

特集 「類似性に基づく推論」

言語と類似性

Language and Similarity

内海 彰
Akira Utsumi

電気通信大学電気通信学部システム工学科
Department of Systems Engineering, The University of Electro-Communications.
utsumi@se.uec.ac.jp, <http://www.utm.se.uec.ac.jp/~utsumi/>

Keywords: language, similarity, linguistics, natural language processing, word's meaning, metaphor.

1. はじめに

似ているかどうか、またはどのくらい似ているかを判断する能力は、人間の思考・推論・問題解決などの認知活動において最も基礎となるものである。例えば、新奇なものごとに出会ったとき、記憶の中にある似ているものごとを探索し、それに関する経験をもとに新奇なものごとの振舞いを予測したり、それに対する反応や行動を決定したりする。つまり、人間の認知機構には類似性判断が不可欠である [Goldstone 99, 鈴木 96]。

よって人間の知を最も特徴付ける言語に関しても、類似性はその処理に重要な役割を果たすことは明らかであろう。例えば、子供がことばの意味（概念）を獲得する際に概念どうしの類似性判断が影響する [Gentner 97]。また大人でも、例えば第二言語（母国語以外の言語）習得の際に、言語表現どうしの類似性判断に基づいて学習していない表現を理解したりする。さらに、言語表現の類似性を利用することによって、比喩などの創造的な表現の生成や、駄洒落などのことば遊びが可能になる。

言語の類似性が人間の言語処理に影響を及ぼすことは、プライミング効果という形でさまざまな心理実験で示されている [Balota 94, 齊藤 97]。プライミング効果とは、一般に先行刺激が後続刺激の処理に何らかの影響を与えることをいう。例えば、単語 **dog** に続いて呈示される単語 **cat** に対する反応（例えば単語・非単語の判断）は、単語 **pen** に続いて呈示される単語 **cat** に対する反応よりも促進される（反応時間が短い）。つまり単語間の意味的な類似性（**dog** と **cat** は意味的に類似しているが、**pen** と **cat** は類似していない）が単語認知に影響を及ぼすのである。

一方、計算機による自然言語処理においても、近年、言語の類似性を用いた手法が積極的に利用されている。実例に基づく機械翻訳 [佐藤 97] の成功は、その最たる例である。実例に基づく機械翻訳では、大量の翻訳例の中から翻訳したい表現と類似した翻訳例を見つけて、それを参考に翻訳を行う。このような実例に基づく手法は、多義性解消などの機械翻訳以外の処理でも用いられている。

以上のように、言語の類似性は人間・計算機いずれの言語処理にとっても大変重要な概念である。そこで本稿では、言語学と自然言語処理の両面から、言語の類似性研究を概観する。まず2章では、言語の類似性にはどのようなものが考えられるのか、それはどのような言語現象や学問分野に関係するのかを概観する。3章では、その中の一つであり、言語研究で最も根本的な話題である単語の意味の類似性について、比喩との関係にも触れつつ論じる。そして4章では、工学的な観点から、単語の意味の類似度計算の手法とそれを用いた自然言語処理について紹介する。

2. 言語の類似性とは

2.1 記号の類似

言語表現を（意味を取捨した）単なる文字列・記号列として見れば、それらの類似性は文字列間の距離として定式化できる。文字列間の距離は、例えば一方の文字列を他方の文字列に変形するために必要な基本操作数（文字の削除・置換・追加など）として定義できる [黒橋 98]。この定義によれば、'surgery' の **g** を **v** に置換して **r** を削除すると 'survey' になることから、'surgery' と 'survey' の距離は 2 ということになる。

このような記号列としての類似性は、情報検索などで文字列の近似照合を行う場合に用いられる。またことばの記号列としての同一性は、回文、アナグラム、ダブルレットなどのことば遊びの基盤になっている [安井 92]。例えば、以下に示すパナマ運河開通を報じた新聞の見出しは有名な回文である。"A man, a plan, a canal — Panama."

