Gestures as effective teaching tools: Are students getting the point?

- A study in pointing gesture in the English as a Second Language classroom -

Paul RUMME Hirofumi SAITO Hiroshi ITO Misato OI Alberto LEPE

Graduate School of Information Science, Nagoya University

Furo cho, Chikusa-ku, Nagoya, 464-8601 Japan

Abstract To examine the effectiveness of human pointing gestures and laser pointers in teaching English to L2 learners, we assigned three junior high school classes to conditions with 1) pointing gestures, 2) laser pointer, or 3) without additional visual support. The students' tasks were to complete a pretest to evaluate their knowledge of prepositions, to observe a video illustrating the use of prepositions, and to complete a posttest to evaluate the learning effect of the video. Students in the human gesture condition had more correct responses on the posttest than students in the laser condition and the control condition, suggesting gestures are effective attractors for teaching prepositional concepts.

Keywords: pointing gesture, laser pointer, English as a Second Language, affective function, substantive function

Hand gestures that co-occur with verbal expression are basic elements of our daily communication and often appear to play a role in conveying a message to a listener. Gestures as conveyors of information have been the focus of a great deal of study and debate in recent years, as researchers attempt to disentangle the complex roles gestures play in spoken language.

The majority of gesture research has thus far examined the *spontaneous production* of gesture in experimental settings (e.g., Alibali, Heath & Meyers, 1999; de Ruiter, 2000; Jacobs & Garnham, 2007; Kita, 2000; Kita & Ozyurek, 2003).

Less research has investigated the degree to which *comprehension* of a speaker's message is actively facilitated by the *perception* of the gestures that the speaker produces (e.g., Beattie, G. & Shovelton, H., 1999; Krauss, Chen & Chawla, 1996; Levy, E., & Fowler, C., 2000).

Far less research has examined the perception of *intentionally produced* gesture in pragmatic contexts, such as in a language classroom (Allen, 1995; Sueyoshi, A. & Hardison, D. , 2005).

Gestures as part of non-verbal behaviour have previously been subdivided into four, non-exclusive types according to their semiotic functions: *emblems*, *representational* gestures, *deictic* gestures, and *beats* (McNeill, 1992).

The focus of the present study is on *deictics*, or *pointing* gestures, especially as perceived from the listener's perspective. In particular, we examine how human *pointing* gestures differ from artificial pointing devices when used to convey abstract content in spoken messages. *Pointing* gestures can be used intentionally to indicate concrete referents in speech. However, pointing may function to imply referents not in the gesturer's or the listener's immediate vicinity. Our aim is to clarify whether pointing gestures are effective as external supports for teaching abstract concepts (*prepositions* referring to spatial relationships) in the English as a Second Language (ESL) classroom.

Allen (1995) speculated that the use of *emblems* accompanying teacher's explanations aided students' retention of French phrases via an internalization process in which mental representations were formed at the moment of perception.

McNeil, Alibali and Evans (2000) demonstrated that native English speaking preschoolers' posttest scores improved after observing matching *iconic gestures* that accompany spoken instructions, while mismatching gestures had no significant effect on preschoolers' scores. Kindergarten children's scores, by contrast, were not affected by matching gestures, but were negatively affected by mismatching gestures. These results imply that the reinforcing effect of iconic gestures depends on the relation of the gesture to speech, and as well on the participants' ability to understand complex spoken messages.

Valenzeno, Alibali and Klatzky (2002) showed that children who observed a video-recorded lesson that incorporated pointing and tracing gestures to explain the concept of symmetry scored higher on a posttest. The positive results were attributed to the 'grounding' effect of pointing and tracing gestures in the teacher's explanations about symmetry.

Searches of the literature have not revealed results of research into the effectiveness of laser pointers as attractors in the language classroom.

The current article focuses on the effectiveness of *pointing* gestures used in the English as a Second Language classroom to explain prepositions referring to spatial relationships. Prepositions were chosen as they are often problematic for foreign learners of English because of ambiguities in their usage and as such can prove to be stumbling blocks to achieving fluency.

