

指示手段の動きによる対象特定効果～人の指さしと矢印の比較～ Specificity effect by the movement of pointing tool: Comparison between human pointing and an arrow

五味 渡海¹, 安田 哲也², 小林 春美^{2*}
Tokai Gomi, Tetsuya Yasuda, Harumi Kobayashi

¹東京電機大学大学院理工学研究科, ²東京電機大学
¹Graduate School of Tokyo Denki University, ²Tokyo Denki University
¹22rmd16@ms.dendai.ac.jp, *h-koba@mail.dendai.ac.jp

概要

本研究では、指さしの指示対象となる事物を聞き手に見せずに話し手が教示行為を行った場合、意図共有にどのような影響があるかを、映像刺激を利用し調べた。結果、話し手が小さく旋回しながら指さした場合と、矢印で指し示した場合で、聞き手は指示対象を事物の部分だと解釈した。聞き手は指示対象物を見なくても、話し手の教示行為によって推測を行うことができること、この解釈は指と矢印では異なる点があることが示唆された。

キーワード : 指さし推論、対象特定、指さしジェスチャー

1. 目的

対面コミュニケーションにおいて、話題の対象物を他者と共有するために、人は言語のみならず非言語情報を使うことが多い。中でも指さしは、意図を伝達する際に有用な非言語コミュニケーション手段である (Kita, 2003)。ある対象物について他者に伝える際、言語情報のみならず非言語情報を用いることは、聞き手の意図推測を助けるために重要である。この指さしにおいて聞き手による指示解釈については2つの仮説がある。1つ目はベクトル(vector)仮説(Butterworth & Itakura, 2000; idea by McGinn, 1981; discussed by Lücking, Pfeiffer, & Rieser, 2015)である。これは、例えば、指さし指が示す対象物については、指さしの方向に人指し指から直線状の光が出ていると仮定し、その線が最初に当たった対象物が指示されたとして推論させるものである。もう1つは語用論的な仮説(Lücking, Pfeiffer, & Rieser, 2015)である。これは、指さしの方向に人指し指から(指示する範囲が)円錐状に放射していると仮定し、その範囲において対象物を推論させるものである。これまでの研究では、指さしの対象物は基本的に聞き手による推論により理解されると考えられているが、ほとんどの研究では、指示対象を共有した状態での推論に重きを置いている。本研究では、この推論について、

対象となる事物を聞き手に見せずに話者が対象指示を行う場合でも語用論的解釈が発生するのかを検討することにした。

これまでの対面による指さし解釈実験では、指さしにおける手・指の動きによって指示解釈が変わると報告している研究がある。Kobayashi and Yasuda (2015)では、弧を描くような指さしを行った場合に遠くの対象が指示されたと解釈されるかについて報告した。弧を描くような指さしの方が典型的な指さしより、遠くの対象を指示していると解釈されることを示唆した。Takahashi, Yasuda, and Kobayashi (2017)の研究では、Kobayashi and Yasuda (2015)に似た状況で、指さしの形は変えず、指さしの動きのみを操作した実験を行った。指先がやや下向きに曲がった指さしを行ったが、直線的な動きの場合には、参加者が遠くの対象物を指示したと解釈することは難しく、弧を描くような動きの場合では遠くの対象物を指示したと解釈したことを報告した。

対象特定は、手首も含むジェスチャーの動きだけではなく、指さしにおける指自体の動き(Kobayashi, Yasuda, & Liszkowski, 2022)や視線シフト(Yasuda & Kobayashi, 2022)の要因も重要な役割を果たす。Kobayashi, Yasuda, and Liszkowski (2022)では、部分名称学習場面を利用し、指さしの推論が部分名称獲得(i.e., 対象特定)に寄与するかを調べた。対象に対して、接触指さしを行い、かつその対象周辺を擦るような強い語用論的手がかりを使用した場合は、2歳半児でも精緻な対象特定をすることができた。その一方、典型的な指さしでは、指示された対象を部分と解釈することは難しく、精緻な対象特定はなされなかった。視線シフトに関して、Yasuda and Kobayashi (2022)では、相手の顔を見た後に対象を見るような視線シフト、相手の顔のみを見る(i.e., アイコンタクト継続)という2種類の語用論的手がかりが対象特定に寄与するのかを調べた。

