LEBS
Vol. 13 No. 1 (2022) 1–5
Letters on Evolutionary Behavioral Science

Understanding Ostensive Behavior in Making Inferences of Referential Intentions

Harumi Kobayashi[1,*], Oga Kobori[2], Yasuo Ihara[3], Hiroyuki Yaguchi[1], Tetsuya Yasuda[1]

[1]Tokyo Denki University, Hatoyama, Hiki-gun, Saitama, 350-0394, Japan

[2]Gradute School of Tokyo Denki University, Hatoyama, Hiki-gun, Saitama, 350-0394, Japan

[3]The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

[*]Author for correspondence (h-koba@mail.dendai.ac.jp)

In ostensive-inferential communication, the speaker's communicative intention must be explicitly presented to the listener, which can be done with both verbal and nonverbal cues. Pointing gesture is an important tool in human ostensive communication. Its interpretation and potential roles in human evolution, however, have not yet been explored. When people produce referential actions, more effortful referential actions may promote different interpretations in the recipient than facile, less effortful actions. In this study, the experimenter demonstrated an effortful pointing, and the participant interpreted its meaning. When effortful pointing was used (i.e., pointing that entails touching an object while tilting one's body forward looking directly at the stimulus), participants thought that a part of the object was being indicated rather than the whole object. This study discussed the relationship between these findings and pedagogy on stone-tool making.

Keywords
pointing gesture, ostensive behavior, relevance, pedagogy

Introduction

Among primates, human communication is unique and can be characterized by its "ostensive-inferential" nature (Gergely, 2013; Scott-Phillips, 2015; Tomasello, 2008). Sperber and Wilson (1995) described ostensive-inferential communication as: "Ostensive-inferential communication consists in making manifest to an audience one's intention to make manifest a basic layer of information. It can therefore be described in terms of an informative and a communicative intention" (p. 54). Thus, in ostensive-inferential communication, the speaker's communicative intention must be explicitly presented to the listener, which can be done using language as well as nonverbal cues such as eye gaze, facial expressions, and gestures including pointing. Then, the listener assumes that some valuable information was conveyed by the speaker; therefore, the listener must pay attention to the information to interpret the message. This ostensive nature of human communication seems to suggest that, before the emergence of language, human nonverbal communication already had some unique aspects of ostensive communication. Csibra and Gergely (2009) proposed the theory of natural pedagogy, which stated that humans have an inherent ability to effectively teach others and to be effectively taught by others. Csibra and Gergely (2011) suggested that natural pedagogy based on ostensive signals developed as an evolutionary adaptation and made the transmission of various human technologies possible, such as stone-tool making skills. Based on archaeological evidence, Gärdenfors and Högberg (2017) discussed different levels of teaching and associated stone-tool making technologies. Thus, to study the origins and the evolution of human communication, it is crucial to understand ostensive communication that involves nonverbal behaviors, such as pointing gestures. In the field of evolution of human behavior, the importance of ostensive communication is not fully recognized.

The pointing gesture is one of the fundamental referential actions (Kita, 2003). There are two major types of pointing gestures depending upon the intention with which they were employed. One can be named socio-centric use, and the other, ego-centric use. Socio-centric use involves inferential communication. People may convey information about objects by using a canonical pointing gesture with extended-index finger aiming at an object, by pointing in various ways using one's elbow, foot, jaw, or face (Cooperrider et al., 2018), or using other referential gestures such as showing or placing (Clark, 2003). Ego-centric use involves thinking and controlling one's attention. Garber and Goldin-Meadow (2002) showed that both adults and children use pointing when they solve problems such as Tower of Hanoi. In terms of attention control, "finger-pointing and calling" is a method in which operators point and call an object (target) by its name or status aloud before, after, or during its operation. It is a well-known method for human error prevention (Haga et al., 1996; Shinohara et al., 2009).

The main focus of this study was to understand the socio-centric use of pointing gestures and to assert that pointing gestures are commonly interpreted using their accompanying actions and contexts. People use at least two types of socio-centric pointing gestures according to the effortfulness of pointing. One is *facile canonical pointing*, which is performed adroitly without effort and typically refers to whole objects, and the other is *effortful close pointing*, which typically refers to object parts. To test whether effortfulness is effective for specifying referential intentions, a situation of learning part-names was used. Learning part-names can be challenging, even for adults, if only facile pointing is used to refer to object parts. In general, facile pointing is ambiguous, because it

doi: 10.5178/lebs.2022.90
Received 30 August 2021.
Accepted 17 September 2021.
Published online 14 February 2022.

can only direct the listener's attention to a certain referent, and the referent can be a whole object, a part of an object, or some features of the referent.

