# Language-Driven Motor Simulation is Sensitive to Social Context

**Heeyeon Y. Dennison (yoon@hawaii.edu)**
Department of Linguistics, 569 Moore Hall, 1890 East West Road,
Honolulu, HI 96822 USA

**Benjamin K. Bergen (bkbergen@ucsd.edu)**
Department of Cognitive Science, University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093-0515

## Abstract

People use their bodies differently in different social situations. In Korea, Japan, and Thailand, for example, there are culture-specific conventions for how to transfer small objects. People use one hand to transfer objects to people of equal or lower social status, but two hands with people of higher status. But does individual knowledge of these conventions for how to use one's body extend to other aspects of cognition? For instance, it is known that understanding action language involves internally simulating what it would be like to perform described actions. Do people mentally simulate actions appropriate to the social context described in a sentence? We report on a behavioral experiment, conducted with people born and raised in Korea, that investigated whether cultural practices affect the actions that people represent during language comprehension. We report evidence that motor simulations do indeed reflect social constraints on action.

**Keywords:** sentence processing; mental simulation; motor simulation; embodiment; social cognition; culture.

## Introduction

What are the cognitive processes involved in understanding the meaning of a sentence that you hear or read? Behavioral and brain imaging evidence over the past several decades has revealed that – among other mechanisms – one process that is engaged routinely and mostly unconsciously is mental simulation (Barsalou, 1999). When you hear or read a sentence about an event, you use your visual system to simulate what the mentioned entities would look like: how they would move (Kaschak et al, 2005), what color (Connell, 2007), orientation (Stanfield & Zwaan, 2001, Zwaan et al., 2004) and shape (Zwaan et al., 2002) they would have, and so on. Similarly, when a sentence describes actions, you engage mental simulations of the described actions (Glenberg & Kaschak, 2002; Tettamanti et al., 2005; Buccino et al., 2005; Kaschak & Borreggine, 2008; Bub et al., 2008; Bergen & Wheeler, In press, Bergen et al., In press).

Key evidence that understanding motor language engages motor routines comes from the so-called "Action-sentence Compatibility Effect" (Glenberg & Kaschak, 2002): manual responses to make sentence sensibility judgments are facilitated when the motion of the physical response is compatible with the sentence's implied direction. For instance, *You handed the puppy to Katie* speeds manual responses away from the body, while *Katie handed the puppy to you* speeds movements toward to the body.

Research in this area has shown that implied features of a described scene – even when only implicit – show up in the comprehender's mental simulation. For example, people simulate handshapes specifically afforded by the objects mentioned. For instance, *Mary caught the marble* primes a grasping handshape, while *Mary caught the watermelon* primes an open palm handshape (Bergen & Wheeler, 2005). However, it is unknown at present whether language-driven motor simulation is sensitive to not only physical but also social constraints on action.

Part of acculturation is for people to learn culture-specific prescriptions for motor action (Mauss, 1934). For instance, in Korean culture, people learn to use both hands when giving an object to someone of higher social status, but only one hand with peers or social inferiors. This is obviously a learned behavior – many other cultures around the world do not share this particular convention. What's more, it's a cultural action convention that's specific to social context. In order to transfer an object appropriately to someone, you have to determine what your relative social status is, so as to engage an action using the right number of hands.

In the study described below, we ask whether socially-contingent prescriptions for motor action reach into other aspects of cognition as well, in particular, language comprehension. Does hearing about an action performed in a particular social situation elicit simulation of the prescribed physical behavior that conforms to the cultural constraint? More specifically, do Koreans who hear sentences about object transfer simulate using one hand or two hands, depending on the relative social status of the mentioned recipient? If they do, this would suggest that the motor simulations are flexibly tailored to the social situations comprehenders would encounter in the described situations. In addition, it would suggest that motor simulations have features specific to culture-specific constraints on action.

