

## 幾何的要因に基づく空間語選択モデルに関する考察

小林竜己 寺井あすか 徳永健伸  
東京工業大学

## 1. はじめに

人間の空間理解と言語の間に存在する認知メカニズムを明らかにし、その計算モデルを作ることは、認知科学や言語学の領域のみならず、幅広い応用が期待される工学においても重要な問題である。従来研究においては、心理学実験や提案された計算モデルの扱う範囲が、主に、物体配置の視覚情報と空間表現の静的な組み合わせを対象とした空間語の適合度計算に限定されていた。これは一つ一つの空間語の知覚レベルでの性質を分析する上ではよい方法であったし、この結果、空間語の適合度を空間テンプレートを用いて表現する方法[4, 11]や、二物体関係を対象とした様々な空間語適合度計算モデル[3, 5, 10, 12, 13]が提案され、一定の成功が得られてきたといえる。しかしながら、実世界の複雑な問題への適用を考えたとき、空間認知の研究は、未だ、基礎的な段階にとどまっているといえる。

空間認知の現実的な問題に取り組む場合、取り扱わなければならない多くの要素が存在していることに気づく。対話者間での視覚情報と知識状態の違いを考慮する必要があるだろうし、対話の文脈も扱わなければならないだろう。物体自体や周囲の幾何的要因の多様性や複雑性に対応できるだけのモデルでなければならない。物体間の機能的要因[2]を考慮する場合は、コミュニケーションにおいて言及されている話題や、人々の意図やプラン、さらには常識や領域知識までを視野に入れる必要が出てくる。明らかに、個別の空間語の適合度計算モデルを集めただけでは解決できない多くの状況がある。しかしながら、いきなり問題のスコープを広げてみても、適切な研究の戦略がなければ一般性をもった現実的な計算モデルを得ることは難しい。

著者らは、空間語計算モデルの実世界への適用を考え、以下の問題を解決する糸口を探ってきた[8]。

問題の一つは、適合度分布の計算モデルを実環境にフィッティングするための情報がモデルそのものに備わっていないという欠点である。適合度関数が用意されても、パラメータ決定をするための情報が示されていない。著者らは実験を通して、実環境の視覚情報を取り込んだモデルを作り、モデルを実環境にフィッティングするための手がかりを掴んだ。

二つ目の問題点は、実世界の多様性と複雑さへの適用に関するものである。今までに提案された計算モデルの多くは、参照物体周りに指示物体が存在す

る単純な二物体関係での適合度分布を扱ってきた。二物体に加えてディストラクタが存在する場合は、今まで、投射型空間語や位相型空間語“near”では適合度が低下することが確認されているが[1, 6, 9]、著者らは、空間語毎に、ディストラクタが視覚情報内の異なる幾何的要因に依存する可能性を発見した。

三つ目の問題は、参照表現生成における認知負荷理論[7]の扱いである。この考え方によると、参照枠を決める必要のない位相型空間語の方が、参照枠の決定を必要とする投射型空間語よりも優先的に選択されると説明される。しかし、様々な要因が関係する状況で、認知負荷のみで画一的な判断を下すことは無理があり、実際には「左」が優勢であったり、「近い」が優勢な状況が存在するはずである。著者らはこの問題についても、幾何的要因の多様性に依存して変化する各空間語の適合度の変化を分析した。

本論文は、著者らによるこれらの研究をベースに、従来の空間言語理解の計算モデルの拡張の可能性を検討したものである。特に、各空間語が実世界の対象に対して異なる注意メカニズムを持つことと同時に、言語表現により与えられる参照物体と指示物体への注意が、空間認知の計算モデルにおいて重要な役割を果たしていることを述べる。

以下では、まず、著者らの今までの研究成果を紹介し、上で述べた問題の解決の糸口が得られたことを説明する。次に“near”に関して、Kelleher らの計算モデル[6]と著者らの実験結果[8]の違いを考察する。その上で、言語表現と視覚情報内の幾何的要因によって与えられる注意のモデリングの重要性を導き、空間語選択モデルに必要な要件を検討する。最後に、結論と今後の課題について述べる。

