を含むとき，それらの照応詞のうち，効用 $Ut(w)$ が高い照応詞の方が低い照応詞よりも参照確率 $Pr(e)$ の高い指示対象 e を参照しやすい」(白松他 2005)(図 1)

選好 1a の検証のために，コーパス中で複数の照応詞を含む述語節を集め，同一述語節内の照応詞対を作る．この照応詞対のうち，選好 1a に合致する事例の割合が高いほど選好 1a は強い予測であると言える²．ただし，照応詞対の効用の差，あるいはそれらの指示対象の参照確率の差が小さい場合，選好 1a は成り立ち難くなると考えられる．図 2 は，その差の閾値を変動させることにより描いた recall-precision 曲線である．日英両コーパスにおいて，全ての照応詞対を考慮した場合でも合致事例の割合が 60% を超え，再現率 25% 付近で合致事例の割合が約 80% まで上がった．これは両コーパスにおいて選好 1a が有効かつ妥当であることを示している．

選好 1b : 「照応詞 w の効用 $Ut(w)$ とその指示対象 e の参照確率 $Pr(e)$ の間には正の相関がある」(白松他 2005)(図 1)

$Pr(e)$ と $Ut(w)$ の相関係数は，日本語コーパスで +0.369 (95% 信頼区間 [0.356, 0.383])，英語コーパスで +0.223 (95% 信頼区間 [0.217, 0.229]) であった³．つまり，日英両コーパスにおいて選好 1b の妥当性が裏付けられた．英語コーパスにおける相関係数が相対的に低い理由は，図 3 が示唆している．同図は参照確率毎に代名詞化の割合をプロットしたもののだが，参照確率 0.75 よりも右側で英語だけが大きく下がっ

1 指示対象候補 e の最近参照箇所と当該述語節の距離 $dist$ ， e の最近参照箇所の文法役割 $gram$ ，先行文脈中で e が参照された回数 $chain$ の 3 素性．

2 選好 1a の合致事例の割合は，効用の基準値 Ut_0 の影響を受けない．
3 選好 1b における $Pr(e)$ と $Ut(w)$ の相関係数は，効用の基準値 Ut_0 の影響を受けない．

ている．これが英語における $Pr(e)$ と $Ut(w)$ の相関係数を下げた原因であろう．英語コーパス中で参照確率が 0.75 よりも大きいのに非代名詞の参照表現を使用していた事例は 11,367 例あり，そのうち 4,735 例 (41.7%) が固有名詞 (Mr. で始まる略称など) を用いていた．本稿の定義ではこのような略称の効用は低くなってしまい，改良の余地がある．

一方日本語コーパスでは，参照確率が大きい領域ではゼロ代名詞を参照表現として用いる事例が圧倒的に多かった．この差異が，日本語と英語の相関係数の差として現れたと考えられる．

選好 2 : 「期待効用 $\sum Pr(e)Ut(w)$ が大きい照応解ほど選ばれやすく，Transition の順序は期待効用の大きさに対応する」(白松他 2005)

Transition の順序とは，中心化理論で定義された 4 つの中心遷移状態の順序である．本稿では，期待効用と Transition 4 状態 (CONT.=4, RET.=3, S.S.=2, R.S.=1 の値を割り当てた) の相関係数を調べた．図 4 は，効用の基準値 Ut_0 を変動させて期待効用と Transition の相関係数への影響を調べたものである．これによると， Ut_0 が $I(w)$ の最大値よりも大きい右側の領域 (常に $Ut(w) > 0$) では，相関係数はほぼフラットである．しかし，これより左側の領域では相関係数が急激に下がるため，効用の基準値 Ut_0 は $I(w)$ の最大値よりも大きく設定すべきである．なお，日英コーパスの $I(w)$ の最大値はそれぞれ約 11.4 と約 11.8 なので，基準値 Ut_0 は日英ともに 12 に設定した．

この設定の下での期待効用と Transition の相関係数は，日本語コーパスで +0.585，英語コーパスで +0.407 であった．また，Transition 毎の期待効用の平均値も中心化理論で定義された順序と一致し，Wilcoxon の順序検定で統計的に有意であった．よって，日英両コーパスで選好 2 の妥当性が裏づけられた．

5 まとめと今後の課題

従来のわれわれの提案モデルにおいて，参照表現の効用には先験的な値を与えていた．本稿では自己情報量を用いて参照表現の認知的コストを定義し，日英コーパス上でその定義の妥当性を示し，それに基づいて参照表現の効用を定量的に定義し，提案モデルを改良した．更に，日英コーパスを用いて提案モデルの検証を行い，われわれが再定式化した選好 1a, 1b, 2 の妥当性を両コーパス上で示した．

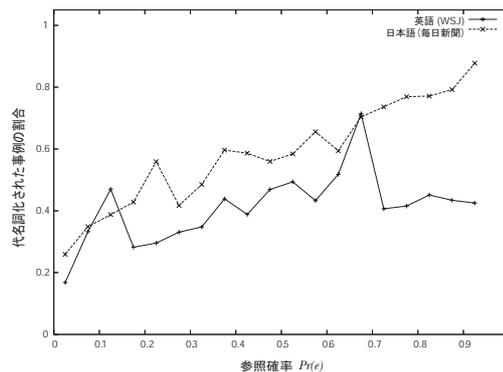


図 3 日英コーパスにおける参照確率毎の代名詞化事例の割合

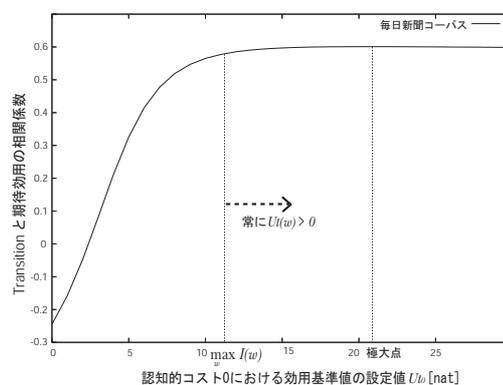


図 4 認知的コスト 0 における効用基準値 Ut_0 に対する選好 2 (Transition と期待効用) の相関係数の変化

ただし，英語コーパスの参照確率が大きい事例で多く使用されていた固有名詞の略称は，本稿の定義では認知的コストが高くなるが，実際は先行文脈の影響からコストが低くなると考えられる．そのような認知的コストの定義の改良は今後の課題とする．

謝辞 本研究で用いた GDA コーパスの作成に係わった方々に深謝する．本研究の一部は 21 世紀 COE プログラム，および学振科研費の支援を受けた．

参考文献

- Grosz, B., Joshi, A., and Weinstein, S. (1995). “Centering: A Framework for Modeling the Local Coherence of Discourse.” *Computational Linguistics*, **21** (2), 203–225.
- Hasida, K. (1996). “Issues in Communication Game.” In *Proc. of COLING’96*, pp. 531–536.
- 白松俊, 宮田高志, 奥乃博, 橋田浩一 (2005). “ゲーム理論による中心化理論の解体と実言語データに基づく検証.” *自然言語処理*, **12** (3), 91–110.