

# 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

白松 俊<sup>†</sup>      宮田 高志<sup>††</sup>  
奥乃 博<sup>†</sup>      橋田 浩一<sup>†††,††</sup>

中心化理論 (centering theory) は、注意の中心、照応、結束性の間の相互作用を説明する談話構造の理論である。しかし、照応現象の背後にあるはずの基本原則を明らかにするものではない。また、中心化理論で重要な役割を担う顕現性 (salience) が、客観的に計量可能な尺度として定式化されていないという問題もある。一方、Hasida ら (1995, 1996) は、ゲーム理論に基づく意図的コミュニケーションのモデルとして意味ゲーム (meaning game) を提唱し、「照応等の現象はゲーム理論で説明できる」と主張しているが、この主張は実言語データに基づいて検証されていない。われわれは、顕現性を計量可能な尺度として定式化し、中心化理論の 2 つのルールに対応する意味ゲームに基づく選好を日本語のコーパスを用いて検証した。その結果、中心化理論の予測を越える部分も含めてこれらの選好が成立することがわかった。したがって、基本原則の明確さおよび予測能力の強さゆえに、中心化理論よりも意味ゲームの方が優れた作業仮説であり、この意味において、中心化理論等の照応や焦点に特化した理論は不要と考えられる。

キーワード: 中心化理論, ゲーム理論, 顕現性, 照応

## Dissolution of Centering Theory Based on Game Theory and Its Empirical Verification

SHIRAMATSU SHUN<sup>††</sup>, MIYATA TAKASHI<sup>††</sup>, HIROSHI G. OKUNO<sup>†</sup>  
and HASIDA KÔITI<sup>†††,††</sup>

Centering theory is to explain relations among focus, anaphora, and cohesion. However, it fails to address any general principle behind anaphora. Moreover, although the salience of discourse entities plays a critical role in centering theory, it is not defined as an objectively measurable quantity. On the other hand, Hasida et al. (1995, 1996) propose *meaning game* as a model of intentional communication, and claim that it derives centering theory, but this claim has not yet been verified on the basis of substantial linguistic data. In this paper, we formulate salience in terms of reference probability (as measurable quantity). Under this formulation, meaning game derives preferences subsuming two rules of centering theory. Those preferences, entailing stronger predictions than centering theory, are verified based on a Japanese corpus. Meaning game is hence a better working hypothesis than the centering theory in terms of both theoretical clarity and predictive power. Domain-specific accounts such as centering theory are probably not necessary to explain anaphora, focus, and so on.

**KeyWords:** *centering theory, game theory, salience, anaphora*

# 1 はじめに

照応現象に関する理論のうち，最も広く論じられているのは中心化理論 (centering theory) である．中心化理論は，注意の中心，照応，結束性の間の相互作用を説明している．しかし，照応現象等の背後にある基本原理を明らかにするものではない．

もし中心化理論の背後に何らかの基本原理が存在するならば，それは談話における発話者と受話者の行動決定を説明する原理であろう．その基本原理は，客観的に計量可能な尺度に基づいて述べられるべきである．しかし，中心化理論において重要な役割を担っている顕現性 (salience) という概念は，客観的に計量可能な尺度として定式化されていない．顕現性とは，人間の注意状態に関連する何らかの尺度であるが，従来研究では Cf ランキングというヒューリスティクスで近似される．本稿では，参照確率という計量可能な尺度として顕現性を定式化し，その計測手法を示す．

一方，中心化理論の背後にある基本原理の説明として，Hasida ら (1995, 1996) が提唱する意味ゲーム (meaning game) がある<sup>1</sup>．意味ゲームとは，ゲーム理論に基づいて意図的なコミュニケーションを説明するモデルであり，発話者と受話者をプレイヤーとする 2 人ゲームである．Hasida らは，顕現性を上記のように参照確率とみなし，照応詞の単純さをプレイヤーの利得の一部とみなすと，この意味ゲームモデルから中心化理論が導けることを示した．彼らはコミュニケーションの一例として特に照応を取り上げて，照応現象の説明はゲーム理論に帰着できると主張している．しかし，この主張の根拠は特定の事例に関する思考実験であり，実言語データに基づいて検証されていない．本稿では日本語の新聞記事コーパスを用いて照応の意味ゲームモデルを検証し，この主張が正しいことを示す．

## 2 中心化理論の概略と問題点

以下では，中心化理論の概略を述べ，基本原理の欠如という問題点と顕現性に関する問題点を指摘する．

### 2.1 理論の概要

中心化理論では，談話を発話 (utterance) の列  $[U_1, U_2, \dots, U_n]$  として扱う．各発話において注意が向けられている実体のことを中心 (center) と言い，発話ごとに中心が更新される．また，中心の決定の際に用いられる尺度である顕現性 (salience) は，ある文脈で具現化 (realize)<sup>2</sup>され

† 京都大学大学院 情報学研究科 知能情報学専攻, Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

†† 科学技術振興機構 戦略的創造研究推進事業, CREST, JST

††† 産業技術総合研究所 情報技術研究部門, ITRI, AIST

1 Hasida らのアプローチを最適性理論の上で発展させる試みも行われている (van Rooy 2003; Kibble 2003)

2 言語表現 (ゼロ照応を含む) によって実世界の实体を参照すること．

た実体の「目立ち具合」を表す。

中心には以下のような種類がある。

- $Cb(U_i)$ :  $U_i$  の後向き中心 (backward-looking center) . 先行文脈で具現化され,  $U_i$  でも引き続き具現化されている実体 . そのような実体が複数ある場合は,  $U_{i-1}$  において最も顕現性が高かった実体 .
- $Cf(U_i)$ :  $U_i$  の前向き中心 (forward-looking centers) .  $U_i$  で具現化された実体を顕現性の順にソートしたリスト .
- $Cp(U_i)$ :  $U_i$  の優先中心 (preferred center) .  $Cf(U_i)$  の要素のうち最も顕現性の高い実体 . 中心化理論は以下の 2 つのルールから構成される (Walker, Iida, and Cotes 1994) .