In order to evaluate their effectiveness, pointing gestures have been categorized according to two main functions. First, pointing can be used to identify specific referents in a speaker's utterance and focus a viewer's attention on a particular referent. This is referred to as the gesture's *substantive* (i.e., analytical, or "cool") function.

Second, the vector produced by human pointing gestures is an *attractor* and functions to hold the listeners' attention. This role of pointing is referred to as the gesture's *affective* (i.e., emotive, or "hot") function.

Combining the two previously mentioned functions results in a combined substantive and affective (i.e., "warm") function that drives the listener's attention to identify and observe the intended referent.

The present experiment comprised three tasks. The participants first completed a pretest designed to assess their English listening comprehension ability.

The second task was to observe a demonstration video with an English narrative that provided examples in the use of prepositions referring to spatial relationships. Under the observation task, the three conditions comprised the identical narration, but differed in the visually superimposed portion.

The third task was a posttest to evaluate the learning effect of the presentation video. The participants received the same spoken messages as in the pretest. Correct response rates for the multiple-choice tests were compared in three conditions.

We hypothesized that participants in the three conditions would show different response rates on the posttest depending on the type of demonstration video shown in their respective condition.

The hypothesis diverged in whether or not the laser pointer shared the same function as the human pointing gesture. If the gesture condition incorporated both hot and cool functions (i.e., 'warm' function) and the pointer condition incorporated only the cool function, the incorrect (pretest) to correct (posttest) response rates (i.e., the *positive* or *facilitating* effect of the training video) were expected to show the highest rate in the gesture condition and the lowest in the control condition, with the pointer condition revealing an intermediate rate. These magnitudes can be described as follows: Gesture (G 'warm') > Pointer (P 'cool') > Control (C).

Method

Participants

Ninety-seven first-year junior high school students (47 girls and 50 boys) in three classes (n = 33, 30, 34) attending a Japanese public school participated in the experiment. The students were all monolingual Japanese speakers, 12 to 13 years of age (M = 12.1 years) and were learning English within the conventional school system for the first time. The students participated in the experiment on a regular school day approximately two months after the beginning of the school year.

Materials

Presentation materials

A video projector (NEC HT 1000) and a projection screen (Kikuchi 80-GUP), and a laptop computer (Panasonic Let's Note CF-W2) programmed with an automated experiment procedure (Explayer) were set up in each of the classrooms. Each setup included a pair of satellite speakers (audio-technica, AT-SP13AV) for the audio portion of the presentation.

The projection screen for the video presentation was set up at the front of each of the classroom in full view of the class. The students sat at their desks, facing the screen, in rows of six students across.

Pretest and Posttest materials

The *pretest* and the *posttest* consisted of a set of 16 prerecorded spoken English messages and printed pages with a series of four graphic illustrations lettered *a*, *b*, *c*, or *d* for each message. Figure 1 shows an example set of illustrations used on the pretest. The pretest comprised an additional two training messages to help orient the student towards the task.

Responses were marked on separate response sheets that were illustrated with the same illustrations presented with the spoken messages.

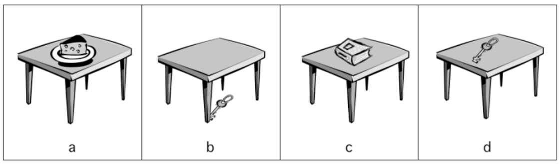

Fig. 1. A sample set of illustrations on the response sheet for the pretest.

Training materials

The *demonstration* videos had been recorded using graphic illustrations of objects and people in everyday situations. These images were used as background illustrations for the video and were displayed on a 32” LCD monitor (Mitsubishi LDT321V) and recorded with a digital video camera (Sony DCR-SR300).

The demonstration video for the *human gesture condition* was recorded with a human hand superimposed (the forearm, wrist, hand and fingers, with the forefinger extended) to indicate portions of the images in the digital video recording. The timing of the pointing gestures coincided with the narrator’s spoken comments about the images. Figure 2 shows an example from the human gesture condition.