2.2 音の類似

前節の「文字・記号」を「音（音韻）」にすると、言語表現間の音に関する類似性を考えることができる。例えば、同音異義語（homonym）である 'pail'（手おけ）と 'pale'（青白い）は音韻列という点では同一であるし、'pale' と 'tale' は最初の子音以外は同じ音韻列なので、かなり類似していることになる。

このような音韻列としての類似性は、欧米語の詩（やその他の言語表現）の技巧である頭韻（**alliteration**）や脚韻（**rhyme**）として利用されている [赤羽 98]。例えば、米大統領アイゼンハワーが選挙で用いたスローガン “I like Ike.” (Ike はアイゼンハワーの呼称) には、頭韻と脚韻が巧妙かつ効果的に用いられている。また、地口（**pun**）（「着た切り雀（舌切り雀）」のように、成句などに音の類似した別の語を当てて別の意味を含ませる語法）や駄洒落（「トイレに行っといれ」のように類似した音韻列を近い位置に併存させる語法）も音の類似を用いたことばの使用（ことば遊び）であり [滝澤 96, 安井 92]、音韻列の類似度を計算することによって地口や洒落の検出や生成を行う計算機システム [ピンステッド 98, 滝澤 96] も開発されている。

2.3 構造の類似

文や句どうしの類似関係として、統語構造の類似を考察することができる。例えば「黒い髪の女」と「青い目の子供」という二つの表現は、「形容詞＋名詞＋の＋名詞」という品詞列として見ると同じであり、さらにそれらの係り受け構造（形容詞が直後の名詞に係り、「名詞＋の」が最後の名詞に係る）が同じである。このような構造的な類似性・同一性を利用すると、例えば長い日本語文における並列構造の自動推定などが可能になる [黒橋 92]。

2.4 単語の意味の類似

言語の類似性として最も重要なのが意味の類似性である。その中でも、意味を担う基本単位である単語どうしの意味の類似は言語研究にとって重要である。例えば、類義語（**synonym**）である「夜」と「晩」は、どちらも一日のうち太陽の出ていない暗い時間帯を指す点で非常に類似しているとともに、指す時間帯の範囲において差異がある。「犬」と「猫」は類義語ではないが、同じカテゴリ（例えば「獣」）のメンバという点で意味的に類似している。この点では、「男」と「女」のような反意語（**antonym**）も、一見すると意味的に非常に離れているように感じるが、意味的に類似していると考えられる。さらに、「親」と「父」のような上位下位関係も意味の類似とみなすことができる（このような関係を包摂性（**hyponymy**） [池上 75] という）。

なお、ここでいうカテゴリは、このような辞書的（分類学的）カテゴリに限らず、アドホックカテゴリ [Barsalou 83] のように動的に生成されるカテゴリも含んでいる。例えば「鰻」は魚という点では「林檎」よりも「鮫」と類似しているが、食物という視点からは「林檎」のほうが「鮫」よりも類似している。このように単語の意味の類似性を広く解釈していくと、隠喩や直喩などの比喩表現に突き当たることになる [Glucksberg 90, 内海 01]（単語の意味の類似については、3章で詳しく述べる）。

2.5 文の意味の類似

単語より大きい単位である文（もしくは句や節）の間でも意味の類似性を考えることができる。例えば、「彼の鼻は大きい」と「彼は大きい鼻をしている」は、構造は異なるがそれらの文が表す意味（文意）は同じである。このような文意の類似も、言語処理において重要である。言い換え（**paraphrasing**）はある表現を文意が類似した別の表現に変換する処理であるし、言い換えを異なる言語間で行えば翻訳になる。

さらに、文意の類似では捉えられない文間の意味的な類似として、発話の意味（話し手の意味）の類似が考えられる。発話の意味とは話し手が意味（意図）することであり、話し手が使用する言語表現そのものの意味である文意とは異なる [Blakemore 92]。例えば、「ちょっと寒いよ」という文と「窓を閉めてほしいな」という文は、文意の点ではほとんど類似していないが、両文とも窓を閉めてもらうという意図で発せられている点で、発話の意味としては類似している。

関連性理論 [Sperber 95] では、このような発話の意味の類似性を共有される含み（**implication**）の多さとして捉える。つまり二つの文から得られる含みの重なりが大きいほどそれらは類似していると考えられる。この考え方は、ことばのゆるい使用（**loose talk**）やアイロニーなどの多くの言語現象の説明に適用されている [Sperber 95]。

2.6 テキストの類似

複数の文からなる文章やテキストの類似性を考えることも可能である。一つには、文章内容の点からのテキストの類似が考えられる。テキスト内容の類似は、最も単純には、文章中の単語（キーワード）の出現頻度分布の類似として捉えられるし、句・文のパターンなどその他の情報を用いて類似度を定義することもできる。テキストの自動分類やクラスタリングは、このようにして得られるテキスト間の類似度を用いて行われる [佐藤 98]。また、テキストの要約はそのテキストと内容が類似した短いテキストに変換する処理と考えられる。

