# Whose Perspective Do Language Comprehenders Adopt During Motor Simulations?

Yusuke Tanaka
Kyoto University
yusuke.tanaka.07@gmail.com

**Abstract:** When processing sentences about performable actions, language comprehenders activate motor systems to perform *mental simulations* (Barsalou 2008) of those events. For instance, Glenberg and Kaschak (2002) demonstrated the *action-sentence compatibility effect* (i.e., comprehending a sentence that implies that action in one direction interferes with real action in the opposite direction). This effect can be observed in the interference between the mental simulations of the content of the sentence the participants read and the real action being performed. However, little is known regarding whose perspective is adopted during such *motor simulations* (i.e., mental simulations in motor areas). I conducted an experiment using two Japanese giving verbs, *age-ru* and *kure-ru*, to address this question. Both verbs mean that the referent of the subject gives something to the referent of the dative object: *age-ru* and *kure-ru* adopt the subject's perspective and the dative object's perspective, respectively. Using the action-sentence compatibility effect methodology, *age-ru* and *kure-ru* sentences were demonstrated to facilitate manual action away from the body rather than toward the body. These results suggest that *age-ru* and *kure-ru* adopt the subject's perspective and that language comprehenders consistently adopt the subject's perspective when simulating another individual's action.

**Keywords: language processing, perspective, mental simulation, motor simulation**

## 1. Introduction

Over the past several decades, many studies have demonstrated that cognition shares systems with perception and action at neural levels. For instance, the primary motor cortex is active when individuals perform and imagine actions (Ehrsson et al. 2003). Similarly, the processing of language that describes perceivable scenes or performable actions results in the activation of neural structures dedicated to actual perceptions or actions. Hauk et al. (2004) demonstrated that reading verbs such as *kick*, *pick*, and *lick* activates motor areas adjacent to or overlapping with areas activated by performing actions with the foot, hand, or mouth. Additionally, the generation of color words or action words activates a region anterior to the area involved in the perception of color or action, respectively (Martin et al. 1995).

Such activation in sensory-motor areas during language processing can be interpreted as reflecting a *mental simulation* of the content of language for the comprehenders. Mental simulation


*Papers in Linguistic Science*, No. 24 (2018), pp. 99–109

is the reenactment of perceptual, motor, and introspective states acquired during experience with the world, body, and mind (Barsalou 2008). Merely hearing or reading language expressions automatically activates sensory-motor areas to simulate the experience of the described scenes. Mental simulation is not a mere epiphenomenon, but an essential process for language comprehension. One piece of strong evidence to support this perspective is that action verb production and comprehension, compared with the processing of nouns, are compromised in patients with motor neuron disease (Bak et al. 2001).

In the early researches on mental simulation, the main concern was what was simulated during language processing, and results revealed that language comprehenders simulate some properties of objects such as color, orientation, visibility, and shape (Connell 2007, Stanfield and Zwaan 2001, Yaxley and Zwaan 2007, Zwaan et al. 2002). Afterward, the main concern shifted to which components of a sentence were the impetus for which aspects of mental simulation. Bergen and Wheeler (2010) suggested that content words tell language comprehenders what to mentally simulate and what brain regions to use in performing these simulations, and grammatical constructions such as aspect modulate how those simulations are performed.

Other researches focused on further aspects of mental simulation: The perspective of language comprehenders during mental simulation. Brunyé et al. (2009) demonstrated that mental simulations adopt a particular perspective, and this perspective is modulated by a grammatical person. In their experiment, participants read sentences that differed only in whether the subject was *I*, *you*, or *he*. After reading each sentence, the participants saw an image that depicted the event from either an internal perspective (i.e., a person performing the action) or external perspective (i.e., a person observing the action; Figure 1). They observed that the participants were faster at responding to a picture of an internal perspective when the subject of a sentence was either *I* or *you*, whereas they were faster at responding to a picture of an external perspective when the subject of a sentence was *he*. These results demonstrated that language comprehenders engage with mental simulations from an internal perspective when reading sentences about themselves and from an external perspective when reading sentences about another individual.