## 2. 日本語の空間語についての実験

著者ら[8]は、日本語の「近い」、「遠い」、「左」を対象に、幾何的要因が変化したとき、二物体の空間関係を記述した説明文の適切さがどのように変化するかを調べた。実験では、コンピュータディスプレイを用いて、説明文と 3-D CG を同時に被験者に呈示した。説明文に用いた空間語は、二つの位相型空間語「近い」、「遠い」と一つの投射型空間語「左」の三つである。一つの画像に対して、

「赤いボールは緑のボールの左にあります」

「赤いボールは緑のボールの近くに 있습니다」

「赤いボールは緑のボールから遠くに 있습니다」

の3パターンを作り、画像の上に表示した。画像には、図1に示すように、参照物体の左方向の基準軸上に、指示物体、参照物体、ディストラクタの三つの物体を描き、指示物体位置(3ケース)、指示物体サイズ(3タイプ)、ディストラクタ位置(表示しないケースを含めて5ケース)を変化させた条件を計画した。これにより、合計135トライアル(=3x3x5x3)を被験者毎にランダムに呈示した。参照物体とディストラクタの物体サイズは、どちらも直径4の球形で、指示物体は直径2, 4, 6の3パターンの球を用意した。色は、参照物体を緑、指示物体を赤、ディストラクタを青とした。被験者は、説明文が画像を適切に表現している度合いを「1:あてはまらない」から「9:あてはまる」までの9件法で評定を行った。

実験の結果、指示物体がL1とL2の位置では「左」、L3では「遠い」の評定が高く、「近い」はL2で「左」に次いで高かった。さらに、以下の結果を得た。

- 1) 「遠い」は、参照物体と呈示画像中の左端の二つに挟まれた領域に被験者の注意が当てられ、左端で最も高い評定値であり、ほぼ直線的に減少し、参照物体近くで最も低かった。
- 2) 「近い」は「遠い」とほぼ逆の形で参照物体から離れるにしたがって適合度が単調に減少した。「遠い」と同様に、画像左端は、近さの程度を規定する基準点として使われることがわかった。
- 3) 「左」は、参照物体から離れるにしたがって緩やかに適合度を下げてはいるが、左端が何らかの基準点であるとは認められなかった。

以上から、「遠い」と「近い」は、言語表現には陽に含まれない左端という画像内の最遠点が参照点として関与するため、その位置情報を適合度計算モデルに取り込まなければならないことが分かった。先に提示した問題1に関連して、位相型空間語の計算モデルを実環境にフィッティングする一つの情報として左端を利用できる可能性が示された。

ディストラクタが各空間語に与える影響の性質もいくつか明らかになった。ここで、Tnは「遠い」の性質(n=1~4)、Cnは「近い」の性質(n=1~3)、Hnは「左」の性質(n=1~3)である。

- T1) ディストラクタが参照物体に近いほど、適合度が高くなる。
- T2) ディストラクタが指示物体と参照物体の間にあると、適合度が高くなる。
- T3) 指示物体サイズがディストラクタより大きいと適合度が高くなる。

T4) ディストラクタが参照物体から見て指示物体よりも外側に置かれると、適合度が低くなる。

C1,H1) ディストラクタが参照物体に近いほど、適合度が低くなる。

C2, H2) ディストラクタが指示物体と参照物体の間にあると、適合度が低くなる。

C3) 指示物体サイズがディストラクタより大きいと適合度が低くなる。

H3) ディストラクタによる影響は、指示物体サイズの違いの影響を受けにくい。

これらのリストは、先に挙げた二番目の問題点に対してディストラクタ効果が空間語毎の規則性を持つ可能性を説明する。また、三番目の認知負荷理論[7]に対しては、単純な空間語の優先順位だけでは解決できない問題の存在を示している。