一方、文章の書き手やジャンル（新聞記事、論文、小説など）の点からテキストの類似を考えることもできる。例えば真贋判定のための計量的な文体研究では、テキストの計量可能な諸要因（e.g., 文の長さ、品詞の使用率、比喩の使用度、句読点の使用度）によって定義される類似度が高い複数のテキストは同じ書き手のものと判断するのである [村上 94]。

2.7 言語システムの類似

よりマクロな視点として、日本語、英語などの個別言語間での類似性を考えることもできる。言語類型論は、このような言語間の類似性を手掛かりとして、言語の普遍的特性を探求する学問分野である [Comrie 89]。例えば基本語順を考えると、英語、仏語、独語などは SVO

である点で似ている(同じである)のに対して、SOVの語順をもつ日本語などとは似ていない(異なる)ということになる。しかし「主語が目的語の前にくる」という点では、これらの語順をもつ言語やさらにVSOの語順をもつアメリカ・インディアン諸語はすべて類似していることにもなる。

3. 単語の意味の類似性

3.1 単語の意味の類似性の分類

2.4節でも述べたように、単語の意味の類似といても、さまざまな観点からの類似を考えることができる。ここでは、野内[野内 00]の類似性の分類を援用・修正して、単語の意味の類似性を以下のように分類する。

- (1) 直接的類似性：辞書的カテゴリの類似
 - 例：類義語、反意語、上位語・下位語(包括性)
- (2) 間接的類似性：辞書的カテゴリ以外での類似
 - a 特徴・イメージ的類似性
 - 例：「女性」と「バラ」(美しさ)
 - b 構造的類似性
 - 例：「愛」と「旅」[Lakoff 80]
 - c 機能的類似性
 - 例：「鰻」と「林檎」(食べられるもの)

直接的類似性はソーラス(類義関係を中心に語を分類し、それらの上位下位関係を木構造で表した辞書)に基づく類似性のことである。例えば、2.4節であげた類義語「夜」と「晩」は、日本語の代表的なソーラスである分類語彙表[国研 64]では同じ項目(1.1635)に分類されている。一方、間接的類似性はソーラスや辞書には記述されていないその語(の表す概念)がもつと一般的に思われる性質、属性、構造、機能などの点での類似性である。ただし、直接的・間接的という分類は便宜的なものであり、2.4節で述べたように、単語が属するカテゴリの類似として統一的に扱うことができる。例えば、「女性」と「バラ」はどちらも「美しいもの」というアドホックカテゴリに含まれる点で類似している。

さらに単語の意味を広く考え、ある単語の意味をその語から連想される語によって規定する考え方もある[池上 75]。しかし実際には、もとの単語(刺激語)と連想語の関係のほとんどは、上記の意味的類似性で説明可能である。Miller [Miller 91]によると、連想実験における連想語の種類は、刺激語と連想語の意味的關係から以下の四つのカテゴリに分類することができる*1(下記の例はすべて刺激語‘chair’の連想語である)。

- (1) 上位、等位、下位語(例：‘table’, ‘furniture’)
- (2) 属性を表す語(例：‘comfortable’, ‘wooden’)

(3) 部分・全体の関係を表す語(例：‘seat’, ‘cushion’)

(4) 機能に関する語(例：‘sit’, ‘rest’)

これらのうち、(3)以外は上述した意味的類似性による連想である((1)は直接的な類似性、(2)は特徴・イメージ的類似性、(4)は機能的類似性)。なお(3)は近接性(proximity)(または隣接性(contiguity))と呼ばれる関係の一種であり、一般的に類似性と区別される。近接性とは二つの語概念が何らかの点で近いということであり、空間・時間的な近さから因果関係的な近さまで含んでいる。

以上の議論では、単語の意味はそれが指示する概念・カテゴリによって規定されるとみなしている。これは言語能力は認知能力と不可分な関係にあるとする認知言語学[Ungerer 96]の考え方と同じである。一方では、現実とは切り離れた言語記号内の差異関係から意味が生じるとするソシユール言語学[丸山 81]や、言語能力は一般的な認知能力から独立していると考えられる生成文法理論[福井 01]のような考え方も存在する。前者の立場であれば、言語(単語の意味)の類似性(判断)を考えるには、言語を含む認知過程一般の中での類似性について議論しなければいけないことになる。後者の考えに従うならば、言語の類似性を考えることは言語に特殊な知識の類似性やその利用法を考えることになる。ここではどちらの考えがより適切かは議論しないが、本稿では(個人的にも)前者の立場を取っているため、単語の意味やその類似性を論じるには、人間の認知過程一般の中での類似性(判断)[楠見 02]を考えていく必要があると考えている*2。