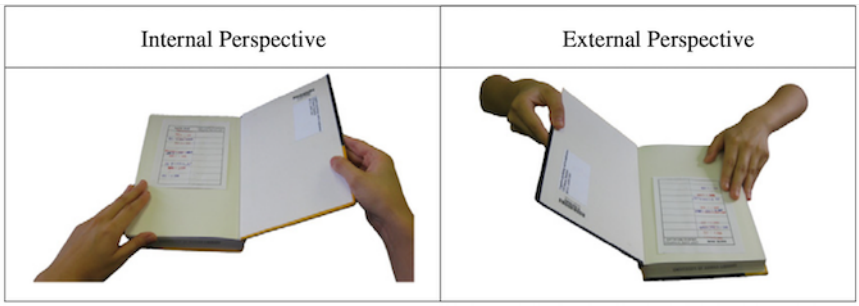


Figure 1 Internal and external perspective images (Sato and Bergen 2013: 364)

Brunyé et al. (2009) used English for their experiment. Additionally, Sato and Bergen (2013) conducted experiments in the same manner as Brunyé et al. (2009), but used Japanese where grammatical subject pronouns could be omitted when the subject was inferable from the context. They observed that the same results were replicated when the sentences contained a subject; however, when sentences without a subject were used, no such effect was observed.

These researches suggest that mental simulations adopt either an internal perspective or an external perspective at least when sentences have a subject. However, the internal perspective in the second-person condition and the external perspective in the third-person condition are the same perspective, in that both are the comprehenders' perspective. Since mental simulation is the reenactment of perceptual, motor, and introspective states, language comprehenders naturally adopt their perspective during mental simulations. Nonetheless, because individuals can imagine a scene from another individual's perspective, engaging with mental simulations of a described scene from another individual's perspective is possible.

To the best of my knowledge, no research has investigated the possibility of adopting another individual's perspective during mental simulations. However, with the *action-sentence compatibility effect* (ACE), this possibility is considerable. ACE is when a sentence implies an action in one direction, and individuals have difficulty making a sentence-sensibility judgment requiring a response in the opposite direction (Glenberg and Kaschak 2002). In the experiment, participants read sentences that implied hand actions toward their body (e.g., *Andy delivered the pizza to you*) or away from their body (e.g., *You delivered the pizza to Andy*) and judged if the sentences made sense by moving their hand toward or away from their body. It was observed that the participants had difficulty moving their hand in the direction opposite to that of the sentence.

Glenberg and Kaschak (2002) used sentences including the word *you* that referred to the participants (language comprehenders), but what if no such word is in a sentence? Bergen and Wheeler (2005) conducted an experiment in the same manner as Glenberg and Kaschak (2002) using sentences about another individual's action instead (e.g., *John is opening the drawer*). If language comprehenders adopt the external perspective when simulating another individual's action, ACE cannot be observed. However, Bergen and Wheeler (2005) reported that ACE was observed even when sentences were not about the participants.

These results demonstrate that language comprehenders mentally simulate the direction of the action described in a sentence. However, these can also be interpreted in terms of perspective: Language comprehenders adopt an internal perspective not only when reading sentences about their action but also when reading sentences about another individual's action. If the participants had adopted an external perspective when reading a sentence about another individual's action, ACE would not have been observed in the experiment in Bergen and Wheeler (2005). Therefore, although language comprehenders adopt an external perspective when engaging with visual simulations of another individual's action (Brunyé et al. 2009, Sato and Bergen 2013), they adopt an internal perspective when engaging with motor simulations of another individual's action.

One question that arises here is as follows: Whose perspective do language comprehenders

adopt during motor simulations? When a sentence has more than one participant (e.g., a subject and a dative object), language comprehenders can choose whose perspective to adopt. The most likely possibility is that language comprehenders consistently adopt the perspective of the subject, which has been given a special status in theoretical linguistics. Notably, the *Surface Structure Empathy Hierarchy* prescribes that "it is easier for the speaker to empathize with the referent of the subject than with the referents of other NPs in the sentence" (Kuno 1987: 211). In this study, *empathy* is the degree to which a speaker adopts the perspective of a participant in a sentence. Therefore, this hierarchy suggests that the subject's perspective is most likely to be adopted.

However, Kuno (1987) demonstrated the exception to this hierarchy: The *Empathy Condition on Giving Verbs in Japanese* as in (1). Kuno (1987) described the difference between two of the Japanese giving verbs, *yar-u* and *kure-ru*. Both verbs mean that the referent of the subject gives something to the referent of the dative object. The two verbs differ only in the perspective the speaker adopts when using these verbs.