ルール 1:  $Cf(U_{i-1})$  の要素の幾つかが  $U_i$  において代名詞によって具現化されているならば, そのうちの一つが  $Cb(U_i)$  である .

ルール 2: 中心の遷移には下記の 4 種類があり, その選好の順序は

Continue > Retain > Smooth-Shift > Rough-Shift である .

	$Cb(U_i) = Cb(U_{i-1})$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	Continue	Smooth-Shift
$Cb(U_i) \neq Cp(U_i)$	Retain	Rough-Shift

ルール 1 は, 同一発話内に代名詞と非代名詞がある場合, 代名詞の方が  $Cb$  を指しやすいという選好である .

ルール 2 は, 結束性 (cohesion), すなわち発話間の語彙的つながりの強さに関するルールである . 発話  $U_{i-1}$  から  $U_i$  への中心遷移の仕方を 4 種類に分け, それらを結束性が高い順に並べた選好である .

## 2.2 問題点

中心化理論の第 1 の問題点として, 上記の 2 つのルールは, 中心, 照応, 結束性の相互作用について述べているものの, そのような現象の根底にある基本原理を説明してはいない .

また, 第 2 の問題点として, 中心化理論において本質的な役割を担う顕現性が客観的に計量可能な尺度として定式化されていない . すなわち, 中心化理論の先行研究において顕現性の意味を明確に定義した研究は無く<sup>3</sup>, 顕現性は注意状態に関係する何らかの尺度として, Cf ランキングと呼ばれるヒューリスティックな順序によって近似されてきた . しかし, そもそも顕現性の定義に上記のような不備があるので, Cf ランキングの妥当性を経験的に検証することは不可能である .

基本的に Cf ランキングは文法機能のみによって決まり, その順序は下記のように言語によって異なる .

<sup>3</sup> 文生成の研究においては, 視覚における鮮やかさや, 語自体が持つ印象の強さ, 言及された至近性などによって顕現性が決定されるとした研究例がある (Reed 2002) .

英語の Cf ランキング:

主語 > 直接目的語 > 間接目的語 > 補語 > 付属語

日本語の Cf ランキング (Walker et al. 1994):

主題 (文法 or ゼロ) > 視点 > 主語 > 間接目的語 > 直接目的語 > その他

しかし、文法機能以外の要因も顕現性の決定に関わっているとする研究もある (Strube and Hahn 1999; Reed 2002)。また、文法機能の順序によって決定できるのは同じ発話内で参照されている実体の順序のみであるが、先行詞の候補が異なる発話に分散していることも多いので、異なる発話間の実体も順序付けする必要がある。

顕現性の問題点を以下にまとめる。

- A 客観的に計量可能な尺度として定義されていないので、Cf ランキング等の妥当性の経験的検証が不可能である。
- B 顕現性に影響する要因は文法役割以外にもあるが、Cf ランキングはそれを捉えていない。
- C Cf ランキングでは直前の発話で参照された実体しか扱えない。

問題点 B に関して、Strube ら (1999) は聞き手にとっての情報の新しさという要因を導入して Cf ランキングを拡張した機能的中心化理論 (Functional Centering) を提案した。問題点 C に関して、Nariyama (2001) は、先行文脈中の名詞句から成る SRL (Salient Referent List) を発話単位毎に更新する手法を提案した。しかし、そもそも問題点 A を解決しなければこれらの試みの妥当性を論ずることも困難である。

### 3 顕現性の定式化と計測

2 節で述べたように、中心化理論では顕現性は Cf ランキングと呼ばれるヒューリスティックによって近似されるが、そもそも顕現性が客観的に計測できる尺度として定式化されていないため、Cf ランキングの妥当性を経験的に検証することは不可能である。これは、理論としての不備である。

本節では、顕現性を参照確率 (reference probability) という計測可能な尺度として定式化することを提案する。参照確率とは、実体が次の発話で参照 (具現化) される確率である。これにより、2 節で述べた問題点が解消される。

#### 3.1 計測方法

以下に顕現性の定式化と計測方法を示す。

顕現性の定式化

顕現性の定義 発話列  $[U_1, U_2, \dots, U_i]$  のどこかで実体  $e$  が参照されているとき, 発話  $U_i$  における実体  $e$  の顕現性とは,  $e$  が  $U_{i+1}$  で参照される確率 (参照確率) である.

参照確率の計測

- 発話列  $\{U_1, U_2, \dots, U_i\}$  のどこかで実体  $e$  を参照している表現  $w$  がある.
- 特徴量ベクトル  $feature(w, U_i)$  を抽出する.
- 十分に大きな言語コーパスにおいて,  $feature(w, U_i)$  と等しい特徴量ベクトルを持つ事例  $(w_x, U_j)$  を全て抽出する. そのうち,  $w_x$  が指し示す実体が  $U_{j+1}$  においても参照されている事例の相対出現頻度を計測する.
- この相対出現頻度が,  $U_{i+1}$  における  $e$  の参照確率である.
- この参照確率を,  $U_i$  における  $e$  の顕現性とする.

ここで, 以下のような具体例において「太郎君」が  $U_{i+1}$  で参照される確率の計測方法を説明する.

$U_{i-2}$ : さきほど ( $\phi_0$  ガ)太郎君 を見かけたが、  
 $U_{i-1}$ : ( $\phi_1$  ガ) 眠そうだった。  
 $U_i$  : 昨夜はとても暑かったし、  
 $U_{i+1}$ : ..... 参照確率  $Pr(\text{太郎君}, U_{i+1})$

事例の特徴量として, たとえば次の 3 素性のみを用いたとする.