3.2 単語の類似性と修辞・比喩

語の意味的類似性は、以下の対応関係が示すように修辞表現(レトリック)と密接な関係がある。

- 直接的類似性—提喩(synecdoche)
- 間接的類似性—隠喩(metaphor)、直喩(simile)
 - 特徴・イメージ的類似性—属性比喩
 - 構造的類似性—構造比喩
- 近接性—換喩(metonymy)

提喩は「花見をする」のように「花」(上位概念、類)で「桜」(下位概念、種)を表したり、逆に「ご飯」(下位概念)で「食事」(上位概念)を表すといった上位下位(類と種)の關係に基づく修辞表現である。なお、部分で全体を表す(例えば「青い日」で「外人」を表す)ように、近接性を利用した修辞は換喩と呼ばれる。一方、隠喩や直喩は間接的類似性に関連する修辞である。

- (1) 僕の恋人は赤いバラ(のよう)だ。

*1 出現頻度は低いが、そのほかの關係として、「たばこ」と「煙」のような部分・全体以外の近接性[阿部 94]や「白」と「城」のような音(形式)の類似性[池上 78]なども考えられる。

*2 なお、単語の意味のすべてが概念によって説明できるとも考えていない。例えば音や形式の類似が意味をつくり出すことも(特に詩などでは)可能である[赤羽 98]。

(2) 愛は旅だ。

文(1)は「恋人」を「バラ」で例える隠喩(「~のよう」などの指標を明示した場合は直喩)であり、特徴・イメージの類似性が主に関係している。(2)は「愛」を構造的に類似している「旅」でたとえる隠喩である。

以上のように、語の意味や類似性を議論していくうえで、修辞表現は不可欠な存在である。よって、隠喩(の意味)とは何であり、それはどのような過程を通じて理解されるのかを解明することは、語の意味やその類似性判断に関する研究でもある[赤羽 98]。

アリストテレス以来の古典修辞学では、隠喩は現実世界に関する既存の類似性の効果的な装飾表現と考える[Lausberg 63]。すなわち、(1)の例で言えば、あらかじめ「恋人」と「バラ」の類似性が認識されていて、隠喩(1)は「僕の恋人は美しい」ことを効果的に述べていることになる。この考え方をさらに推し進めた概念メタファ論[Lakoff 80, Lakoff 89]では、このような我々の心に存在する概念間の類似性(写像)そのものが隠喩(概念メタファ)であると捉える。例えば(2)のような概念メタファが心の中に存在し、これに基づいて“*Our relationship has hit a dead-end street*”などの隠喩表現が生成される。

確かに多くの慣習的な隠喩はこのような説明で問題ないが、隠喩は必ずしも既存の類似性に基づいた表現というわけではない。現代の隠喩論の多くは、隠喩そのものが新たな類似性を創造すると主張する[Black 62, Indurkha 91, Tourangeau 82, 内海 01]。例えば、隠喩文(1)では、もともと似ていない「恋人」と「バラ」が結び付けられることによって、美しさ、冷淡さ、無邪気さ、秘密などに関する類似性が動的に生成され、それが認知され、詩的效果が生じるのである。言い換えれば、語の意味は弾性[佐藤 96]をもっていて、語の新たな使用によってその意味が拡張されるということである。よって、語の意味や類似性を考えるには静的な類似性判断だけではなく、語の意味やその類似性の動的な側面にも目を向ける必要がある。

4. 単語の類似性の工学的処理

4.1 単語の類似度の計算

単語間の意味的類似度の計算手法は、シソーラスを用いる方法とベクトル空間モデルによる方法に大別できる。

§1 シソーラスを用いる方法

この手法は3.1節で述べた直接的類似性を計算することに相当し、二つの単語間の類似度はシソーラス上での近さとして定義される。具体的には、二つの語に共通する上位概念(ノード)の深さ・位置[黒橋 92, 長尾 96, Sumita 92]や二つの語の間のパスの長さ[Fujii 98, Uramoto 94]によって類似度が計算される。

これらの計算法では分類語彙表[国研 64]などの分類

シソーラス(木構造の葉の部分だけに語が存在する)が用いられるが、EDR 概念辞書や WordNet [Fellbaum 98] などの上位下位シソーラスでは、木構造の内部にも語が対応するため、この計算法による類似度が適切な尺度にならないことがある。このような場合には、共通上位概念の深さとともに語自身の深さを考慮する方法[長尾 96]や、共通上位概念が有する下位概念の総数を用いる方法[大井 97]などが提案されている。