Another important aspect in comprehension of the pointing gesture is the context or situation in which the pointing is presented (Tomasello, 1999, 2008). If the situation provides sufficient reasons for rational interpretation of certain pointing gestures, effortful pointing may be more easily understood. For example, if a person points at a stone in water using effortful pointing such as repetitive pointing action (i.e., tapping motion), the observer may think that the referent is the seashell stuck on the stone, rather than the stone as a whole. The observer may think the pointer has a good reason for not using close pointing because the referent object is in water. Any surrounding situation seems to have the potential to be informative and be used to interpret the pointing gesture, and people will pick up relevant information for interpretation.

An evolutionary perspective on pointing gestures explores the possibility that pointing might have played an important role in pedagogy for stone-tool making. We can assume that hominins used pointing gestures to teach about important parts, in particular, parts of stones in stone tool making. If this is the case, then, pointing must be proven to be important to teach object parts as well as whole objects. In addition, it must be explored how and when people recognize the speaker's communicative intent through observation of pointing gestures according to context.

In this study, we investigated whether adults interpreted a combination of referential gestures and postures as effortful information. We used the learning part-names paradigm (Kobayashi, 2007) with some modifications. We considered three conditions: *pointing, accessibility, and test*. There were two types of pointing conditions: facile and effortful. In the facile condition, the experimenter kept one's posture straight, looking down at an object, and pointed at a part of the object (a nut attached to a U-shaped bolt) inside a transparent plastic display box. In the effortful condition, the experimenter tilted his posture forward and looked at the object while pointing at it. There were two situations in the accessibility condition: more accessibility and less accessibility. An obstacle (transparent plastic cover over the object) was openable, and the object was either accessible (more accessibility) or it was not (less accessibility). Two types of test conditions were used: isolated test and transfer test. In the isolated test, the participant chose a part choice from a set of an isolated part, a whole object (that does not include the target part), and a distractor. In the transfer test, the participant must recognize that the target part was a part of another whole object of a different shape. Thus, the transfer test was more complex and difficult than the isolated one. We predicted that if the experimenter pointed at the object part (nut) with more effort, and if the stimulus could not be directly touched or pointed at in close range such as less than 1cm, that is, it was less accessible, the participants will infer that the object part is being referred to rather than the whole object (bolt). The reason is that the participant may think the speaker has a proper reason to engage in effortful pointing.

Methods

Thirty-six undergraduate and three graduate students participated in this experiment (*M* age: 21.5). All participants were Japanese monolingual speakers. The experimental stimulus comprised video clips that were incorporated into a questionnaire survey created using Microsoft Forms. This study was examined and approved by Tokyo Denki University Research Ethics Committee (Number 31–103).

A condition of referential actions (i.e., postures) consisted of two styles of pointing gestures: effortful and facile. In the effortful condition, the experimenter pointed at the object in a plastic cover (box), along with a forward tilted posture, looking into the side of the box. In the facile condition, the experimenter pointed toward the object part in the same way, but with a straightened posture, looking straight down at the box. A condition of accessibility consisted of two types: more accessibility and less accessibility. In the more accessibility condition, the experimenter easily removed the plastic cover over the object and subsequently covered the object again. In the less accessibility condition, the experimenter rigorously tried to open the box twice but failed. Thus, it was impossible to remove the cover. In addition, each test consisted of the four sets of objects (e.g., nut set) used by Kobayashi (2007) with a slight modification. Stimulus video clips consisted of various combinations of the three aspects—context-, training-, and test-phase—and was created using Adobe Premiere Pro (version 13.1). For instance, in the nut set stimulus with less accessibility and with effortful pointing, the experimenter first demonstrated that the box was locked, and then pointed to the object in a more effortful posture and said in Japanese, "Kore wa muta (i.e., nonsense word) desu" ("This is (a) muta") twice. Next, in the test phase, the isolated test was conducted, followed by the transfer test. There were four object sets, and for each set, there were two postures; additionally, for each posture, there were two accessibilities, resulting in a total of 16 video clips.