- *dist*: 現発話と, 先行文脈中で実体  $e$  を最近参照した表現の発話距離
- *gram*: 先行文脈中で  $e$  を最近参照した表現に係る助詞 ( 文法役割)
- *chain*: 先行文脈中の  $e$  の共参照連鎖 (coreference chain) の長さ

このとき, この事例の特徴量ベクトルは以下ようになる.

$$feature(\text{太郎君}, U_i) = (dist = 2, gram = \text{ガ}, chain = 2)$$

このとき, コーパス中で  $feature(w, U_j) = feature(\text{太郎君}, U_i)$  となるような事例, つまり

$U_{j-k}$ :  $w_0$  \_\_\_\_\_  
 $\vdots$   
 $U_{j-1}$ :  $w$  が \_\_\_\_\_      条件 C:  $feature(w, U_j) = (dist = 2, gram = \text{ガ}, chain = 2)$   
 $U_j$  : \_\_\_\_\_  
 $U_{j+1}$ : ..... 参照確率  $Pr(w, U_{j+1}) = \frac{C \wedge (U_{j+1} \text{ で } w \text{ を参照している}) \text{ が成り立つ事例数}}{C \text{ が成り立つ事例数}}$

という表層的なパターンを持つ事例の出現頻度と, そのうちの  $U_j$  において  $w$  が参照する実体が参照される相対出現頻度を計測しておいたとすると, この相対頻度が,  $U_i$  において「太郎君」が参照される確率の近似値となる.

ただし, 実際のコーパスの事例数には限りがあり, 特徴空間上の全てのベクトルについて十分な事例数があるわけではないので, 任意の特徴量ベクトルを持つ事例の確率を外挿する必要

S V R	多重 logi- stic	<i>dist</i> <i>gram</i> <i>chain</i>	参照表現と指示対象候補の最近参照箇所との発話距離の自然対数 指示対象候補の最近参照箇所の文法機能 (助詞) (指示対象候補の先行文脈中の共参照連鎖の長さ+1) の自然対数
		<i>exp</i> <i>last_topic</i> <i>last_sbj</i> <i>p1</i> <i>pos</i>	指示対象候補の最近参照箇所の表現種別 (ゼロ/代名詞/定名詞/一般) 指示対象候補が最近のトピックであるか否か 指示対象候補が最近の主語であるか否か 指示対象候補が一人称であるか否か 指示対象候補の最近参照箇所の品詞 (名詞/述語)

表 1 本稿で参照確率の回帰分析に用いる素性

がある。そのためには、コーパス中の事例集合を用いて回帰分析を行えばよい。回帰分析のアルゴリズムについては限定しない。ただし、複数の要因を統合して参照確率を外挿するためには複数の説明変数を扱う多重回帰が可能なアルゴリズムでなければならない。本稿では、以下の2つの回帰アルゴリズムによって参照確率の計測を行う。

- 3素性による多重ロジスティック回帰
- 8素性によるSVR (Support Vector Regression)

本稿で多重ロジスティック回帰に用いる3素性とSVRに用いる8素性を表1に示す。多重ロジスティックで用いる説明変数を3素性に絞ったのは、モデルの表現能力に対して問題空間が疎になることを防ぐためである。

## 4 意味ゲーム

本節では意味ゲームの概要を説明し、意味ゲームから中心化理論のルール1,2に相当する選好を導出する。

### 4.1 意味ゲームの概要

Hasidaら(1995, 1996)が提唱した意味ゲームは、意図的なコミュニケーション(非自然的意味の伝達)のゲーム理論的な定式化である。Hasidaらは、意図的・言語的コミュニケーションの一例として特に照応現象を論じた。意味ゲームでは、発話者による意図決定および受話者による解釈の組み合わせの、コミュニケーションの成功以外の要因による期待効用は以下の式で表される。

$$\sum_{w \text{ が } e \text{ を参照する}} Pr(e) Ut(w)$$

$Pr(e)$  は実体  $e$  が参照される確率(参照確率)、 $Ut(w)$  は  $e$  を参照する表現  $w$  の効用である。

ここでは簡単のため以下のように仮定する。

- コミュニケーションは確実に成功する。すなわち、発話者の意図した意味を受話者は必

ず理解する。

- 実体  $e$  の参照確率  $Pr(e)$  は、先行文脈を含む発話者・受話者の間での共有信念に基づいて定まり、それ自身共有信念に属する。したがって、発話者、受話者双方にとって  $Pr(e)$  は等しい。
- 参照表現  $w$  が単純なほど表層的処理（発話/筆記/聞き取り/読み取り）のコストが低く、発話者、受話者双方の利得  $Ut(w)$  が高い。表現が複雑であればコストが高く、 $Ut(w)$  が低い（ただし、 $Ut(w)$  の値は発話者と受話者において異なってもよい）。

これらの仮定により、期待効用を最大化する解が発話者・受話者間で共通となる。副作用として、誤解が起こりやすいような発話はモデルの対象外となるが、文法的知識などの共有信念に基づいて理解可能な談話現象をモデル化できれば充分であり、理解不能な発話はそもそも扱う必要が無いと我々は考える。これら 3 つの仮定をおくことにより、プレイヤー間で共有された期待効用が最大になる解（発話意図と解釈の組み合わせ）が Pareto 最適解<sup>4</sup>となる。

## 4.2 ルール 1 の導出

Hasida らは、照応の意味ゲームから中心化理論のルール 1 を導いている。たとえば次のような意味的制約の影響が小さい談話においては、he が Fred を指し、the man が Max を指す場合が多い。

談話 (1)  $U_1$ : Fred scolded Max.

$U_2$ : He was angry with the man.

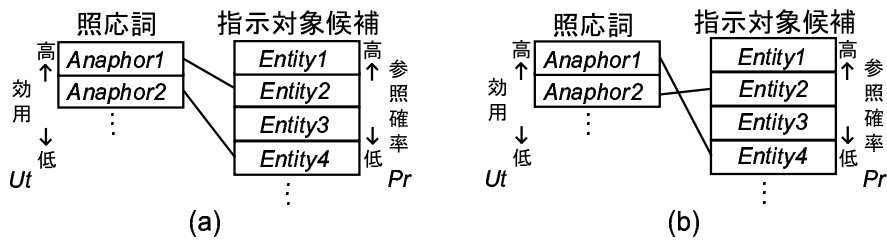
$p_1$ : Fred(主語) の参照確率 >  $p_2$ : Max(目的語) の参照確率