§2 ベクトル空間モデルによる方法

この手法は、単語の意味(語義)を多次元ベクトルとして表現し、それらの類似度を計算する方法である。この手法はもともと情報検索におけるクエリ(検索質問)と文書の関連度を計算するために開発されたもので、ベクトルの各次元は索引語に対応し、その値は例えば索引語の出現頻度が割り当てられる[Salton 83]。したがって、ベクトル空間モデルを用いて単語の類似性を計算するには、「単語ベクトル(の表現)をどのように決定するか」という問題と「ベクトル間の類似度の計算をどのようにするか」という問題を考えなければならない。

前者の問題(単語ベクトルの表現方法)に関しては、以下の手法が提案されている。

- コーパス中の語の共起情報を用いる方法：共起関係にある語をベクトルの各次元に対応させる。用いられる共起関係には以下のようなものがある。
 - ある範囲内(固定長単語列内、同一文内など)における共起[Niwa 94, Schütze 92, Schütze 98]
 - 文法的な関係に基づく共起：主語や目的語と述語の共起[Fujii 98, Hindle 90]、複合名詞における名詞の共起[稲子 00]
- ベクトルの各次元の値には、共起頻度[稲子 00]、tf-idf 値[Fujii 98]、相互情報量[Hindle 90, Niwa 94]などが用いられる。
- 辞書の語義文を用いる方法[笠原 97, 小嶋 97, 渡部 01]：語義文に含まれる語をベクトルの各次元に対応させる。しかしこれだけだと data sparseness (ベクトルのほとんどの次元が0になる)という問題があるので、再帰的参照や逆引き参照[笠原 97]、活性伝搬[小嶋 97]、次元の縮小[小嶋 97, 渡部 01]などの対処法が提案されている。

一方、ベクトル間の類似度計算に関しては、最もよく使われるのが両ベクトルの成す角の余弦(cosine)を類似度とする方法である。その他にはユークリッド距離[小嶋 97]や Kullback-Leibler 情報量(ダイバージェンス)[稲子 00]なども類似度計算に用いられる。

なお、ベクトル空間モデルに基づいて計算される類似度の意味合いは、(3.1節の直接的類似性に対応する)シソーラスによる類似度に比べて、それほど明確ではない。これに関しては、単語ベクトルの表現方法による類似度の違いを比較した研究[Niwa 94]やベクトル空間モデルでの類似度が心理学的な意味での類似性(例えばブ

ライミング効果の度合い)に対応することを実験的に示す研究 [Burgess 97] などが行われている。

また、類似性判断は文脈や観点などの違いにより変化する [Goldstone 99]。これを考慮した、文脈や観点に依存した類似度の計算方法の研究も行われている。小嶋ら [小嶋 97] は、文脈を表す語集合がベクトル空間上で均質に配置されるようにベクトルの各次元を拡大・縮小する手法により、文脈に依存した類似度計算を実現している。また笠原ら [笠原 97] は、観点を表す単一語のベクトルを特徴付ける次元について、すべての語ベクトルのその次元における値を大きくすることによって観点に応じた類似度の計算を行っている。

4.2 類似性を利用した自然言語処理

類似性を利用する言語処理としてまず考えられるのが、蓄積されている事例に基づいて新たな対象(言語表現)を処理するという事例に基づく言語処理である。類似性は現在の対象を処理するうえでどの事例を参考にするかを決定する際の尺度として用いられる。

事例に基づく言語処理の最初の研究は、長尾 [Nagao 84] によるアナロジーによる機械翻訳手法の提案である。この論文での提案を実際にプログラムとして計算機上で動作させ、事例に基づく機械翻訳を成功させたのが佐藤の一連の研究 [佐藤 91a, 佐藤 91b, 佐藤 97] である。この機械翻訳システム (MBT1 [佐藤 91a, 佐藤 97]) では、まず与えられた翻訳対象の可能な翻訳候補をすべて生成する。例えば "He plays cards" を翻訳する場合、翻訳候補として「彼はトランプをする」、「彼はトランプをひく」、「彼ははがきをする」、「彼ははがきをひく」などが生成される。次にそれぞれの候補に対して最も類似した翻訳例を探す。このときに類似度の計算が必要であり、MBT1 では翻訳対象と同じ動詞を含む事例に対して、シソーラスから計算した単語(名詞)どうしの類似度をもとに計算している。そして翻訳候補の中で事例との類似度が最も大きい翻訳候補を翻訳結果として出力する。この研究を契機にして、事例に基づく機械翻訳の研究は多く行われている [Sumita 92, Sumita 00]。