In the experiment, each participant accessed an experimental site organized using Microsoft Forms (Office 365); then, they proceeded to watch each stimulus and chose one of three choices, that is, part choice, whole choice, and distractor, for a total of four times. The details of the experimental method are shown in the supplemental material.

Results

Each participant response that guessed the nonsense word as the object part was scored. Figure 1 shows part response scores in each condition.

A mixed three-way ANOVA (2 *pointing* × 2 *accessibility* × 2 *test*) was conducted on part response scores using R software (3.5.1; R Core Team, 2020) and statistical function of ANOVA (Iseki, 2021). *Pointing* and *accessibility* were between-participant variables, and *test* was a within-participant variable.

The ANOVA revealed that main effects of *pointing*, F (1,32) = 5.370, p = .027, η^2 = .103, and *test*, F (1,32) = 10.390, p = .003, η^2 = .53, were significant. The main effect of *accessibility* was not statistically significant, F (1,32) = 3.315, p = .078, η^2 = .64. In addition, *pointing* had a large

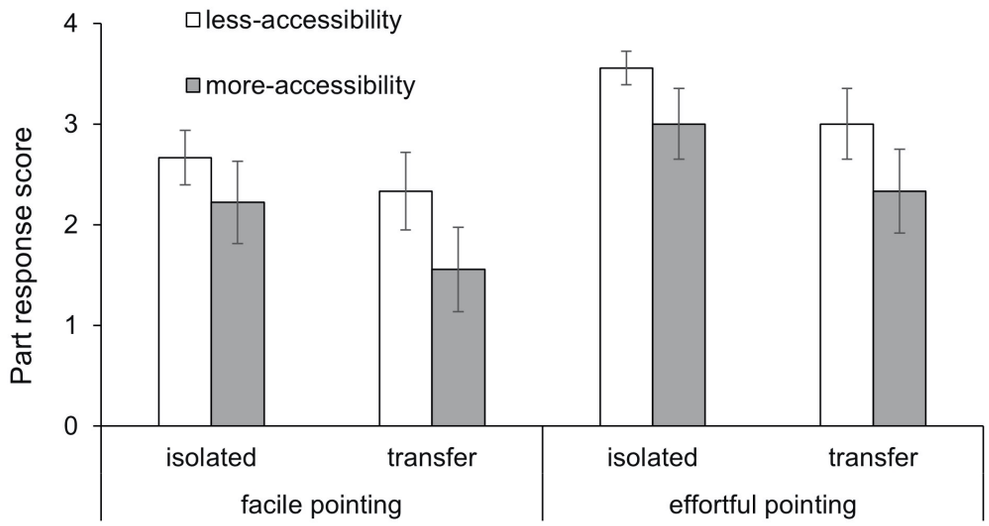

Figure 1. Part response scores in pointing each condition. In facile pointing, the experimenter pointed with less effort. In effortful pointing, the experimenter pointed with more effort.

effect size, and *accessibility* and *test* had medium effect sizes. There were no interaction effects between *pointing* and *accessibility*, F (1,32) = 0.000, p = 1.00, *accessibility* and *test*, F (1,32) = 0.416, p = .52, η^2 = .002, and *pointing* and *test*, F (1,32) = 0.104, p = .75, η^2 < .001, and among *pointing*, *accessibility*, and *test*, F (1,32) = 0.104, p = .75, η^2 < .001.

Regarding *pointing*, the effortful pointing (M = 2.97; SD = 0.90) elicited more part response score than the facile pointing (M = 2.19; SD = 1.08). Regarding test, the isolated test (M = 2.86; SD = 1.06) elicited more part response score than the transfer test (M = 2.31; SD = 1.29). In addition, the difference in part response scores in the condition with less accessibility (M = 2.89; SD = 0.94) and the condition with more accessibility (M = 2.28; SD = 1.11) was not statistically significant.

Overall, participants interpreted the given names as the object part in the effortful condition and in the less complex test. Accessibility was not demonstrated to be effective.