A simulated red laser-pointer dot was used to indicate portions of the images in the video recording for the *laser pointer* condition. (The laser pointer dot in the video was added digitally as the point of light from a laser pointer was not visible on the surface of the LCD monitor.) The superimposition of the laser pointer dot was timed to coincide with the narrator’s spoken comments about the images and followed the hand motion from the demonstration video for gesture condition as closely as possible.

The video for the *control* condition showed the identical background images used in the *human gesture* and the *laser pointer condition*, but without additional visual support.

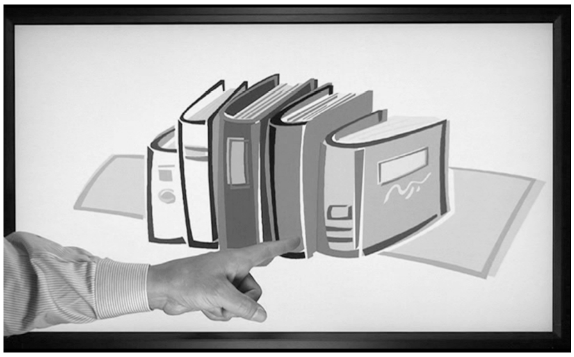

Fig. 2. An illustration from the demonstration video for the *human gesture* condition.

In the postproduction stage, the three presentation videos were edited to the same length (i.e., 7 minutes).

The audio portions for each condition were identical for all parts of the experiment. All of the prerecorded English messages and Japanese instructions were narrated by native speakers.

The spoken messages and illustrations were designed to use vocabulary and images that reflect the syllabus of English classes taught in Japanese public junior high schools. To minimize the priming effect for messages with prepositional content, the content of the spoken messages alternated between *target* messages (n = 8) using prepositions to describe the position and location of objects, and *filler* messages (n = 8) with messages about daily life situations.

Procedure

Pretest The experiment was carried out in the students’ homeroom classrooms with 33 (gesture condition), 30 (pointer condition), and 34 (control condition) participants in each classroom, respectively. The students were instructed to listen to 16 spoken English messages (8 target and 8 filler messages) and to identify one illustration in each row (a, b, c, or d) that most closely matched the message they heard. Responses were recorded on a response sheet using the same series of graphic illustrations shown on the screen with each message.

Assistants in each classroom advised participants regarding the procedure and collected the response forms following the pretest.

Training session: The students in each condition were instructed to observe a seven-minute video presentation about prepositions indicating spatial relationships. The illustrations (n = 24) were organized into four pairs of prepositions (on - under, next to – between, in front of – behind, near - at), with three examples for each half of the four pairs. Illustrations presented in the video were accompanied by an audio track with spoken English messages describing the prepositional relationships between objects and people.

Posttest: The posttest consisted of the same 16 examples and prerecorded messages from the pretest. The messages were presented in different order to that used on the pretest.

As in the pretest, the students were instructed to listen to the 16 prerecorded messages, and to circle the letter of the one illustration in each row that most closely matched the spoken message.

Following the posttest, the students completed an evaluation questionnaire.

Coding: A score of 1 was awarded for each of the correct responses to the sixteen messages, a score of zero (0) was given for incorrect responses. Possible total scores ranged from 0 (zero) to 16 (sixteen) on both the pretest and the posttest.

Results

The mean *correct* response rate across conditions for filler messages was 72% and the correct response rate for the target messages was 45%. The correct response rate for filler messages was higher than for the target messages, and therefore filler messages were excluded from analysis.

The mean correct response rate on the pretest for the target messages in each condition were: *gesture*, 47%, *pointer* 43%, and *control*, 44%. To confirm the homogeneity of the responses on the pretest, a one-way ANOVA with condition (gesture, pointer, control) was carried out. The ANOVA revealed that response rates did not differ significantly as a function of the pretest, $F < 1$.