機械翻訳以外の事例に基づく言語処理としては、多義性解消への適用 [Fujii 98, Niwa 94, Schütze 98, Uramoto 94] があげられる。事例に基づく多義性解消では、多義語を含む処理対象(入力)と事例の類似度を計算し、最も類似度の大きい事例に付与されている語義をその多義語の語義とする。空間ベクトルモデルによる方法では、ベクトル表現された多義語を含む事例をクラスタリングして語義ベクトルをつくり、処理対象の多義語のベクトルとの類似度を最大にする語義ベクトルを選ぶ方法 [Schütze 98] や、個々の事例ベクトルと処理対象の単語ベクトルとの類似度を計算する方法 [Niwa 94] などがある。一方シソーラスを用いる手法では、動詞の格要素(名詞)の類似度を用いて計算した事例と入力の類似

度を最大にする動詞の語義を選択する [Fujii 98, Uramoto 94]。

事例に基づく言語処理以外にも、類似性を用いた研究は多く行われている。例えば長文における並列構造の推定 [黒橋 92] では、並列している二つの部分を同定するために二つの文節間の類似度を定義して、類似度が最大となる並列構造を取り出す。この類似度計算において、語そのものの一致や品詞の一致のほかにシソーラスに基づく語の類似度が用いられている。また情報検索においては、検索質問に含まれる語と類似した語を付け加える検索質問拡張などに単語の類似度が援用されている [大井 97, 徳永 99]。

5. おわりに

本稿では、類似性という観点から言語や言語処理を概観した。3.1節でも述べたように、単語(言語)の意味の類似性は、概念・カテゴリ化、類推などの類似性に関わる人間の認知過程と切り離して考えることはできない。その意味でも、本稿と合わせて本特集の楠見論文 [楠見 02] も参照されたい。類推に関して日本語で読める書籍としては [鈴木 96] がお勧めであるし、概念研究については本学会誌の解説記事 [河原 01] も参考になる。

事例に基づく言語処理については、やはり [佐藤 97] を読むことをお勧めする。また類似性に基づく言語処理は、現在の主流である統計的言語処理とも深く関係している。統計的言語処理については文献 [Manning 99] などが参考になる。なお本稿ではふれることができなかった比喩の計算機処理に関しては、文献 [Martin 96, 諏訪 93, 内海 00, 内海 01] を参照されたい。

◇ 参考文献 ◇

- [阿部 94] 阿部純一, 桃内佳雄, 金子康朗, 李光五: 人間の言語情報処理, サイエンス社 (1994)
- [赤羽 98] 赤羽研三: 言葉と意味を考える [I] 隠喩とイメージ, 夏目書房 (1998)
- [Balota 94] Balota, D.: Visual word recognition, in Gernsbacher, M. ed., *Handbook of Psycholinguistics*, pp. 303-358, Academic Press (1994)
- [Barsalou 83] Barsalou, L.: Ad hoc categories, *Memory & Cognition*, Vol. 11, No. 3, pp. 211-227 (1983)
- [ピンステッド 98] キム ピンステッド, 滝澤 修: 日本語駄洒落なぞなぞ生成システム "BOKE", 人工知能学会誌, Vol. 13, No. 6, pp. 920-927 (1998)
- [Black 62] Black, M.: Metaphor, in Black, M. ed., *Models and Metaphors*, pp. 25-47, Cornell University Press (1962)
- [Blakemore 92] Blakemore, D.: *Understanding Utterances*, Oxford: Basil Blackwell (1992); 武内, 山崎 (訳): ひとは発話をどう理解するか—関連性理論入門—, ひつじ書房 (1994)
- [Burgess 97] Burgess, C. and Lund, K.: Modelling parsing constraints with high-dimensional context space, *Language and Cognitive Process*, Vol. 12, pp. 177-210 (1997)
- [Comrie 89] Comrie, B.: *Language Universals and Linguistic Typology*, Chicago: University of Chicago Press (1989); 松本, 山本 (訳): 言語普遍性と言語類型論, ひつじ書房 (1992)
- [Fellbaum 98] Fellbaum, C., ed.: *WordNet: An Electronic Lexical Database*, MIT Press (1998)