(1) a. *yar-u* E (subject) ≥ E (dative)
b. *kure-ru* E (subject) < E (dative)

(Kuno 1987: 248)

In (1), E(x) represents the degree to which a speaker adopts the perspective of the referent of x. Accordingly, this condition means that the speaker's perspective is closer to the subject (or neutral) when he/she uses *yar-u*, whereas the speaker's perspective is closer to the dative object when *kure-ru* is used by him/her. Therefore, *kure-ru* is not acceptable in (2a) because it is difficult to evoke the context where the speaker adopts the stranger's perspective instead of his brother's perspective. For the same reason, *yar-u* is not preferable in (2b).

(2) a. *Ototo=ga shiranaihito=ni hon=o yat-ta/??kure-ta.*
brother=NOM stranger=DAT book=ACC give-PAST
'My brother gave a book to a stranger.'
b. *Shiranaihito=ga ototo=ni hon=o kure-ta/??yat-ta.*
stranger=NOM brother=DAT book=ACC give-PAST
'A stranger gave a book to my brother.'

Based on the Empathy Condition on Giving Verbs in Japanese, a prediction is that *kure-ru* yields mental simulations from the perspective of the dative object, and *yar-u* yields mental simulations from the perspective of the subject. If these perspective-takings are psychologically real, language comprehenders would be faster at moving their hands toward their body when reading *kure-ru* sentences and away from their body when reading *yar-u* sentences. However, if language comprehenders always adopt the perspective of the subject, they consistently move their hand away from their body faster. The experiment that I report in Section 2 was designed to provide convincing evidence to support the latter by demonstrating that the actions designated by *yar-u* and *kure-ru* sentences promote real action away from the body.

# 2. Experiment

## 2.1 Participants

A total of 16 right-handed individuals aged between 18 and 28 years participated in this experiment. The participants received 500 yen for their participation.

## 2.2 Materials

I created 48 meaningful critical sentences and 48 non-meaningful filler sentences. The meaningful sentences included either *age-ru* or *kure-ru* such as (3). In modern Japanese, *yar-u* is rarely used as a giving verb. Thus, sentences such as (2a) are unfamiliar to most Japanese individuals. For this reason, *age-ru*, which can be used almost interchangeably with and more frequently than *yar-u*, was used in this experiment.

Two classes of non-meaningful filler sentences were used. Filler sentences in the first class were unacceptable because of the unnatural colocation of the verb and the object, as in (4a). If there had been only this type of filler sentence in the experiment, participants could have judged sensibility by reading only the verb phrases and would not have had to read the whole sentences. To prevent the participants from reading merely some parts of the sentences, another set of filler sentences such as (4b) was prepared. Filler sentences in this class were unacceptable because the referents of the dative objects were not living things but physical objects.

(3) a. *Ototo=ga Hanako=ni ame=o age-ta.*
brother=NOM Hanako=DAT candy=ACC give-PAST
'My brother gave a piece of candy to Hanako.'

b. *Hanako=ga Ototo=ni hon=o kure-ta.*
Hanako=NOM brother=DAT book=ACC give-PAST
'Hanako gave a book to my brother.'

(4) a. *Hanako=ga Ototo=ni niku=o ui-ta.*
Hanako=NOM brother=DAT meat=ACC float-PAST
'Hanako floated a piece of meat to my brother.'

b. *Ototo=ga bakudan=ni hoshi=o age-ta.*
brother=NOM bomb=DAT star=ACC give-PAST
'My brother gave a star to a bomb.'

## 2.3 Procedures

The participants were asked to sit in front of a personal computer and read sentences displayed in the middle of the screen. They were then instructed to decide as quickly as possible if each

sentence made sense by pressing the right-shift key or the left-shift key, labeled either "natural" or "unnatural," respectively. These two keys were equidistant from a middle key (the N key). The sentences were revealed while the middle key was held down. The keyboard was rotated to be perpendicular to the participant and the three keys placed in a straight line.