$u_1$ : he(代名詞) の効用 >  $u_2$ : the man(定名詞) の効用

Fred	Max		Fred	Max
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
			/	\
			\	/
</				

ルール 1	意味ゲーム
代名詞	効用の高い参照表現
非代名詞	効用の低い参照表現
顕現性	参照確率
後向き中心	参照確率の高い実体

表 2 ルール 1 と意味ゲームの対応



$$Pr(Entity2)Ut(Anaphor1) + Pr(Entity4)Ut(Anaphor2) > Pr(Entity2)Ut(Anaphor2) + Pr(Entity4)Ut(Anaphor1)$$

図 1 交差の有無

上記の Hasida らによる例証は，一般の事例では図 1(b) の解よりも (a) の解の方が選ばれやすいことを表し，以下のように記述できる．

選好 1a: 同一発話内に複数の参照表現があるとき，そのうち効用が高い参照表現が，参照確率の高い実体を参照しやすい．

また，この選好は一般<sup>5</sup>には以下の選好と等価である．

選好 1b: 参照表現の効用とその指示対象の参照確率の間には正の相関関係がある．

この意味ゲームから導かれた選好 1b は 1 つの発話の中に参照表現が 1 つしかない場合にも及ぶので，中心化理論のルール 1 よりも強い予測を導く．

### 4.3 ルール 2 の導出

ルール 2 は，中心の遷移 (transition) と結束性 (cohesion) に関する選好である．中心遷移は 2 つの条件式の組み合わせによって 4 種類に分けられ，結束性に関する優先順位がつけられる．1 つ目の条件式  $Cb(U_i) = Cb(U_{i-1})$  は，後向き中心  $Cb$  が直前発話からそのまま受け継がれていることを表す．2 つ目の条件式  $Cb(U_i) = C_p(U_i)$  は， $Cb$  が  $U_i$  中で参照されている実体のうちで

<sup>5</sup> 参照表現が 3 つ以上の場合，及びそれらが同一発話内に無い場合にも予測の範囲を広げる，という意味での一般



最も顕現性が高いことを表す<sup>6</sup>。

われわれは、ルール 1 の場合と同じく、ルール 2 における順序も発話の期待効用の高さの順序として導かれると考える。1 つ目の条件式が成り立つとき、 $C_b$  の参照確率が高くなると同時に、選好 1b の予測から  $C_b$  を参照する照応詞の効用も高くなると考えられ、したがって現在の発話の期待効用が増すからである。また、2 つ目の条件式が成り立つときも、やはり  $C_b$  の参照確率と効用が高くなり、期待効用が高くなると考えられるからである。さらに、Retain と Smooth-Shift は共に一方の条件式のみが成り立つタイプであるが、1 つ目の条件式が直前の発話から現在の発話への結束性を表す（したがって現在の発話の期待効用を直接高める）のに対し、2 つ目の条件式は現在の発話から次の発話への結束性の予測に過ぎないので 1 つ目の条件式の方が 2 つ目の条件式よりも期待効用への影響が強く、したがって Retain の方が Smooth-Shift よりも期待効用が大きくなると予想される。これらの予想が正しいならば、ルール 2 は意味ゲームに基づき以下のように一般化できると考えられる。

選好 2: 期待効用の高い解（発話意図と解釈の組み合わせ）が選ばれやすい。この期待効用の高さが結束性の強さに対応する。

本稿では、コーパス中の事例を中心遷移 4 タイプに分類し、各タイプの期待効用の平均の順序がルール 2 の順序と合致するという予想を検証する。

## 5 統計的検証

本節では、意味ゲームから導出した選好 1a, 選好 1b, 選好 2 を統計的に検証する。その際、統語構造や照応を表す GDA タグ (Hasida 1998) を人手で付与した毎日新聞の記事 1356 記事から成るコーパスを用いる。表 3 にコーパスに含まれる事例数と、正例・負例の頻度分布を示す。正例は、先行文脈で参照された実体  $e$  が次の発話  $U_{i+1}$  でも参照されている事例、負例は、 $e$  が  $U_{i+1}$  では参照されていない事例である。表 4 に参照表現の種類別の頻度分布を示す。

ところで、文を発話単位と見なす手法では、複文における文内の照応を扱うことができない。Kameyama(1998) は時制節を発話単位とする拡張を提案した。本稿では時制節か非時制節かの区別は行わず、1 つの述語が統率する述語節を発話単位と見なして統計的検証を行った。

また、Hasida らは参照表現の効用について、参照表現が単純なほど発話者・受話者の表層的処理のコストが低く、効用が高いと仮定した。将来的には参照表現のコストの低さを計量可能な尺度として定式化する必要があるが、現段階では中心化理論と同様に代名詞と非代名詞の区別だけを考慮する。代名詞の方が非代名詞よりもコストが低いので、代名詞の効用が非代名詞の効用よりも高いと仮定する。具体的な値の設定においては、代名詞と非代名詞という 2 タイプのみを仮定しているため 2 タイプの効用の値の大小だけが問題であり、少なくとも選好 1a,

<sup>6</sup> 前者の条件式は直前発話  $U_{i-1}$  と現発話  $U_i$  の結束性に対応し、後者の条件式は現発話  $U_i$  と次の発話  $U_{i+1}$  の結束性の予測に対応すると考えられる。

	事例数	割合
正例	16728	1.6%
負例	1057053	98.4%
全事例	1073781	100.0%

表 3 全事例数と正例・負例の割合

参照表現の種類	事例数	割合
ゼロ代名詞	5876	35.1%
代名詞	843	5.0%
指示詞に係る名詞句	1011	6.0%
その他の名詞句	8998	53.8%
計	16728	100.0%

表 4 参照表現のタイプ別分布

助詞	は	が	の	を	に	も	で	から	と
出現頻度	35329	38450	88695	50217	46058	8710	24142	7963	19383
参照頻度	1908	1107	1755	898	569	105	267	76	129
参照確率 (%)	5.40	2.88	1.98	1.79	1.24	1.21	1.11	0.954	0.666
助詞	その他の助詞		助詞無し						
出現頻度	512006		153197						
参照頻度	8027		1315						
参照確率 (%)	1.57		0.858						