結果、アイコンタクト継続による推論のみでは、成人でさえ精緻な対象特定をすることができなかったが、アイコンタクトをした後に対象物を見るような推論の手がかりが与えられると、4歳半児と大人が精緻な対象特定を行ったことを示した。

対面コミュニケーションにおける指示対象の特定は、通常、話し手と聞き手が対面状態で、環境を共有して指示対象物への共同注意を成立させて行われることが多い。しかし、近年普及が著しい zoom 等を利用した遠隔コミュニケーションでは、話し手と聞き手が空間を共有しておらず、映像から話し手の指示行為の様子を聞き手が充分確認することが難しい。このことから、互いに事物を視認し、共同注意しているという前提でコミュニケーションを行うことは難しいことが考えられる。加えて、昨今新型コロナウイルスの影響などにより、リモートワークや遠隔授業が導入され、同じ空間を共有しないコミュニケーションの機会が増加している。果たして、この遠隔状況下において、指さしについての語用論的推論は行われるのだろうか。

本研究では、このような状況下においても、語用論的解釈が起こるのかを調べるために、社会的手がかりや指示表現を操作し、語用論的解釈が起こる要因を検討した。指示表現においては、人間の指さしと矢印(→)の違いを調べた。矢印の手がかりは、知覚的に顕著な手がかりであり、先端部分の近くの対象を指示していると解釈されることが様々な研究で示されている(Finke & Pinker, 1982; Kingstone, et al., 2003 等)。遠隔対話で矢印を使って指示対象物を示す場合があるため、この手がかりを検討することにした。なお、矢印を用いたこれまでの人間の認知に関する研究では、反応時間や注意を検討する研究が多く存在する一方、推論に関する研究は少なかった。もし、ビデオ映像において、矢印手がかりに語用論的推論が生じるならば、新奇な知見を提示できる可能性がある。

本研究は、指示行為により対象を特定するためには対話の参加者による推論が必要であるという立場を取り、推論に寄与すると考えられる手がかりとして、1) 指示表現(指さし、矢印)、2) 視線シフト(シフトあり、なし)、3) 指示の動き表現(円運動あり、なし)を操作し、参加者による対象特定がどのように起こるかを調べた。先行研究では、指示を行う事物が対話の参加者間で互いに見える状況で調べられていたが、本研究では、遠隔対話で想定される、どの図が指示対象物か聞き手には見えていないが、図が複数含まれた図版全体は

共有できているという状況で調べた。さらに語用論的手がかりが推論に寄与するのかについて、指さしと矢印という異なる手段を利用して調べた。

本研究の予測は、2つ挙げられる。1) 指示が指さしにより行われた場合には、語用論的な推論がなされるために、指さしの手がかりが強くなれば、対象物が見えない場合であっても語用論的推論が発生し、精緻な対象特定が行われる。よって、参加者は指示された無意味語が対象とされる範囲が狭まっていると推論する結果、指示された対象は事物の部分であると解釈するだろう。2) 指示が矢印により行われた場合には、知覚的手がかりによる推論がなされ、矢印という人工物であるがゆえ、他の手がかりは推論を行う際に寄与しにくくなり、精緻な対象特定はなされない。よって参加者は指示された対象は事物の全体であると解釈するだろう。

これらを検討するために、それぞれの要因を操作した刺激を作成し、遠隔コミュニケーション状況で起こりうる映像を作成した。指示対象の事物はタブレットに提示されており、話し手は聞き手に見えない方向でタブレットを保持し、タブレットには対象事物の図版が提示されている状況であった。よって、タブレットに提示されている図版(計5個の対象物が描かれている)は、話し手は確認できるが、聞き手は映像では確認できなかった。そのため、聞き手には話し手が見ている図版全体をあらかじめ提示しておいた。