To detect the effects of the training session, we computed two types of correct responses to the messages on the posttest, a) responses that were correct on both the pretest and the posttest (*correct to correct*), and b) responses that were incorrect on the pretest and correct on the posttest (*incorrect to correct*). Figure 3 shows the correct to correct and the incorrect to correct effect in the three conditions.

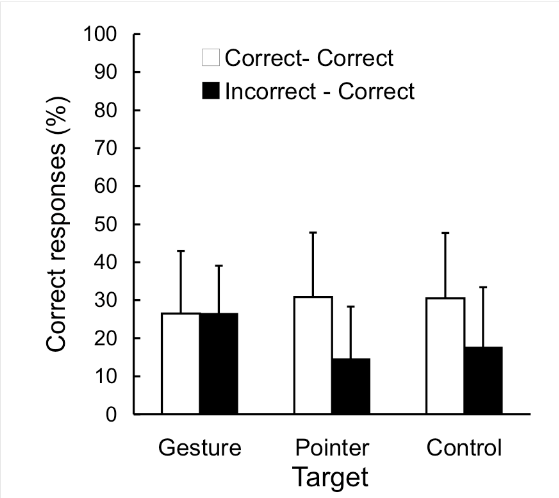

Fig. 3. *Correct to correct* and *incorrect to correct response rates* for the three conditions.

The results of the computation revealed that while the participants in the gesture condition showed a *higher incorrect to correct* response rate (27%), the participants in the pointer condition and the control condition showed a *lower correct response* rate (15% and 18%, respectively).

The results of a 3 (condition: gesture, pointer, control) X 2 (response: correct to correct, incorrect to correct) ANOVA yielded a significant effect for *response*, F (1, 94) = 13.14, p < . 01, whereas *condition* did not yield a significant effect, $F < 1$. The interaction between *condition* and *response* was significant, F (2, 94) = 3.42, $p < .05$.

The significant interaction effect was explored using a simple main effects analysis of the correct responses within each response type (constant positive or facilitatory effect). The simple main effects of the *incorrect to correc*t response rate were significant, F(2, 94) = 6.01, p < .01. The simple main effect of the *correct to correct* response rate was not significant, $F < 1$. An LSD post hoc comparison showed that the *incorrect to correct* response rate for the gesture condition was higher than the both the pointer and the control condition.

The significant interaction effect was further explored using a simple main effects analysis of the correct responses within each condition. The simple main effects of the response type (correct to correct or incorrect to correct response rate) under the pointer condition and the control condition were significant, $F(1, 94) = 7.70, p < .01$, $F(1, 94) = 12.27, p < .01$, respectively, but not significant under the gesture condition, $F < 1$.

The *incorrect to correct* response rate in the gesture condition was greater than in the pointer and control conditions, while the *correct to correct* response rate in the three conditions showed no significant differences. Moreover, the *incorrect to correct* response rate between the pointer and control conditions showed no significant differences. Therefore, the results may be formulated as 'G > P = C'. That is to say that the pointer condition had no substantive function in the present study with regard to teaching abstract concepts.

Discussion

The present study was designed to examine whether gestures and pointers function in the same way by drawing listeners' attention to an abstract subject matter.

We hypothesized that human pointing gestures and artificial means of pointing have cognitively distinct functions. In accordance with the affective ('hot' =1) and the substantive ('cool' = 1) dichotomy, we predicted that human pointing gestures would have an affective plus a substantive function, that is, incorporating both 'hot' and 'cool' functions. Further, we predicted that the laser pointer would have a substantive function, but not an affective function. The results of the control condition were our baseline, without either an affective or a substantive function.

Given that the hypothesis is correct, we would expect to obtain G ('hot' + 'cold' = 2) > P ('cold' = 1) > C (0). However, we obtained the unexpected result revealing G > P = C.

The results obtained in the present experiment make it difficult to determine what exact function G has, as G (2) > P (0) and G (1) > P (0). There are three possible functions within G: a combination of hot and cool, or one of the two functions, either 'hot' or 'cool'.