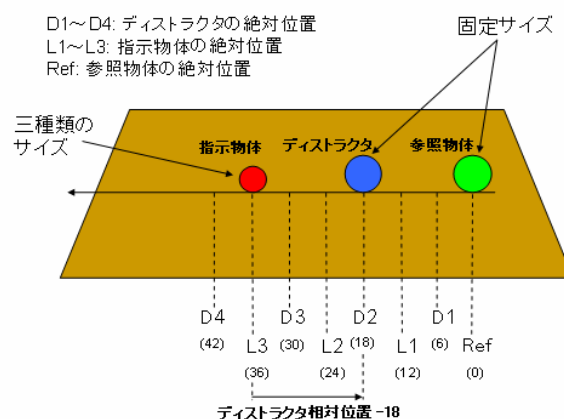


図1 参照物体(緑)、指示物体(赤)、ディストラクタ(青)の配置計画

### 3. 実用的な計算モデルに向けて

#### 3.1 Relative proximity model との比較

次に、著者らの実験における「近い」の実験結果を、英語の“near”に関する Relative Proximity Model (RPM)[6]で再現できるかを検証する。RPMは、ある点xでの物体L固有の絶対近接性の値  $P_{abs}(L,x)$  から、周辺に存在する他物体の中から点xで最も高い絶対近接性の値を持つ物体の絶対近接性の値を引くことにより相対近接性  $P_{rel}(L,x)$  を計算する。絶対近接性  $P_{abs}(L,x)$  は、物体サイズ  $S_{vis}(L)$  と対話における情報新規性  $S_{disc}(L)$  による顕現性パラメータ  $S(L)$  により調節される。 $dist_{norm}(L,x)$  は物体Lから点xまでの距離を正規化したものである。

$$P_{abs}(L,x) = (1 - dist_{norm}(L,x)) * S(L) \quad \dots(1)$$

$$S(L) = (S_{vis}(L) + S_{disc}(L)) / 2 \quad \dots(2)$$

$$P_{rel}(L,x) = P_{abs}(L,x) - \underset{\forall L_n \neq L}{MAX} P_{abs}(L_n,x) \quad \dots(3)$$

表 1 は、著者らの実験結果を RPM による計算結果と比較したものである。参照物体の絶対近接性 (a) は、ディストラクタがない条件での実験で得られた被験者評定値の平均である。同様に、参照物体の相対近接性 (d) は、ディストラクタがある条件での実験結果を用いる。ディストラクタの絶対近接性 (b) は実験で得られた値ではなく、ディストラクタ位置 (=18) での評定値を 9、呈示画面の左端と右端を評定値 1 として結んだ直線を仮定した “near” の適合性をほぼ直線近似してよいことは、実験で得られた (a) より裏付けられている。著者らの条件では、参照物体とディストラクタはサイズが同じであったので、顕現性パラメータ  $S(L)$  は同じ (=1) とした。以上より、(a) - (b) は式 (3) に相当し、RPM の定義に基づき、参照物体から指示物体位置 12, 24, 36 での参照物体の相対近接性 (c) が求まる。(a) - (d) は実験におけるディストラクタの影響、すなわち、RPM におけるディストラクタの絶対近接性に相当する値を計算したものである。

表 1 での比較から即座に、RPM から計算される (c) と著者らの実験結果 (d) がまったく異なる結果であることがわかる。さらに、実験でのディストラクタによる影響 (e) と RPM が仮定するディストラクタ絶対近接性 (b) の値のオーダーが一桁違うことに気がつく。これらの原因を考察した結果、RPM において、参照物体とディストラクタの間の顕現性パラメータを等しいとすることに原因があると考えられる。このようなケースは例えば、視覚情報の中に、言語指示された参照物体に類似した区別不可能な二つ以上の物体が存在する場合と考えられ、その結果、

どの物体もほぼ同じ程度の顕現性の値を得るだろう。このように参照物体を特定できない場合、指示物体に近い方の物体（ここではディストラクタ）が参照物体として適当か、あるいは指示不可能という結果になる。一方、著者らの実験は、言語表現から参照物体が明確に認識できる場合であり、相対近接性が (d) のように高い値ため、RPM の定義に従えば、式 (1) より、ディストラクタが持つ顕現性はかなり低くならなければならない。以上から、[6] で実施された被験者実験は、参照物体とディストラクタが区別できないような設定であった可能性がある。