## Method

The logic of the experiment was relatively simple. Participants listened to sentences in Korean about transferring small objects to recipients who were of either high status or low status. After the end of each sentence, they made a meaningfulness judgment about the sentence, which required them to press buttons either with two hands

or just with their right hand. We predicted that if the participants were automatically engaging motor representations of actions that used either one hand or two hands (appropriate to the social status of the mentioned recipient) then the status of the recipient and the number of hands they had to use to respond should show an interaction in their effect on measured response times.

## Participants

Thirty-two native Korean speakers at the University of Hawai'i participated in this experiment and received $5 in compensation. They all had been born and raised in Korea but moved to the U.S. for their higher education. All but 2 people had lived in Korea for at least 20 years. Their mean length of residency in Korea was 24 years (range: 13.5–35, std: 4.6), while the mean of their age was 27.8 (range: 20–45, std: 6.38). All had normal or corrected-to normal vision and hearing. All but one were right-handed.

## Materials

We constructed twenty pairs of critical sentences in Korean, all of which are meaningful and describe transfer of a small object. The object in all these sentences is conventionally transferred to people of higher status using two hands and to people of equal or lower status using one hand. Examples of high status and low status sentences are in (1) and (2), below.

(1) [High Status]
*Ne-nun cikum kyoswu-nim-kkey phyenci-lul tuli-koisse.*
'You are now (humbly) giving a letter to (your) professor.'

(2) [Low Status]
*Ne-nun cikum tongsayng-hanthey phyenci-lul cwu-koisse.*
'You are now giving a letter to (your) younger sibling.'

The sentences in each pair (like (1) and (2)) differed in two ways. First, while High-status sentences mentioned people of conventionally high status as recipients (such as professors, doctors, lawyers, etc.), low status sentences had recipients who were of lower status or close peers, (like younger siblings, nieces, friends, etc.). Second, the sentences about transfer to people of high status were accompanied by grammatical/lexical markers called *honorifics*, which indicated that the status of the recipient is higher than the status of the subject (*you*). These markers are present on the recipient noun, the dative case marker, and the verb, and as confirmed in a norming study, this is the most natural and proper way to describe an object transfer to a social superior in Korean. (We'll discuss these honorific markers and their implications in more detail in the Discussion section below.) All critical sentences had *you* as the subject, were in the present tense, and used active dative sentence structure.

Beyond the critical stimuli, we created some additional materials. In order to disguise the intent of the experiment, we prepared twenty meaningful filler sentences that were not about transfer of a small object. Because participants were performing a forced choice task, we needed forty non-sensible fillers to balance the twenty critical and twenty meaningful filler sentences. By including these, we ensured that each participant was expected to respond "Yes" half of the time overall. So that participants could not learn to use syntactic properties of the sentences to make judgments, half of the non-meaningful sentences were dative sentences, while the other half included sentences varying in structure and length – just like the meaningful sentences. Also, orthogonally, approximately half of all items mentioned high-status people in somewhere in the sentences, whereas the other half mentioned equal- or lower-status people.

## Design, Procedure, and Predictions

Participants performed a sentence meaningfulness judgment task. They were seated in front of a computer with two keyboards aligned side-by-side in front of them (as in Figure 1 below). They keyboards were oriented such that the long axis of the keyboards projected out directly in front of the participant, at their midline. Participants first pressed the yellow keys with their two thumbs to begin auditory presentation of a sentence – these are at the bottom of the image in Figure 1, and were on the edge of the keyboard closest to the participant. Once they decided whether the sentence made sense or not, they released the yellow buttons to press either the "Yes" or "No" buttons on the keyboard to indicate their decision. The response buttons were positioned to require the participant to use either both hands (the green buttons) or only their right hand (the pink buttons) to press the buttons. Participants were instructed to press the pink buttons with two fingers of their right hand and the green buttons with the index fingers of their two hands. We measured the Reaching Time – the time from the yellow-button release until the "Yes" or "No" button press, because we were interested in seeing if simulation of the socially expected action would influence subsequent motion execution.