Discussion

The results indicated that the interpretation of pointing gestures changed depending on the speaker's posture and the type of the test. When effortful pointing was used, participants interpreted that the object part was being indicated more frequently than when facile pointing was used. Participants interpreted the given name of the object part when the object was simple (i.e., less complex) rather than when it was embedded in the other object (i.e., more complex). These results suggest that recognizing the object part when it was embedded in a new whole object was more difficult than recognizing the object part alone. If additional ostensive cues such as pointing with circular motion around parts used by Kobayashi et al. (in press) were used, participants could have recognized parts more easily. Therefore, a follow up study on effortful pointing may be needed to refer to the part related to the new whole object. Finally, accessibility was not effective. One reason for this result may be that, in this experiment, the cover's "openable" nature was only demonstrated before the act of pointing and not at the occurrence of pointing. Ostensive signals, such as pointing gestures, may recruit relevant information on here-and-now basis. Thus, while the pointing posture was utilized, the proceeding demonstration of the openability of the cover was not.

Another possibility regarding the interpretation of the results is that the experimenter's pointing action in the *effortful* condition may have attracted the participant's attention to the task as a whole. Specifically, similar to what Samuelson and Smith (1998) discussed regarding the cues related to highlighting and focusing a learner's attention, it can be argued that the tilted upper body may have attracted the participant's attention. One way to separate these predictions pertaining to ostensive cues and attentional focus may be to test participants' responses to accidental (non-intended) actions (Diesendruck et al., 2004)—such as the speaker accidentally tilting their body during the experiment, for example, by realizing that the chair was broken or avoiding a flying bug. Although tilting the upper body does capture participants' attention, this posture itself should not be interpreted as a referential cue; especially since ostensive communication may inherently include the manipulation of the recipient's attention. In everyday ostensive communication, the recipient's attention focus is usually manipulated using pointing gestures, posture, gaze direction, and language. Therefore, perceiving whether the attentional focal point is simply observed or manipulatively presented by the speaker may be important. Thus, future studies must further explore how people recognize ostensive communication based on various information provided by the speaker and the environment. This study showed that when the speaker's action is perceived as intentional, effortful referential actions facilitate a more specific interpretation, such as referring to parts of an object. Overall, this experiment indicated that reasoning, in the case of a complex action

such as pointing, tracking eye gaze, and considering the nature of the stimulus objects, leads the observer to guess what is being indicated given the ostensive context. Thus, effortful pointing can promote specific inferences such as the referent being part of the object, not the whole object.

Our findings can provide an insight to ancient human stone-tool making. Based on experimental results on adults teaching how to make stone tools to other adults, Lambao et al. (2017) suggested that using language in addition to gesture made learning lithic knapping more effective. Morgan et al. (2015) examined teaching tool making by comparing among imitation/emulation, basic teaching, gestural teaching, and verbal teaching including gestures. In their experiment, they described basic teaching as "in addition to demonstrating tool production, tutors could also manually shape the pupil's grasp of their hammerstone or core, slow their own actions and reorient themselves to allow the pupil a clear view (this condition replicates teaching reported in non-human primates.)" (p. 7). They described gestural teaching as "tutors and pupils could also interact using any gestures, but no vocalizations" (p. 7). They showed that gestural teaching doubled and verbal teaching quadrupled the probability of a viable flake per hit compared to reverse engineering. However, they did not report the types of gesture that the teachers used. In this gestural teaching condition, we can assume that teachers effectively used effortful pointing to convey information about parts of stones that must be dismissed to produce sharp flakes.

The results of this study can also be related to Relevance Theory (Sperber & Wilson, 1986). The theory claims that every act of ostensive communication must be interpreted on the part of the communicator to establish the optimal relevance of the act in the communicative situation (Sperber, 2016). The results of the current study seem to accord with the theory because the speaker's pointing action, including the posture, was taken to be relevant for making inferences. Other information such as accompanying actions of pointing gestures (Kobayashi et al., in press) and immediate contexts may also be used as relevant information in understanding the communicator's intention. In this study, effortfulness was evaluated by controlling the posture that accompanied the speaker's pointing gesture. Tilted posture, indicating more effortful pointing, was interpreted as referring to an object part. Furthermore, this evaluation of effortful pointing was influenced by task complexity. When the task was more complex, such as a target part being embedded in a completely different object, another effortful pointing was required. A more detailed and advanced method of communication, such as language, may play an important role for specification. This may be a reason why language emerged in the course of cultural transmission. Future studies must explore the mechanisms of the relevance of non-verbal behavior, that is, how human actions are evaluated to be effortful and how such information contributes to effective natural pedagogy for human evolution. In addition, exploring such human factors will contribute to designing smooth and barrier-free communication.