### 3.2 注意のモデリングの重要性

著者らの実験では、ディストラクタが各空間語に与える影響の性質のリストを明らかにしたが、このリストは、各空間語が互いに異なる注意のメカニズムを持つことを示している。

「遠い」の場合、左端にディストラクタが置かれたときに、「遠い」の適合度を負にすることから、注意の範囲は参照物体から左端までと考えられる。

「左」と「近い」の場合、参照物体から見て、指示物体よりも外に置かれたディストラクタは、ほとんど影響を持たなくなり、注意の影響は急速に減少する。さらに [8] では、ディストラクタの影響下では、指示物体サイズは「遠い」と「近い」には影響するが、「左」には影響しないことが判明している。同様に、呈示画像の左端への注意は、「遠い」と「近い」には利用されるが、「左」には利用されないことが分かった。

表 1 “near”に関する Kelleher らの RPM を用いた相対近接性と小林らの実験結果の比較  
(ディストラクタ位置 18, 指示物体サイズ小の条件)

指示物体位置	(a) 実験での評定値 (ディストラクタ無)	(b) ディストラクタ絶対近接性 (RPM)	(c) RPM を用いた相対近接性 (a) - (b)	(d) 実験での評定値 (ディストラクタ有)	(e) 実験でのディストラクタの影響 (a) - (d)
12	6.929	7.0	-0.071	6.286	0.643
24	3.857	7.667	-3.810	3.429	0.428
36	2.214	5.0	-2.786	2.0	0.214

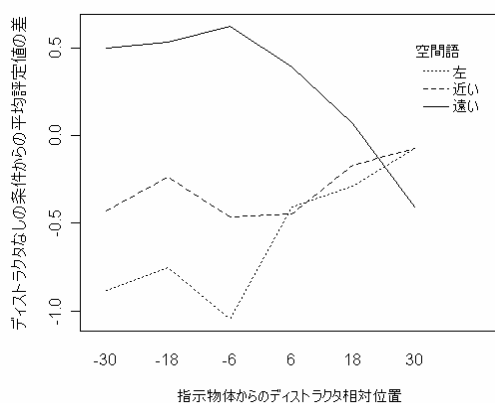


図 2 空間語とディストラクタ相対位置の交互作用

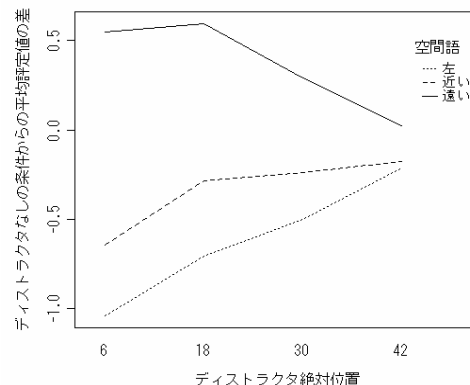


図 3 空間語とディストラクタ絶対位置の交互作用



以上から、幾何的要因に限定した上で、「左」「近い」「遠い」を対象とした空間語選択モデルを考える場合、空間内のどの部分に、どの程度、注意が向けられているかに関するモデリングの重要性が示唆される。これによって、実用的な空間語選択の問題においては、従来までの単純な参照物体周りの適合度分布を扱う計算モデルを、注意のメカニズムをベースにして拡張する必要がある。

### 3.3 言語表現と物体認識の顕現性への影響

Kelleher らの RPM と著者らの実験結果との比較で分かったことは、ディストラクタが持つ顕現性パラメータの大きさが、計算モデルの適用条件に大きく影響するという事実だった。この顕現性決定を左右する要因として、人間の空間語処理における二つのステップの重要性を指摘したい。