- [Fujii 98] Fujii, A., Inui, K., Tokunaga, T. and Tanaka, H.: Selective sampling for example-based word sense disambiguation, *Computational Linguistics*, Vol. 24, No. 4, pp. 573-597 (1998)
- [福井 01] 福井直樹: 自然科学としての言語学—生成文法とは何か, 大修館書店 (2001)
- [Gentner 97] Gentner, D. and Medina, J.: Comparison and the development of cognition and language, *Cognitive Studies*, Vol. 4, No. 1, pp. 112-149 (1997)
- [Glucksberg 90] Glucksberg, S. and Keysar, B.: Understanding metaphorical comparisons: Beyond similarity, *Psychological Review*, Vol. 97, pp. 3-18 (1990)
- [Goldstone 99] Goldstone, R.: Similarity, in Wilson, R. and Keil, F., eds., *The MIT Encyclopedia of the Cognitive Sciences*, pp. 763-765, MIT Press (1999)
- [Hindle 90] Hindle, D.: Noun classification from predicate-argument structures, in *Proc. of 28th Annual Meeting of the ACL*, pp. 268-275 (1990)
- [池上 75] 池上嘉彦: 意味論, 大修館書店 (1975)
- [池上 78] 池上嘉彦: 意味の世界, NHK ブックス (1978)
- [稲子 00] 稲子希望, 笠原 要, 松澤和光: 複合語内単語共起による名詞の類似性判別, 情報処理学会論文誌, Vol. 41, No. 8, pp. 2291-2298 (2000)
- [Indurkha 91] Indurkha, B.: *Metaphor and Cognition: An Interactionist Approach*, Kluwer Academic Publishers (1991)
- [笠原 97] 笠原 要, 松澤和光, 石川 勉: 国語辞書を利用した日常語の類似性判別, 情報処理学会論文誌, Vol. 38, No. 7, pp. 1272-1283 (1997)
- [河原 01] 河原哲雄: 概念の構造と処理, 人工知能学会誌, Vol. 16, No. 3, pp. 435-440 (2001)
- [小嶋 97] 小嶋秀樹, 伊藤 昭: 文脈依存的に単語間の意味距離を計算する一手法, 情報処理学会論文誌, Vol. 38, No. 3, pp. 482-489 (1997)
- [国研 64] 国立国語研究所: 分類語彙表, 秀英出版 (1964)
- [黒橋 92] 黒橋禎夫, 長尾 真: 長い日本語文における並列構造の推定, 情報処理学会論文誌, Vol. 33, No. 8, pp. 1022-1031 (1992)
- [黒橋 98] 黒橋禎夫: テキスト処理, 言語の科学 9, 言語情報処理, pp. 1-50, 岩波書店 (1998)
- [楠見 02] 楠見 孝: 類似性と近接性—人間の認知の特徴について—, 人工知能学会誌, Vol. 17, No. 1, pp. 2-7 (2002)
- [Lakoff 80] Lakoff, G. and Johnson, M.: *Metaphors We Live By*, The University of Chicago Press (1980); 渡辺, 楠瀬, 下谷 (訳): レトリックと人生, 大修館書店 (1986)
- [Lakoff 89] Lakoff, G. and Turner, M.: *More than Cool Reason: A Field Guide to Poetic Metaphor*, The University of Chicago Press (1989); 大堀俊夫 (訳): 詩と認知, 紀伊国屋書店 (1994)
- [Lausberg 63] Lausberg, H.: *Elemente der literarischen Rhetorik*, Max Hueber Verlag (1963); 萬澤正美 (訳): 文学修辞学, 東京都立大学出版会 (2001)
- [Manning 99] Manning, C. and Schütze, H.: *Foundations of Statistical Language Processing*, MIT Press (1999)
- [Martin 96] Martin, J.: Computational approaches to figurative language, *Metaphor and Symbolic Activity*, Vol. 11, No. 1, pp. 85-100 (1996)
- [丸山 81] 丸山圭三郎: ソシュールの思想, 岩波書店 (1981)
- [Miller 91] Miller, G.: *The Science of Words*, W. H. Freeman and Company (1991); 無藤, 青木, 柏崎 (訳): ことばの科学, 東京化学同人 (1997)
- [村上 94] 村上征勝: 計量的文体研究の威力と成果, 言語, Vol. 23, No. 2, pp. 30-37 (1994)
- [Nagao 84] Nagao, M.: A framework of a mechanical translation between Japanese and English by analogy principle, in Elithorn, A. and Barnerji, R., eds., *Artificial and Human Intelligence*, pp. 173-180, North-Holland (1984)