表 5 助詞別の参照確率

1b の検証は絶対値や比の設定に影響を受けない<sup>7</sup>。本稿では、代名詞（ゼロ代名詞含む）の効用の値を 2、非代名詞の効用の値を 1 と仮定して期待効用を計算し、検証を行う。

## 5.1 回帰分析による参照確率の計測

ここでは、回帰分析による参照確率の計測方法として、多重ロジスティック回帰を用いる方法と、SVR を用いる方法について述べる。

回帰分析に用いる説明変数には、助詞（文法役割）を表すパラメタ *gram* が含まれる。コーパス中の 1356 記事における出現頻度上位 9 助詞について、事例数を数えることで参照確率を計測した結果を表 5 に示す<sup>8</sup>。以下に述べる多重ロジスティック回帰と SVR では、パラメタ *gram* の値として表 5 の値を用いる。

<sup>7</sup> 選好 2 の検証は絶対値や比の設定に影響を受ける可能性がある

<sup>8</sup> 幾つかの要因を統合して参照確率を求める場合にはコーパスの事例数が疎になるため回帰分析が必要になるが、文法機能のみから参照確率を求めるためには十分な事例数があるので、ここでは単に相対頻度を数えることで計測できる。

モデル $k$	$b_{k,1}$	$b_{k,2}$	$b_{k,3}$	定数 ( $b_{k,0}$ )
1	-0.7636	9.036	2.048	-2.825
2	-0.7067	10.47	2.270	-3.055
3	-0.7574	6.433	2.399	-2.952
4	-0.5911	9.170	2.129	-3.288
5	-0.6578	4.836	2.178	-3.043

表 6 5 回の多重ロジスティック回帰による 5 つのモデルの係数

### 多重ロジスティック回帰

多重ロジスティックモデルは、ある事象が発生する確率を  $P$  としたとき、その対数オッズ  $\log(\frac{P}{1-P})$  が説明変数の線形結合式  $\lambda$  で表せるという仮定に基づいた回帰モデルである。3 節の表 1 で示した 3 素性 ( $dist, gram, chain$ ) を説明変数とする多重ロジスティック回帰式は以下のようになる。

$$\begin{aligned} P &= (1 + \exp(-\lambda))^{-1} \\ &= (1 + \exp(-(b_0 + b_1 dist + b_2 gram + b_3 chain)))^{-1} \end{aligned}$$

ただし、全 1,073,781 事例を使って多重ロジスティック回帰をするには膨大な時間がかかるため、本稿では 12,000 事例ずつサンプリングして 5 回の多重ロジスティック回帰を行った。多重ロジスティック回帰には統計ソフトウェア R(R-Project 2004) を用いた。表 6 は、その結果得られた 5 つのモデルのパラメータである。これら 5 つのモデルによって求まる確率の平均

$$\frac{1}{5} \sum_{k=1}^5 (1 + \exp(-(b_{k,0} + b_{k,1} dist + b_{k,2} gram + b_{k,3} chain)))^{-1}$$

を参照確率とした。

### SVR

3 節の表 1 で示した 8 素性を用い、参照確率の計測を行うための SVR モデルを作成した。多重ロジスティック回帰の入力は 0 または 1 の値から成る事例集合であったが、SVR によって確率を外挿するためには入力を平滑化しておく必要がある。本稿では全事例から無作為に 60,000 事例を抽出し、 $k = 100$  の  $k$ -NN 法によって平滑化したのち、TinySVM(Kudoh 2002) を用いて 2 次多項式カーネルによる SVR を行った。

## 5.2 意味ゲームから導かれる選好 1a, 選好 1b の検証

まず、4.2 節で述べた選好 1a の検証として、コーパス中の同一発話中に出現する参照表現を 2 つずつ対にし、選好 1a が成り立っているペアの比率を計測し、95%信頼区間を求める。また、

1 発話内の 参照表現数	発話数	参照表現数	参照表現の 割合 (%)	参照表現の ペア数
0	47728	0	0.0	0
1	14960	14960	89.4	0
2	854	1708	10.2	854
3	20	60	0.4	60
計	63562	16728	100.0	914

表 7 1 発話内の参照表現数

		選好 1a 成立	選好 1a 不成立
多重 logi- stic	選好 1a 成立比率 (360 組中)	75.3% (271/360)	24.7% (89/360)
	選好 1a 成立比率 (一人称を含まない組)	76.9% (227/295)	23.1% (68/295)
	選好 1a 成立比率 (一人称を含む組)	67.7% (44/65)	32.3% (21/65)
	全 16728 事例における選好 1b 相関係数	+0.373	
S V R	選好 1a 成立比率 (360 組中)	74.4% (268/360)	25.6% (92/360)
	選好 1a 成立比率 (一人称を含まない組)	74.2% (219/295)	25.8% (76/295)
	選好 1a 成立比率 (一人称を含む組)	75.4% (49/65)	24.6% (16/65)
	全 16728 事例における選好 1b 相関係数	+0.386	

表 8 選好 1a, 選好 1b の検証

選好 1b の検証として、参照表現の効用と指示対象の参照確率の相関係数を計測し、95%信頼区間を求める。

表 7 は 1 発話内の参照表現数である。選好 1a の検証のために同一発話内の参照表現を 2 つずつ対にしたところ、914 組であった。その 914 組のうち、代名詞と非代名詞のペアは 360 組であり、それ以外は同種同士のペアであった。同種の参照表現のペアにおいては効用に差が無いので、代名詞と非代名詞の参照表現対 360 組中で選好 1a が成立している比率を計測する。また、「効用と参照確率には正の相関がある」という選好 1b の検証のため、全 16728 照応詞の効用とその指示対象の参照確率の Pearson 積率相関係数を求める。それらの結果を表 8 に示す。多重ロジスティック回帰による参照確率を用いた場合は 75.3% の事例で選好 1a が満たされており、選好 1b にあたる相関係数は +0.373 であった。SVR による参照確率を用いた場合も 74.4% の事例で選好 1a が満たされており、選好 1b にあたる相関係数は +0.386 であった。ペアのどちらかが一人称である場合に限定しても、それぞれ同程度の比率で選好 1a が成り立っていた。