For each trial, the participants pressed and held the middle key to reveal a written sentence until they decided if the sentence made sense (or not), whereupon they released the middle key and pressed either "natural" or "unnatural." The participants were instructed to use only the index finger of their right hand during the experiment, and were initially randomly assigned to the yes-is-near condition (i.e., participants pressed the proximal key when they considered a sentence sensible) or yes-is-far condition (i.e., participants pressed the distal key when they considered a sentence sensible). Halfway through the experiment, the experimenter swapped the locations of the "natural" and "unnatural" labels; thus, all the participants were exposed to both conditions. Eight practice trials preceded each half of the experiment.

## 3. Results

Two participants and three sentence items that had accuracy lower than 80% were excluded from the analysis.[1] All incorrect responses, one missing trial, and all responses smaller or greater than 3 standard deviations from the mean per participant were also excluded from analysis. This resulted in the exclusion of about 6% of the data.

The participants' reaction times (the number of seconds used to read the sentence and release the middle key) and the movement times (the number of seconds between releasing the middle key and pressing the response key) were analyzed. These two analyses were performed because although Glenberg and Kaschak (2002) reported that ACE was observed regarding reaction time, Bergen and Wheeler (2005) reported that ACE was observed regarding movement time. They suggested that when sentences are not about the language comprehenders, there is a delay in motor activation, and ACE is observed in the movement times rather than in the reaction times. Because the sentences used in this experiment were not about language comprehenders, my hypothesis is that ACE would be observed in movement times rather than reaction times.

I used R (R Core Team 2018) and the lme4 package (Bates et al. 2015) to perform a generalized linear effects analysis of the relationship between the verb (*age-ru* or *kure-ru*) and response direction (toward or away from the body). As fixed effects, I entered verb and response direction with interaction term into the model. As random effects, I had intercepts for participants and items, as well as by-participant and by-item random slopes for the effect of verb and response direction. Because the distributions of the reaction times and the movement times were positively skewed, the distribution and link function were specified with gamma and log, respectively. *P*-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

## 3.1 Reaction times

The *age-ru* sentences and *kure-ru* sentences derived faster responses away from the body, as shown in Figure 2. However, the main effect of response direction as well as of the verb were not statistically significant ($p$ = .787 and $p$ = .967, respectively). In addition, the interaction effect between the two was also not significant ($p$ = .456).

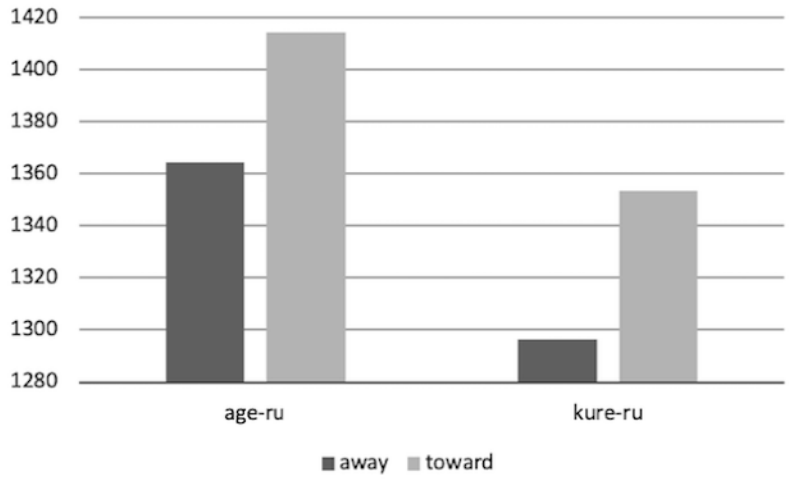


Figure 2 Mean reaction times (in milliseconds) as a product of *age-ru* and *kure-ru* sentences, when manual responses were toward or away from the body.

Table 1 The result of the analysis of the reaction times

| | Estimate | Std. Error | $t$-value | $p$-value |
|---|---|---|---|---|
| Verb | −0.001778 | 0.040879 | −0.043 | 0.967 |
| Direction | 0.019417 | 0.071435 | 0.272 | 0.787 |
| Verb × Direction | −0.034304 | 0.045985 | −0.746 | 0.456 |

## 3.2 Movement times

The movement times are presented in Figure 3. A significant main effect of response direction was observed: The participants took longer to move their hand toward their body compared to when they moved it away from their body ($p$ = .043). Neither the interaction effect between the verb and response direction nor the main effect of the verb was significant ($p$ = .481 and $p$ = .568, respectively).