- [長尾 96] 長尾 真 編: 自然言語処理, 岩波書店 (1996)
- [Niwa 94] Niwa, Y. and Nitta, Y.: Co-occurrence vectors from corpora vs. distance vectors from dictionaries, in *Proc. of COLING'94*, pp. 304-309 (1994)
- [野内 00] 野内良三: レトリックと認識, NHK ブックス (2000)
- [大井 97] 大井耕三, 隅田英一郎, 飯田 仁: 意味的類似性と多義解消を用いた文書検索手法, 自然言語処理, Vol. 4, No. 3, pp. 51-70 (1997)
- [齊藤 97] 齊藤洋典: 心的辞書, 言語の科学 3 単語と辞書, pp. 93-153, 岩波書店 (1997)
- [Salton 83] Salton, G. and McGill, M.: *Introduction to Modern Information Retrieval*, McGraw-Hill (1983)
- [佐藤 96] 佐藤信夫: レトリックの意味論—意味の弾性, 講談社学術文庫 (1996)
- [佐藤 91a] 佐藤理史: MBT1: 実例に基づく訳語選択, 人工知能学会誌, Vol. 6, No. 4, pp. 592-600 (1991)
- [佐藤 91b] 佐藤理史: MBT2: 実例に基づく翻訳における複数翻訳例の組み合わせ利用, 人工知能学会誌, Vol. 6, No. 6, pp. 861-871 (1991)
- [佐藤 97] 佐藤理史: アナロジーによる機械翻訳, 共立出版 (1997)
- [佐藤 98] 佐藤理史: 情報検索, 言語の科学 9 言語情報処理, pp. 51-93, 岩波書店 (1998)
- [Schütze 92] Schütze, H.: Dimensions of meaning, in *Proc. of Supercomputing'92*, pp. 787-796 (1992)
- [Schütze 98] Schütze, H.: Automatic word sense discrimination, *Computational Linguistics*, Vol. 24, No. 1, pp. 97-123 (1998)
- [Sperber 95] Sperber, D. and Wilson, D.: *Relevance: Communication and Cognition, Second Edition*, Oxford, Basil Blackwell (1995); 内田, 中達, 宋, 田中 (訳): 関連性理論—伝達と認知—, 研究社出版 (1999)
- [Sumita 92] Sumita, E. and Iida, H.: Example-based transfer of Japanese adnominal particles into English, *IEICE Transactions on Information and Systems*, Vol. E75-D, No. 4, pp. 585-594 (1992)
- [Sumita 00] Sumita, E.: Lexical transfer using a vector-space model, in *Proc. of 38th Annual Meeting of the ACL*, pp. 425-431 (2000)
- [諏訪 93] 諏訪正樹, 岩山 真: 比喩の計算モデル, 情報処理, Vol. 34, No. 8, pp. 566-575 (1993)
- [鈴木 96] 鈴木宏昭: 類似と思考, 共立出版 (1996)
- [滝澤 96] 滝澤 修: 日本語修辞表現の工学的解析, 大阪大学学位論文 (1996)
- [徳永 99] 徳永健伸: 情報検索と言語処理, 東京大学出版会 (1999)
- [Tourangeau 82] Tourangeau, R. and Sternberg, R.: Understanding and appreciating metaphors, *Cognition*, Vol. 11, pp. 203-244 (1982)
- [Ungerer 96] Ungerer, F. and Schmid, H.: *An Introduction to Cognitive Science*, Addison-Wesley (1996); 池上嘉彦 ほか (訳): 認知言語学入門, 大修館書店 (1998)
- [Uramoto 94] Uramoto, N.: Example-based word sense disambiguation, *IEICE Transactions on Information and Systems*, Vol. E77-D, No. 2, pp. 240-246 (1994)
- [内海 00] 内海 彰: 比喩の認知/計算モデル, *Computer Today*, Vol. 96, pp. 34-39 (2000)
- [内海 01] 内海 彰: レトリックの認知・計算モデル: 隠喩とアイロニー, 認知科学, Vol. 8, No. 4, pp. 8-13 (2001)
- [渡部 01] 渡部広一, 河岡 司: 常識的判断のための概念間の関連度評価モデル, 自然言語処理, Vol. 8, No. 2, pp. 39-54 (2001)
- [安井 92] 安井 泉: ことば遊びのトリックとレトリック, 安井泉 (編), グラマー・テキスト・レトリック, pp. 321-354, くろしお出版 (1992)

2001年11月15日受理

著者紹介

内海 彰 (正会員)



1965年名古屋市に生まれる。1993年東京大学大学院工学系研究科情報工学専攻博士課程修了。博士(工学)。東京工業大学大学院総合理工学研究科システム科学専攻助手, 同研究科知能システム科学専攻講師を経て, 2000年から電気通信大学電気通信学部システム工学科助教授。ことば(言語)やその周辺を対象とした認知や計算に関する研究に従事。日本認知科学会, 情報処理学会, 言語処理学会, ACL, AAAI, Cognitive Science Society など各会員。