一つは、言語表現理解のステップである。言語表現された参照物体と指示物体は、言語表現されないディストラクタと比較して、大きな注意と顕現性を獲得している可能性がある。もう一つは、物体認識のステップである。言語表現された物体を視覚情報内で探索する過程で、一つの物体がユニークに特定できた場合と比べ、類似した参照物体が複数認識された場合は、着目する物体へ与えられる注意とその結果の顕現性は異なるものになる可能性がある。発話者の意図の上ではディストラクタとなるべき物体が、物体認識の曖昧性によって参照物体となる可能性がある時、顕現性は大きなものになると推測できる。物体認識における曖昧性は、話者間での曖昧性解消のための対話が行われるか、または、単純に、指示物体に近い物体が選択されてしまうといった、コミュニケーション上、異なる選択が発生する可能性がある。

このような考察は、従来までの空間認知分野での心理学実験を主体とする研究においてはなされてこなかった。計算モデル考案へのインパクトとしては、言語表現と物体認識の可否が、計算モデルにおける物体の顕現性パラメータへ与える影響を扱わなければならないという点にある。これらも、先に述べた注意のモデリングと同様に、従来の計算モデルを拡張する上で、十分考慮しなければならない。

## 4. おわりに

本論文では、著者らが行った空間認知の実験を元に、実世界への適用可能な空間語選択モデルへの拡張を目指して、いくつかの考察を行った。特に、いままでは考慮されていなかった空間語毎に異なる注意のモデリングと、言語表現と物体認識の可否が物体の顕現性に与える影響の重要性を指摘した。

今後は、今回浮かび上がってきた課題をさらに詳細に分析し、具体的な計算モデルの検討を行っている。

きたい。最終的には、仮想空間ソフトウェア内のエージェントとの対話で、空間表現を計算可能なモデルを実装することを目標にしたいと考えている。

## 参考文献

- [1] Carlson, L. A. & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, **29**, 883--892.
- [2] Coventry, K. R., & Garrod, S. C. (2004). *Saying, Seeing, and Acting: The Psychological Semantics of Spatial Prepositions*. Psychology Press.
- [3] Gapp, K.-P. (1994). Basic meanings of spatial relations: Computation and evaluation in 3D space. In *Proceedings of AAAI-94*, Seattle, WA, 1411--1417.
- [4] Hayward, W.G., & Tarr, M.J. (1995). Spatial language and spatial representation. *Cognition*, **55**, 39--84.
- [5] Kelleher, J., & van Genabith, J. (2005). A Computational Model of the Referential Semantics of Projective Prepositions. In P. Saint-Dizier. (Eds.) *Computational Linguistics: Dimensions of the Syntax and Semantics of Prepositions*, Kluwer Academic Press, 211--228.
- [6] Kelleher, J, Kruijff, G.J., and Costello, F. J. (2006). Proximity in context: An empirically grounded computational model of proximity for processing topological spatial expressions. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, 745--752.
- [7] Kelleher, J. & Kruijff, G.J. (2006). Incremental Generation of Spatial Referring Expressions in Situated Dialogue. In *Proceedings of COLING-ACL'06*. Sydney, Australia Association of Computational Linguistics, 1041--1048.
- [8] 小林 竜己・寺井 あすか・徳永 健伸 (2008). 空間語選択における幾何的要因の影響、『認知科学』, **15**(1). (in press)
- [9] Kojima, T. & Kusumi, T. (2006). The effect of the extra object on the linguistic apprehension of spatial relationship between two objects. *Spatial Cognition and Computation*, **6**, 145--160.
- [10] 小島隆次 (2007). 投影的空間表現語指示領域適合度計算モデル ACAP, 第 24 回日本認知科学会発表論文集, 52--53.
- [11] Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M.A. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space*. Cambridge, MA: MIT Press, 493-529.
- [12] Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General*, **130**, 273--298.
- [13] Tokunaga, T., Koyama, T., Saito, S., & Nakajima, M. (2004). Classification of Japanese spatial nouns. *Proceedings of 4th International Conference on Language Resources and Evaluation: LREC2004*, 1829--1832.