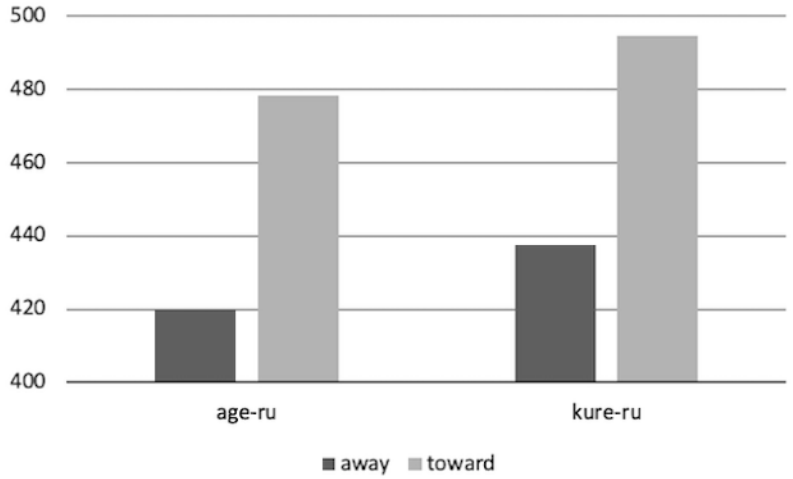


Figure 3 Mean movement times (in milliseconds) as a product of *age-ru* and *kure-ru* sentences, when manual responses were toward or away from the body.

Table 2 The result of the analysis of the movement times

| | Estimate | Std. Error | *t*-value | *p*-value |
|---|---|---|---|---|
| Verb | −0.04768 | 0.08243 | −0.578 | 0.568 |
| Direction | 0.17480 | 0.07597 | 2.301 | 0.043 |
| Verb × Direction | −0.04447 | 0.06271 | −0.709 | 0.481 |

## 4. Discussion

No significant interaction effect was observed between the verb and response direction in the reaction times and the movement times; thus, there is no such process where *age-ru* drives motor simulations from the perspective of a subject and *kure-ru* from that of a dative object. However, regarding the movement times, a significant main effect of response direction was observed, demonstrating that the participants were faster at moving their hand away from their body compared with toward their body when they read *age-ru* and *kure-ru* sentences.[2] These results demonstrated that *age-ru* and *kure-ru* sentences facilitate hand action away from the body, and that language comprehenders consistently adopt the perspective of the subject when they mentally simulate another individual's action described by a sentence.

Notably, the results in this experiment do not mean that language comprehenders never adopt another individual's perspective. Also, two other possibilities should be considered. The first pos-

sibility is that perspective is not simulated in motor areas but in other perceptual or introspective areas such as visual areas. In fact, simulations in motor areas and simulations in visual areas are not always consistent. As mentioned in Section 1, language comprehenders adopt an external perspective when engaging with visual simulations of another individual's action, while they adopt an internal perspective when engaging with motor simulations of another individual's action.

The second possibility is that *age-ru* and *kure-ru* are not relevant to perspective. Yokoyama et al. (2009) found much greater activation in the left premotor area for the processing of *kure-ru* sentences than for the processing of *age-ru* sentences. They suggest that this activation reflects the attentional shifting process of linguistic empathy in *kure-ru* sentence comprehension. Although language comprehenders cannot determine the empathy relationship between the subject and the dative object until they read *age-ru* or *kure-ru*, these words come at the end of the sentence. Therefore, language comprehenders pay attention to the subject at first, and they shift their attention to the dative object if the verb is *kure-ru*. This study indicates that the notable difference between *age-ru* and *kure-ru* may be attributed to the matter of attention rather than of perspective.

## 5. Conclusion

To assess whose perspective comprehenders adopt during motor simulations, an experiment was conducted using the ACE methodology. It was found that the participants were faster at moving their hand away from their body than toward their body when they read *ager-u* and *kure-ru* sentences. The interpretation of the results I consider most notable is as follows: comprehenders consistently adopt the subject's perspective when understanding sentences regarding another individual's action. Combined with the results of Glenberg and Kaschak (2002), where comprehenders took their perspective when sentences were about own action, the *Empathy Hierarchy for Motor Simulation* can be proposed as follows: comprehender > subject > other noun phrases.