また、選好 1a の成立事例の母比率を二項分布と仮定した場合の 95% 信頼区間と、選好 1b で予測する相関係数を t 分布と仮定した場合の 95% 信頼区間を表 9 に示す。これは、母集団における選好 1a 成立比率が 7 割以上である確率、および選好 1b で予測する正の相関係数が 0.343 以上である確率が 97.5% であることを表している。これにより、表 8 の 2 つの回帰手法による参照確率の双方において、意味ゲームから導かれた選好 1a, 選好 1b は有意であることが示せた。

		実測値	95%信頼区間
多重 logistic	選好 1a 成立比率	0.753	[0.705, 0.796]
	選好 1b 相関係数	0.373	[0.360, 0.386]
SVR	選好 1a 成立比率	0.744	[0.696, 0.789]
	選好 1b 相関係数	0.386	[0.373, 0.399]

表 9 選好 1a 成立比率と選好 1b 相関係数の 95%信頼区間

	選好 1b 相関係数の実測値	95%信頼区間
多重 logistic	0.357	[0.343, 0.371]
SVR	0.386	[0.372, 0.400]

表 10 1 発話内に 1 照応詞の場合 (14960 事例) に限った選好 1b 相関係数と 95%信頼区間

#### 中心化理論のルール 1 と意味ゲームから導かれた選好 1b の適用範囲の比較

中心化理論のルール 1 が意味を持つのは、1 つの発話が複数の参照表現を含む場合である<sup>9</sup>。それに対して、意味ゲームから導かれる選好 1b は指示対象の参照確率と参照表現の効用との正の相関関係を予測し、その適用範囲は 1 つの発話内に参照表現が 1 つしかない場合も含む。よって、選好 1b の方がルール 1 よりも一般的である。さらに、表 10 は、1 つの発話内に 1 参照表現の事例に限っても正の相関があることを示している。すなわち、意味ゲームの予測は中心化理論が及ばない範囲においても成立する。

#### 従来の日本語 Cf ランキングの妥当性

意味ゲームから導かれる選好 1a, 1b の予測能力の良さは、顕現性と効用との相関の高さに帰着できる。従来の Cf ランキングは文法機能のみに基づく顕現性の順序であるので、文法機能のみに基づいて相関を最大化するように設定した顕現性の順序と比較することで従来の Cf ランキングの妥当性を検討する。

各助詞の出現を説明変数とし、効用 (代名詞:2, 非代名詞:1) を目的変数とする単回帰分析を行った。相関を最大化する顕現性としてこの回帰係数を用いることができる。表 11 に示す単回帰分析の結果では、直接目的語 (ヲ格) > 間接目的語 (ニ格) という順序が観測される。しかし、Walker ら (1994) による従来の日本語 Cf ランキングは、間接目的語 > 直接目的語という順序を含む点において、誤りと考えられる。

また表 9 より、他の要因も統合した参照確率としての顕現性を用いたモデルの相関係数は、表 11 の相関係数よりも更に高い。このモデルが前提としている文法役割に対応する参照確率

<sup>9</sup> 1 つの発話が 1 つの参照表現しか含まない場合、その参照表現が Cb を指すことは Cb の定義から明らかなので、ルール 1 は何も言っていないに等しい。

助詞	は	も	が	から	を	と	に	の	で
回帰係数	5.46	5.37	5.27	5.14	5.12	5.05	5.05	5.04	4.98
切片	-3.86								

効用との相関係数:+0.248

表 11 文法機能 (助詞) のみによる顕現性

	Continue	Retain	Smooth-Shift	Rough-Shift
ゼロ代名詞	56.0% (1315/2347)	1.7% (41/2347)	38.3% (898/2347)	4.0% (93/2347)
代名詞	43.6% (102/234)	2.1% (5/234)	50.9% (119/234)	3.4% (8/234)
代名詞計	54.9% (1417/2581)	1.8% (46/2581)	39.4% (1017/2581)	3.9% (101/2581)
定名詞句	20.9% (56/268)	3.0% (8/268)	64.2% (172/268)	11.9% (32/268)
一般名詞句	20.0% (522/2611)	1.8% (48/2611)	67.4% (1761/2611)	10.7% (280/2611)
非代名詞計	20.1% (578/2879)	1.9% (56/2879)	67.1% (1933/2879)	10.8% (312/2879)
合計	36.5% (1995/5460)	1.9% (102/5460)	54.0% (2950/5460)	7.6% (413/5460)

表 12 本稿のデータ (新聞記事) による中心遷移の分布

	Continue	Retain	Smooth-Shift	Rough-Shift
ゼロ代名詞	55.9% (76/136)	2.2% (3/136)	25.0% (34/136)	16.9% (23/136)
ゼロ代名詞以外	7.8% (7/90)	43.3% (39/90)	10.0% (9/90)	38.9% (35/90)
合計	36.7% (83/226)	18.6% (42/226)	19.0% (43/226)	25.7% (58/226)

表 13 Iida のデータ (新聞記事) による中心遷移の分布

	Continue	Retain	Smooth-Shift	Rough-Shift
ゼロ代名詞	47.3% (43/91)	4.4% (4/91)	30.8% (28/91)	17.6% (16/91)

表 14 竹井のデータ (小説) による中心遷移の分布

(表 5) でも, やはり直接目的語 (ヲ格) > 間接目的語 (ニ格) という順序が観測された. この意味でも, 従来の日本語 Cf ランキングの妥当性は低いと言えるだろう

ただし, 直接目的語 > 間接目的語という順序が本研究に用いたコーパスに特有である可能性も否定できない. 他の種類のコーパスに関する調査は今後の課題である.