2. 方法

参加者

大学生・大学院生 99 人が実験に参加した(平均年齢: 19.9 歳; 年齢レンジ: 18-24 歳)。参加者の母語はいずれも日本語であった。参加者は、全員実験前に実験同意書に同意して実験に参加した。なお、11 人はスマートフォンからの回答であったため、刺激を十分確認できないと判断し分析から除外した。結果、88 人分のデータを使用した。

実験条件

実験条件は、指示表現(指さし、矢印の2水準)、視線シフト(シフトあり、なしの2水準)、指示の動き(動きあり、なしの2水準)が操作された。

指示表現(RE: Referring Expression)条件における指さし(Pointing)水準では、話し手が、指さしを用いて、タブレット端末の表示されるモニターの中央から約 2cm

離れた地点を指すものであった(図 1a)。一方、矢印 (Arrow)水準では、指さしと同様にタブレット端末が表示されているモニターの中央から約 2cm 離れた地点を指すが、その表現は矢印であった(図 1b)。

視線シフト (GS: Gaze shift) 条件における、視線シフトが起こらない(No-Shifting)水準では、教示に関わらず、話し手が聞き手とアイコンタクトを行っているようにみえる直視を行い続けた(図 1c)。一方、視線シフトが起きる(Shifting to monitor)水準では、話し手はまずアイコンタクトのような直視を行った後、教示時にタブレット端末に映っている画像に視線を向けた(図 1d)。

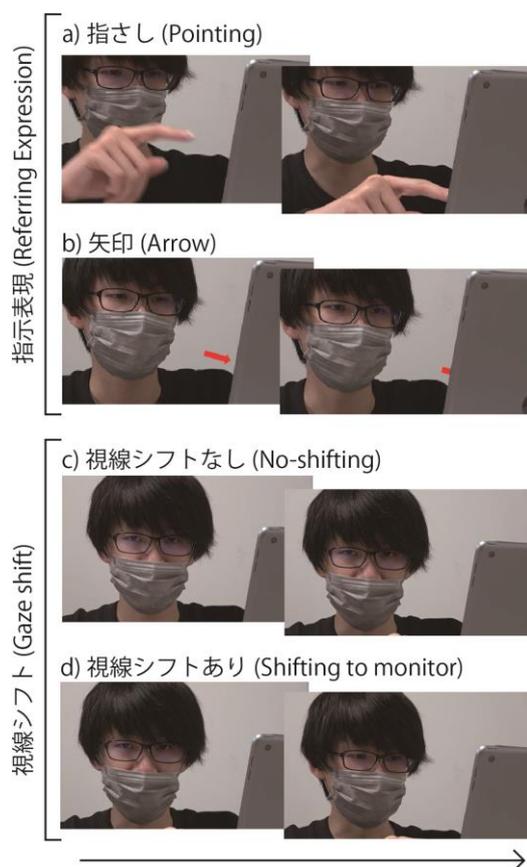


図 1：指示表現条件と視線シフト条件

指示の動き表現(Expression of motion)条件における動きに円運動を伴う(Motion)水準では、指示する行為(指、または矢印)が半径 1 cm 程度の円を、1 秒間に 2 回程度描きながらタブレット端末をさし示した。一方、動きがない条件(No-motion)水準では、動きを加えず単純に、指さしや矢印等で、タブレット端末をさし示した。

実験刺激

刺激は、条件の全ての組合せで 8 通りになるように、指示表現のムービークリップを用意した。このムービークリップは、ビデオカメラ (FDR-AX40, Sony) を使用し、実験者の指示表現を録画した後に、Adobe Premiere Pro で編集を行ったものであった。また、これらのムービークリップはいずれも 16 秒間でできており、5 秒間の沈黙ののち 1 回目の名称教示発話(「これは xx です。」)と指示行為、その後 3 秒間の沈黙ののちに 2 回目の名称教示発話(「これは xx です。」; 1 回目と同様)と指示行為という構成になっていた(図 2)。名称教示には、新奇語(無意味語)を利用した。利用した無意味語に関しては、Kobayashi (1997)より無意味語のリスト(ヘク、ムタなど)を参照した。