## Notes

1. This percentage is based on Bergen and Wheeler (2005).
2. Technically, a possibility is that it was simply easier for the participants to move their hand away from their body than toward their body. However, given that the participants were faster at moving their hand toward their body as compared to moving it away from their body when reading filler sentences, this possibility is not worthy of consideration.

## References

Bak, Thomas H., Dominic G. O'Donovan, John H. Xuereb, Simon Boniface and John R. Hodges. 2001. Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain* 124:

103–120.

Barsalou, Lawrence W. 2008. Grounded cognition. *Annual Review of Psychology* 59: 617–645.

Bates, Douglas, Martin Mächler, Benjamin M. Bolker and Steven C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.

Bergen, Benjamin K. and Kathryn B. Wheeler. 2005. Sentence understanding engages motor processes. In *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society*.

Bergen, Benjamin K. and Kathryn B. Wheeler. 2010. Grammatical aspect and mental simulation. *Brain and Language* 112: 150–158.

Brunyé, Tad T., Tali Ditman, Caroline R. Mahoney, Jason S. Augustyn and Holly A. Taylor. 2009. When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science* 20: 27–32.

Connell, Louise. 2007. Representing object colour in language comprehension. *Cognition* 102: 476–485.

Ehrsson, H. Henrik, Stefan Geyer and Eiichi Naito. 2003. Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *Journal of Neurophysiology* 90: 3304–3316.

Glenberg, Arthur M. and Michael P. Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin & Review* 9: 558–565.

Hauk, Olaf, Ingrid Johnsrude and Friedemann Pulvermüller. 2004. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41: 301–307.

Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago/London: The University of Chicago Press.

Martin, Alex, James V. Haxby, François M. Lalonde, Cheri L. Wiggs and Leslie G. Ungerleider. 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270: 102–105.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Sato, Manami and Benjamin K. Bergen. 2013. The case of the missing pronouns: Does mentally simulated perspective play a functional role in the comprehension of person? *Cognition* 127: 361–374.

Stanfield, Robert A. and Rolf A. Zwaan. 2001. The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12: 153–156.

Yaxley, Richard H. and Rolf A. Zwaan. 2007. Simulating visibility during language comprehension. *Cognition* 105: 229–236.

Yokoyama, Satoru, Kei Yoshimoto, Tadao Miyamoto and Ryuta Kawashima. 2009. Neurophysiological evidence of linguistic empathy processing in the human brain: A functional magnetic resonance imaging study. *Journal of Neurolinguistics* 22: 605–615.

Zwaan, Rolf A., Robert A. Stanfield and Richard H. Yaxley. 2002. Language comprehenders mentally represent the shapes of objects. *Psychological Science* 13: 168–171.

# メンタル・シミュレーションにおける視点に関する実証的研究

田中　悠介

人は動作を描写する文を理解する際，その動作を運動領域でシミュレーションすることで理解している。例えば Glenberg and Kaschak (2002) は，文が描写する動作の方向と，その文を読んだ後に行う動作の方向が一致している場合，その動作がより早く実行されることを示した。この結果は，文を理解する際に運動領域において生じたシミュレーションが，実際の動作を促進した結果と解釈される。その一方で，このようなシミュレーションに際してどの視点が取られるのかに関しては不明な点が多い。本研究では，日本語の授与動詞「あげる」と「くれる」を利用して，Glenberg and Kaschak (2002) と同様の実験を実施した。これらの動詞はともに，「主語の指示対象が与格目的語の指示対象に何かを授与する」という事態を描写する。Kuno (1987) によれば，話し手は「やる（本実験で使用する「あげる」と同義の動詞）」を使用する場合は主語の視点を，「くれる」を使用する場合は与格目的語の視点を取っている。したがって，あげる文を読んだ後は手を奥に動かす動作が，くれる文を読んだ後は手を手前に動かす動作がより早く実行されると予測される。しかし，実験の結果から，「あげる」と「くれる」はともに手を奥に動かす動作を促進することが明らかになった。この結果は，どちらの動詞を理解する場合でも主語の視点が取られていることを示唆しており，他者の動作を運動領域でシミュレーションする際には主語の視点が優先される可能性を支持する。