### 5.3 意味ゲームから導かれる選好 2 の検証

4.3 節で述べた, 意味ゲームに基づく選好 2 の検証を行う. まず, 表 12 は, コーパスから計測した多重ロジスティック回帰による参照確率を顕現性と見なした場合の中心遷移の頻度分

Transition	事例数	期待効用の平均	期待効用の分散
Continue	1995	0.874	0.361
Retain	102	0.473	0.242
Smooth-Shift	2950	0.287	0.175
Rough-Shift	413	0.109	0.0336

Transition(cont.:4, ret.:3, s.s.:2, r.s.:1) と期待効用の相関係数:+0.520

表 15 中心遷移タイプ毎の期待効用の平均と分散

$\chi^2$ 値	自由度	有意確率
1780.7	3	$< 2.2 \times 10^{-16}$

表 16 Kruskal-Wallis の検定

布を示している．比較のため，表 13 に Iida(1997) による中心遷移分布データと，表 14 に竹井ら (2000) による中心遷移分布データを示す．いずれのデータによる頻度分布も，Continue と Smooth-Shift への偏りが顕著である．しかし，頻度分布の順序が選好順序と一致するとは必ずしも言えない．何故なら，各事例において 4 種類の遷移がすべて選択可能とは限らないためである (Kibble 2001) ．

そこで，頻度分布ではなく，期待効用の値によってルール 2 の順序を検証する．すなわち，遷移の種類毎に期待効用の平均値を計測し，その順序がルール 2 の順序と合致するか否かを検証する．表 15 は，遷移の種類毎の期待効用の平均と分散を表している．期待効用の平均値は Continue>Retain>Smooth-Shift>Rough-Shift となっており，従来研究における選好順序と合致する結果となった．また，4 種類の遷移の期待効用の平均値の多重比較を行った．表 16 に，R(R-Project 2004) を用いて Kruskal-Wallis の検定を行った結果を示す．これにより，4 種類の遷移の期待効用の平均値には有意差があることを示した．表 17 に，R を用いて Holm の方法で調整した Wilcoxon の順位和検定を多重実行した結果を示す．これにより，4 種類の遷移の期待効用値の順序が有意であること，すなわちルール 2 の順序と合致するという結果が統計的に有意であることがわかる．

## 6 考察

### 6.1 参照確率としての顕現性の効果

参照確率によって顕現性を定式化することにより，2 節で述べた顕現性に関する問題点を以下のように解消した．

- A 客観的に計測可能になり，実言語データに基づく統計的検証ができるようになった．

比較するタイプ	有意確率
Continue : Retain	$5.89 \times 10^{-13}$
Continue : Smooth-Shift	$< 2.2 \times 10^{-16}$
Continue : Rough-Shift	$< 2.2 \times 10^{-16}$
Retain : Smooth-Shift	$1.64 \times 10^{-6}$
Retain : Rough-Shift	$< 2.2 \times 10^{-16}$
Smooth-Shift : Rough-Shift	$< 2.2 \times 10^{-16}$

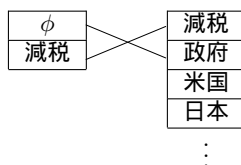
表 17 Holm の方法で調整した Wilcoxon の順位和検定の多重実行

- B 多重回帰が可能な回帰アルゴリズムを用いることにより、顕現性を決定する要因となる複数の素性 (説明変数) の統合がヒューリスティックな手法よりも容易になった。
- C 直前の発話だけでなく、先行文脈全ての実体を扱えるようになった。

## 6.2 選好 1a が成り立たない場合

本稿では、意味ゲームから導いた選好 1a, 選好 1b をコーパスを用いて検証し、統計的に有意であることを示した。選好 1a が成り立っていた事例は多重ロジスティック回帰で 75.3%, SVR で 74.4%であったが、ここでは選好 1a が成り立たない事例について考察する。多重ロジスティック回帰で用いた 3 素性から成るモデルと、SVR で用いた 8 素性から成るモデルには、選択制限や常識的知識などの意味的な素性は含まれていない。選好 1a が成り立っていなかった事例は、多重ロジスティック回帰で 24.7%, SVR で 25.6%であったが、これらの事例はモデルに含まれていない要素の作用によって選好が覆されていると考えられる。つまり、選択制限などの素性をモデルに取り込めば、より強い選好となることが予想される。以下に、選好 1a が覆されていた具体例を示す。

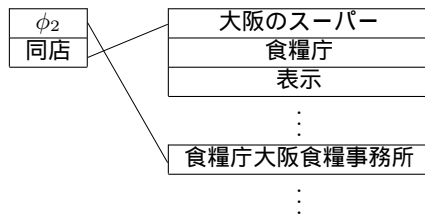
政府は二日、... 政策の骨格を固めた。柱の一つでもあり、米国が強く求めている減税は、来年度以降も今年度に近い規模の所得・住民税減税を恒久的に実施する意向を ( $\phi$  ガ) 表明した。



上記事例では、意味的な制約が含まれていない本稿のモデルによれば Pareto 最適解は  $\{\phi \leftarrow \text{減税}, \text{減税} \leftarrow \text{政府}\}$  であるが、実際の正解は意味的な制約によって覆されている。

同事務所に十日、「国産米があったのでレジに持っていきと一万八千円と言われびっくりした。(  $\phi_1$  ) 売り場に表示もなかった」という主婦からの通報があり、十一日に同店を ( $\phi_2$  ガ) 調査。





上記事例では、意味的な制約が含まれていない本稿のモデルによれば Pareto 最適解は  $\{\phi_2 \leftarrow \text{大阪のスーパー}, \text{同店} \leftarrow \text{食糧庁}\}$  であるが、実際の正解は文脈的知識や常識などの意味的な制約によって覆されている。

以上のように、選好 1a が満たされていない事例においては、本稿で用いたモデルに含まれていない意味的・言語外的な制約が働いて選好が覆されていた。よって、語の意味的な近さや選択制限などの意味的な制約をモデルに追加することにより、選好 1a, 選好 1b はより強い選好になるであろうと予想される。

また、本稿では代名詞・非代名詞という区別に基づいて参照表現の効用を決定したが、顕現性の定義と同様に、参照表現の効用も客観的に計測可能な尺度として定義することが望ましい。これについては今後の課題とする。