Acknowledgements

We would like to thank all the participants. We would also like to thank Editage (www.editage.jp) for English language editing. This study was supported by JSPS/MEXT KAKENHI JP17H06382 (H.K.) and JP17H06381 (Y.I.), and JSPS KAKENHI JP20H01763 (H.K.).

Supplementary Material

An electronic supplementary material is available online.

References

Clark, H. (2003). Pointing and placing. In S. Kita (Ed.). *Pointing: Where language, culture, and cognition meet* (pp. 243–268). Hillsdale: Erlbaum.

Cooperrider, K., Slotta, J., & Núñez, R. (2018). The preference for pointing with the hand is not universal. *Cognitive Science*, *42*(4), 1375–1390. https://doi.org/10.1111/cogs.12585

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157. https://doi.org/10.1098/rstb.2010.0319

Diesendruck, G., Markson, L., Akhtar, N., & Reudor, A. (2004). Two-year-olds' sensitivity to speakers' intent: An alternative account of Samuelson and Smith. *Developmental Science*, *7*(1), 33–41. https://doi.org/10.1111/j.1467-7687.2004.00320.x

Gärdenfors, P., & Högberg, A. (2017). The archaeology of teaching and the evolution of Homo docens. *Current Anthropology*, *58*(2), 188–208. https://doi.org/10.1086/691178

Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science*, *26*(6), 817–831. https://doi.org/10.1016/S0364-0213(02)00087-3

Gergely, G. (2013). Ostensive communication and cultural learning: The natural pedagogy hypothesis. In J. Metcalf, & H. Terrace (Eds.), *Agency and joint attention* (pp. 139–151). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199988341.003.0008

Haga, S., Akatsuka, H., & Shiroto, H. (1996). Laboratory experiments for verifying the effectiveness of "finger-pointing and call" as a practical tool of human error prevention. *Japanese Association of Industrial/Organizational Psychology Journal*, *9*, 107–114. (In Japanese). https://doi.org/10.32222/jaiop.9.2_107

Iseki, R. (2021) ANOVA-Kun version 4.8.5. http://riseki.php.xdomain.jp/index.php

Kita, S. (2003). Pointing: A foundational building block of human communication. In Kita, S. (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 1–8). Erlbaum.

Kobayashi, H. (2007). The effect of touching object parts on learning novel object part names among young children and adults. *Studies in Language Sciences*, *6*, 61–76. https://ci.nii.ac.jp/naid/10021940168/

Kobayashi, H., Yasuda, T., & Liszkowski, U. (in press). Marked pointing facilitates learning part names: A test of lexical constraint versus social pragmatic accounts of word learning. *Journal of Child Language*.

Lombao, D., Guardiola, M., & Mosquera, M. (2017). Teaching to make stone tools: New experimental evidence supporting a technological hypothesis for the origins of language. *Scientific Reports*, *7*, 14394. https://doi.org/10.1038/s41598-017-14322-y

Morgan, T. J., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., Cross, C. P., Evans, C., Kearney, R., de la Torre, I., & Whiten, A. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, *6*, 6029. https://doi.org/10.1038/ncomms7029

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. C*hild Development*, *69*(1), 94–104. https://doi.org/10.1111/j.1467-8624.1998.tb06136.x

Scott-Phillips, T. C. (2015). Nonhuman primate communication, pragmatics, and the origins of language. *Current Anthropology*, *56*(1), 56–80. https://doi.org/10.1086/679674

Shinohara, K., Morimoto, K., & Kubota, T. (2009). The effect of "Finger-pointing and Call" on orientation of visual attention. *The Japanese Journal of Ergonomics*, *45*(1), 54–57. (In Japanese). https://doi.org/10.5100/jje.45.54

Sperber, D. (April, 2016). New perspectives on ostensive communication. In T. Hasegawa (Chair). *Symposium conducted at the meeting of The Evolutionary Origin and Neural Basis of the Empathetic Systems*. University of Tokyo.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (Vol. 142). Harvard University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press. https://doi.org/10.2307/j.ctvjsf4jc

Tomasello, M. (2008). *Origins of human communication*. MIT Press. https://doi.org/10.7551/mitpress/7551.001.0001