参加者が実験の回答を行うにあたって、Google フォームを用いてアンケートを作成し、ムービークリップと動画内の男性の話し手が教示したとする対象の画像を埋め込んだ(図 3)。

指さし (Pointing) + 視線シフトあり (Shifting to monitor)



図 2：実験刺激(指さし+視線シフトあり+動きなし)

※ 参加者に対し話し手による指さしや矢印がどの図をさしているかは、ムービークリップ刺激のいずれにおいても表示されない。



図 3:参加者に提示された指示対象画像

※ オレンジ線の番号は事物の部分で、そうでない番号は事物の全体を指している。

手順

実験は Google フォームを利用した。参加者は心理学を受講する学生であり、Google フォームの実験に参加する前に口頭で実験に関しての確認を行った。参加確認の取れた学生は、zoom のブレイクアウトルーム機能

を用いて、約 25 人ずつの小部屋に分け、参加者に実験の参考となる画像を見せながら実験の事前説明(図 4)と実験同意に関する説明を行い、実験参加の同意が取れた学生のみ、アンケートの URL を配布した。

その後、参加者は受け取った Google フォームアンケートの指示に従って回答を行っていった。なお、動画にアクセスしてもらう前に実験者は、「動画の人物が発言した A は何番の物をさしていたと思いますか？(オレンジ線で指し示しているものは物体の一部を、そうでないものは物体全体に対しての番号です。)」と伝えた。回答の際には、参加者には必ず 0 番から 9 番のいずれかを選択してもらった。

実験概要

Googleフォームを用いています。

「りんご」はアップル(英語)、エパル(マレーシア語)、メーレ(ルーマニア語)、アル(インガリ語)のように様々な言い方があります。

右図のような画像を動画内で指し示すのでどれを指しているかを推測してください。

出題例: アルはAです。

回答例: 1番(実験では選択式です。)

※同じ事物を違う単語で表す場合もあります。

実験にかかる時間は5分前後です。質問などがあればメインルームのチャット欄で聞いていただいで大丈夫です。



図 4:参加者への課題説明画像

分析と統計モデリング

分析は、統計ソフトウェア R (ver. 3.5.1; R Core Team, 2020)、lme4 パッケージ(ver. 1.1-26; Bates, Mächler, Bolker, & Walker, 2015)、lmerTest パッケージ(ver. 3.1-3; Kuznetsova, Brockhoff, & Christensen, 2017)、ggeffects パッケージ(ver. 1.0.2; Lüdtke, 2018)、MuMin パッケージ(ver. 1.43.17; Bartoń, 2020)、emmeans パッケージ(ver. 1.6.3; Lenth et al., 2018) を利用した。目的変数は、事物の部分を選択した場合を 1 とするようなコーディングを行った。要因に関しては、ダミーコーディングを行った後に、エフェクトコーディング(e.g., -0.5, 0.5)を行った。

目的変数を部分名称推測(Part response)、固定効果を指示表現(Referring Expression(RE); 矢印“arrow”, 指さし“pointing”)、視線シフト(Gaze Shift (GS); シフトなし“no-shifting”, シフトあり“shifting-to-monitor”)、指示の動き表現(Expression of Motion(EM); 動きなし“no-motion”, 円運動“motion”)とそれらの交互作用とし、参加者(Participants)と各条件をランダム効果とした最大モデル($\text{glmer}(\text{Part response} \sim \text{RE} * \text{GS} * \text{EM} + (1 + \text{RE} + \text{GS} + \text{EM} | \text{Participants}))$)として、一般化線形混合モデル(GLMM)を構築した。なお、この最大モデルを構築す