### 6.3 選好 2 の検証の厳密性

本稿では選好 2 の検証として、実際のコーパスにおける 4 種類の遷移の期待効用の平均が中心化理論における選好順序と合致していることを示し、分散検定によってこの結果が統計的に有意であることを示した。本稿で示した全体の傾向における順序は、ルール 2 を意味ゲームに帰着できることを示唆している。ただし、この検証によって示したのは全体の傾向における順序であり、各事例における解候補間の順序ではないため、より厳密には各事例における解候補間で期待効用の順序が付けられることを示す必要がある。

## 7 おわりに

中心化理論は広く論じられている照応の理論であるが、照応現象の基本原理の解明には至っておらず、また理論において重要な役割を担う顕現性の意味が不明確だという問題点がある。本稿では、照応現象の背後にある基本原理はゲーム理論によって捉えることができるという観点に立ち、顕現性を参照確率として定式化することにより、意味ゲームから中心化理論のルール 1,2 に相当する選好を導出できることを説明し、この選好の妥当性を以下のように実言語データに基づいて検証した。

まず、意味ゲームから導出したルール 1,2 に対応する選好をコーパスを用いて統計的に検証した。ルール 1 に対応する選好 1b については、指示対象の参照確率と参照表現の効用との正の

相関を観測した。これに関連して、従来提案されてきた日本語の Cf ランキングは参照確率の順序と異なり、誤りであることがわかった。さらに、ルール 1 は 1 つの発話に複数の参照表現を含む場合にのみ意味を持つが、意味ゲームから導いた選好 1b が予測する相関関係は 1 発話に 1 参照表現しか含まない場合においても成り立つことを示した。つまり、意味ゲームはルール 1 よりも強い予測を導く。ルール 2 に対応する選好 2 の検証としては、期待効用の順序がルール 2 の順序と合致することを観測した。

以上より、意味ゲームは基本原理の明確さおよび予測能力の強さゆえに中心化学理論よりも優れた作業仮説である。ゆえに、中心化学理論のような領域に依存した理論は、照応現象に関しては不要と考えられる。

謝辞 有意義なコメントを頂いた査読者の皆様、本研究を進めるにあたってコメントや励ましのお言葉を頂いた旧サイバーアシスト研究センターの皆様に感謝致します。また GDA コーパス作成に携わった方々に深謝致します。

## 参考文献

- Hasida, K. (1996). "Issues in Communication Game." In *Proceedings of COLING'96*, pp. 531–536.
- Hasida, K. (1998). "大域文書修飾 Global Document Annotation (GDA)." <http://i-content.org/gda/>.
- Hasida, K., Nagao, K., and Miyata, T. (1995). "A Game-Theoretic Account of Collaboration in Communication." In *Proceedings of the First International Conference on Multi-Agent Systems*.
- Iida, M. (1997). "Discourse Coherence and Shifting Centers in Japanese Texts." In Walker, M., Joshi, A., and Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 161–180. Oxford University Press.
- Kameyama, M. (1998). "Intrasentential Centering: A Case Study." In Walker, M., Joshi, A., and Prince, E. (Eds.), *Centering Theory in Discourse*, pp. 89–112. Oxford University Press.
- Kibble, R. (2001). "A Reformulation of Rule 2 of Centering Theory." *Computational Linguistics*, **27** (4).
- Kibble, R. (2003). "Towards the Elimination of Centering Theory." In Kruijff-Korbayova, I. and Kosny, C. (Eds.), *Proceedings of the 7th Workshop on the Semantics and Pragmatics of Dialogue*, pp. 51–58.
- Kudoh, T. (2002). "TinySVM: Support Vector Machines." <http://chasen.org/taku/software/TinySVM/>.

- Nariyama, S. (2001). "Multiple argument ellipses resolution in Japanese." *In Proceedings of Machine Translation Summit VIII*, 241–245. Spain.
- R-Project (2004). "The R Project for Statistical Computing." <http://www.r-project.org/>.
- Reed, C. (2002). "Saliency and the Attentional State in Natural Language Generation." pp. 440–444 Lyon, France.
- Strube, M. and Hahn, U. (1999). "Functional Centering: Grounding Referential Coherence in Information Structure." *Computational Linguistics*, **25(3)**, 309–344.
- van Rooy, R. (2003). "Relevance and Bidirectional OT." In Blutner, R. and Zeevat, H. (Eds.), *Pragmatics in Optimality Theory*, pp. 173–210. Palgrave Macmillan.
- Walker, M., Iida, M., and Cotes, S. (1994). "Japanese Discourse and the Process of Centering." *Computational Linguistics*, **20** (2).
- 竹井光子, 高田美佳, 相沢輝昭 (2000). "日本語ゼロ代名詞補完のためのグローバルトピックの役割." 情報処理学会研究報告, **135** (10), 71–78.

#### 略歴

- 白松 俊 (非会員): 2000年東京理科大学工学部情報科学科卒業。2003年同大学院修士課程修了。同年, 科学技術振興機構 CREST 研究補助員。GDA コーパスを用いた照応研究に従事。2005年, 京都大学大学院博士後期課程入学。
- 宮田 高志 (正会員): 1991年東京大学理学部情報科学科卒業。1996年同大学院博士課程修了。理学博士。同年, 奈良先端科学技術大学院大学情報科学研究科助手, 2001年より科学技術振興機構 CREST 研究員, 情報検索・構文解析の研究に従事。
- 奥乃 博 (非会員): 1972年東京大学教養学部基礎学科卒業。博士(工学)。NTT, JST, 東京理科大学を経て, 現在, 京都大学情報学研究科教授。音環境理解・ロボット聴覚研究に従事。1990年度人工知能学会論文賞, 平成14年度船井情報科学振興賞等受賞。IPSJ, JSAI, JSSST, RSJ, ACM, IEEE 等会員。
- 橋田 浩一 (正会員): 1981年東京大学理学部情報科学科卒業。1986年同大学院博士課程修了。理学博士。同年, 電子技術総合研究所入所, 現在, 産業技術総合研究所情報技術研究部門副研究部門長。知的コンテンツの研究開発に従事。