Due to the relatively lower incorrect to correct response rates in the G condition, we interpret the result G (1) > C (0) based on G ('hot'). If G incorporates both hot *and* cool functions, the added incorrect to correct responses due to the *cool* function in the G condition would be expected to produce a higher incorrect to correct response rate than was actually achieved. The relatively higher 'hot' increase that G (1) produced to the 'cool' response rate in P (0) now remains to be explained.

Before accepting the result G ('hot' = 1) > P (0), we need to examine whether or not P (0) is the result of a malfunction in the experimental operations. In the present experiment, we used a small, statically closed dot that was superimposed on the stimulus to simulate the function of a laser pointer. On a level of physical visibility, the laser pointer condition used in this experiment might have been handicapped in comparison with the illustrative functionality of pointing gestures. Minor modifications to the laser pointer condition may enhance its indication functionality and produce a value of P (1) or P ('cold').

A further question is whether G ('cool') is equivalent to P ('cool'). To answer to this question, we need to consider a salient difference between the pointing gesture and the present pointing device: the abstract pointing function in pointing gesture versus the concrete pointing function in laser pointing. Pointing gestures do not indicate their referents directly, or, more precisely, do not touch their targets, while the laser pointer dot comes into directly into contact with its referent. The abstract inference by a pointing gesture may invite a listener to interpret what a speaker intended to indicate. In contrast, the concrete indication by a laser pointer dot may focus an observer's attention to a specific object too closely, and, accordingly, the observer may fail to attend to a looser relationship between a speaker's intention and target object.

In other words, the referential ambiguity in pointing gestures may have the advantage over a laser pointer in driving the listener's 'affective' or hot cognition, and this ambiguity may take on an especially important role in understanding abstract prepositions referring to spatial relationships in the present experiment.

Further research in pointing used to convey abstract concepts in the ESL classroom must be discussed with relation to the results obtained here. To better understand the effect of gesture in teaching abstract concepts, follow-up experiments focusing on concrete concepts are necessary. The relationship between pointing type (e.g. pointing gestures or artificial means of pointing) and the content (concrete or abstract) is an important avenue for further experimentation.

The results of our experiment suggest that pointing functions in conveying abstract concepts in the language classroom. In particular, the affective stimulus in human pointing gestures appears to play a central role in attracting, and keeping, a learner's attention. In order to develop our understanding of how gesture works in conveying new ideas to viewers, it is important to pursue the central dichotomy of substantive and affective functions in gesture.

References

Allen, L. A. (1995) The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal*, 79, pp. 521 -529

Beattie, G. & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? *Semiotica 123 (1/2)*, pp. 001-030.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology, Vol. 91, No. 4*. pp. 720 – 730.

de Ruiter, J.P. (2000). The production of gesture and speech. In D. McNeill (ed.) *Language and Gesture* (pp. 163 -185). Cambridge: Cambridge University Press.

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language, 56*. pp. 291 – 303.

Kita, S., (2000). How representational gestures help speaking. In D. McNeill (ed.) *Language and Gesture* (pp. 163 -185). Cambridge: Cambridge University Press.

Kita, S. & Oezyuerek, (2003) What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, pp. 16 -32.

Krauss, R., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational gestures tell us? M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 389-450). San Diego, CA: Academic Press

Levy, E., & Fowler, C. (2000) The role of gestures and other graded language forms in the grounding of reference in perception. In McNeill, D. (Ed.), *Language and Gesture*, (pp. 215 – 234). Cambridge: Cambridge University Press.

McNeil, N. M., Alibali, M. W., Evans, J. L. The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior.* 24 (2), pp. 131 – 150.

McNeill, D., (1992). *Hand and Mind: What Gestures Reveal about Thought.* Chicago: University of Chicago Press.

Meltzoff, A.N. (2006). The 'like me' framework for recognizing and becoming an intentional agent. *Acta Psychological*, 124. pp. 26 – 43.

Sueyoshi, A. & Hardison, D. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning, Vol. 55, No. 4,* pp. 661 – 669

Valenzeno, L., Alibali, M., & Klatsky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology, Vol. 28*. pp. 187 – 204.