る際に、lme4 パッケージの glmer 関数を用い、最適化アルゴリズムである BOBYQA (Bound Optimization by Quadratic Approximation)を用いた。この最大モデルに対し、dredge 関数を利用してモデルを選定した結果、最大モデルと同様のモデルが選定された($\text{df} = 18, \text{logLik} = -463.583, \text{AIC} = 963.2, \text{weight} = 0.457$)。よって、分析は最大モデルを利用することとした。

3. 結果

指示表現、視線シフト、指示時の動きが部分名称解釈に影響を与えるかについて、GLMM を利用した分析を行なった。その結果、指示時の動き($\beta = 0.509, SE = 0.228, z = 2.235, p = .025$)が有意であった。また、指示表現と指示時の動きの交互作用($\beta = 0.951, SE = 0.343, z = 2.774, p = .006$)、及び指示表現と視線シフトと指示時の動きの二次の交互作用($\beta = 1.791, SE = 0.680, z = 2.633, p = .008$)が有意であった。なお、切片に有意な傾向が認められた($\beta = 0.172, SE = 0.094, z = 1.828, p = .068$)。この 2 次の交互作用に関して、emmeans パッケージの pairwise を利用し、単純主効果を調べた。

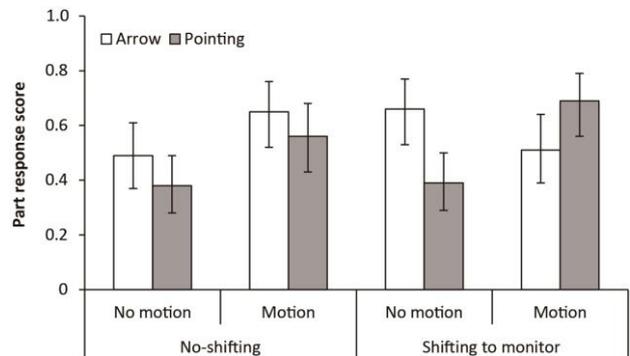


図 5:部分名称得点。各条件の値は、ggeffects パッケージの ggpredict 機能を用い最大モデルから求めた。

指示表現では、視線シフトがあり、かつ動きがない場合($z.ratio = 3.078, p = .002$)に、また視線シフトがあり、かつ動きがある場合($z.ratio = -2.070, p = .0384$)に有意な差が認められた。視線シフトがない場合に違いは認められなかった($z.ratios < 1.0, ps > .19$)。視線シフトでは、動きがなく、矢印表現で指示された場合($z.ratio = -2.049, p = .041$)のみ有意な差が認められた。指示時の動きでは、指さしで指示が表現され、かつ視線シフトがない場合($z.ratio = -1.982, p = .048$)に、また指さしで指示が表現され、かつ視線シフトがある場合($z.ratio = -3.331, p = .0009$)に有意な差が認められた。さらに、矢印で指示

が表現され、かつ視線シフトがない場合($z.ratio = -1.809$, $p = .071$)に有意な傾向が示された。

4. 考察

指さして表現された場合、視線シフトに関わらず、動きがない場合よりも動きがある方が、より多く部分を指示していると解釈された。一方、矢印で表現された場合には、動きがない場合においてのみ、視線シフトがある場合の方が動きがない場合よりも、部分を指示していると解釈された。視線シフトなしでは、指示が矢印でなされた場合は違いがなかったが、視線シフトが起こり、かつ円を描くような指さしが行われた場合は、単に指さしが行われた場合に比べ、より多く部分を指示していると解釈された。一方、視線シフトが起こり、矢印で指示された場合は、矢印が動く場合よりも動かない場合の方がより多く部分を指示していると解釈された。

指さしにおいては視線シフトに関係なく、円運動ありの方が円運動なしに比べて部分解釈が多く行われた。矢印においては円運動が無い場合においてのみ、視線シフトがある方が、視線シフトが無い時に比べて部分解釈が多く行われた。また、視線シフトがある時は、指さしは円運動ありの方が、矢印は円運動なしの方が、それぞれ部分解釈が多く行われた。また、視線シフトがない時も指さしでは同様に円運動ありの方が円運動なしに比べて、聞き手の部分解釈が多く行われた。

これらの結果より、予測「1」参加者は推論の手がかりに応じ指示された対象は部分として解釈する」は支持された。したがって、ビデオ刺激であっても、視線シフトや指の円運動という語用論的な要因の強度が、聞き手の推測に影響を与え、精緻な対象特定を促したと考えることができる。

予測「2」参加者は矢印については、推論のための手がかりの利用が難しいために指示された対象は全体として解釈する」においては、支持されなかった。興味深いことに、矢印で指示表現がなされ、かつ矢印に動きが加わらなかった場合には、視線シフトがある場合の方が、視線シフトがない場合よりも、指示している対象を部分であると多く解釈していた。よって、矢印という知覚的手がかりにおいても視線シフトという社会的手がかりが、推論に影響を及ぼす可能性を見出した。また、矢印が円運動をしながらさし示すことにより、指でさし示した時と違い、矢印によって指示された範囲が、あたかも円柱状のように見えやすくなった

ことが考えられる。これにより、話し手は円柱の範囲である事物全体を示している、と聞き手は推測したのかもしれない。

本研究では、指示をした対象物が見えない場合においても、視覚的共同注意が成立する場合と同様に指さしによる推論が行われることがわかった。よって、対象特定に関する推論は、視覚的共同注意が成立しなくても、意図的な共同注意が成立することで行われる可能性が高い。知覚的手がかりである矢印においても、視線シフトと矢印の動きという手がかりが加わることで緻密な推論が行われていた。このことから、矢印という人工的な手段であっても遠隔コミュニケーションにおいてコミュニケーション意図が見出される可能性を示唆し、またこの知見は遠隔コミュニケーションだけではなく、対面コミュニケーションへも適用できる可能性が考えられる。今後、動画だけではなく、対面場面やAR(Augmented Reality)環境下でも起こりうるのかを検討していきたい。

謝辞

本研究は、MEXT/JSPS 科研費 JP17H06382(HK), JSPS 科研費基盤研究(B) 20H01763 (HK)の助成を受け行なわれた。

参考文献

- [1] Butterworth, G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology*, 18(1), 25-50.
- [2] Finke, R. A., & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8(2), 142.
- [3] Kita, S. (Ed.). (2003). *Pointing: Where language, culture, and cognition meet*. Psychology Press
- [4] Kobayashi, H. (1997). The role of actions in making inferences about the shape and material of solid objects among Japanese 2 year-old children. *Cognition*, 63(3), 251-269.
- [5] Kobayashi, H., Yasuda, T., & Liszkowski, U. (2022). Marked Pointing Facilitates Learning Part Names: A test of lexical constraint versus social pragmatic accounts of word learning. *Journal of Child Language*, First View, 1 – 15, <https://doi.org/10.1017/S0305000921000891>
- [6] Kobayashi H., and Yasuda, T. (September, 2015). Pointing to an invisible object behind a wall: Comprehension of pointing with a bent index finger. *Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science*, pp.477-481. Torino, Italy.
- [7] Lücking, A., Pfeiffer, T., & Rieser, H. (2015). Pointing and reference reconsidered. *Journal of Pragmatics*, 77, 56-79.
- [8] McGinn, C. (1981). The mechanism of reference. *Synthese*, 157-186. <https://doi.org/10.1007/BF00399010>
- [9] Takahashi, H., Yasuda, T., and *Kobayashi, H. (July, 2017). Enforced pointing gesture can indicate invisible objects behind a wall. In G. Gunzelmann, et al. (Eds.), *Proceedings of the 39th*

Annual Conference of the Cognitive Science Society
(pp.3291-3295). Austin, TX: Cognitive Science Society

- [10] Yasuda, T., & Kobayashi, H. (2022). Ostensive gaze shifting changes referential intention in word meanings: an examination of children's learning of part names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <https://doi.org/10.1037/xlm0000859>