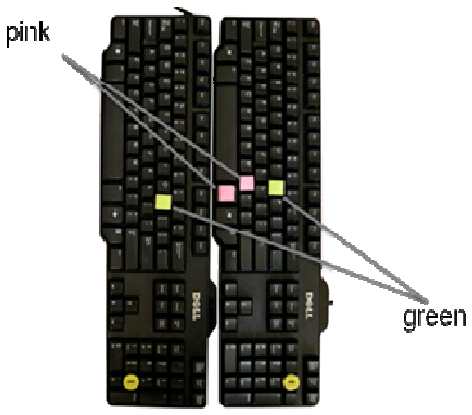


Figure 1. Configuration of keyboards to collect bimanual (green buttons) and unimanual (pink buttons) responses.

Participants pressed down the yellow buttons with the thumbs of their two hands to initiate presentation of a sentence, then pressed either the pink or green buttons to indicate their meaningfulness judgment.

We assigned participants to one of two starting conditions – they began with two-handed responses (the green buttons) meaning either "Meaningful" or "Non-meaningful" (and the pink buttons assigned complementarily). We switched the key assignments for each participant halfway through the experiment. Ten training trials preceded each half. Each participants heard the high-status recipient version of a randomly selected half of the critical sentences (e.g., (1)) and the low-status version of the other half (e.g., (2)). Each session lasted less than 20 minutes.

If native Korean speakers engage motor representations of two-handed actions when processing sentences about object transfer to people of high social status, and one-handed actions when processing sentences about transfer to people of equal or low status, then we should observe an interaction of Hand-Number by Sentence-Type.

But the direction of this interaction effect between language and action is a more difficult matter. Both match advantages (i.e., faster responses in the matching conditions) and mismatch advantages (i.e., slower responses in the matching conditions) have been reported in the literature. A close reading, however, leads us to expect a *mismatch* advantage; that is, two-handed responses will be faster when participants have just heard a sentence about transferring an object to someone of lower status, and conversely, one-hand responses should be faster when the preceding sentence describes transfer of an object to someone of higher status.

We are led to predict this mismatch advantage, rather than a perhaps more intuitive match advantage, for the following reason. Priming effects of action language on motor control are quite sensitive to timing. When there is a delay (more than 500 milliseconds) between the word that denotes the action and the action itself, language about actions facilitates motor actions that have broadly similar characteristics, such as the direction of motion (Glenberg & Kaschak, 2002; Kaschak & Borreggine, 2008; Bergen & Wheeler, In press) or handshape (Bergen & Wheeler, 2005; Bub et al., 2008). However, when the critical action word and the motor action are temporally aligned (within 500 ms), actions that are similar but not exactly the same will in fact be inhibited (Bergen, 2007; Bergen et al., In press, see also the review in Kaschak et al., 2005). That is, there is a *mismatch* advantage when an action verb (e.g., *throw*) immediately precedes activation of a motor routine for a similar but subtly different action (e.g., *push*), as compared with a less similar action (e.g., *kick)* (Bergen et al., In press).

Critically, Korean differs from English in terms of where in the sentence the main verb is placed. In English, the verb in canonically ordered sentences occurs after the subject and before the object. This means that in a sentence about someone acting on something (like *You handed Andy the pizza)*, the verb appears relatively early in the sentence. In Korean, however, the verb occurs at the end of the sentence (as exemplified in (3) below). As a result, when a participant is asked to perform an action immediately after the end of a Korean sentence, it falls within the 500ms window in which we observe mismatch-advantages for similar but non-identical actions. In this experiment, the described actions are indeed different in certain ways from the action of pressing keyboard buttons. Handing someone a business card or a letter, for instance, is different from pressing keyboard buttons in terms of its force, acceleration, palm orientation, handshape, and other motor details. Because participants are asked to process language about an action and then perform a similar but subtly different action very soon thereafter, we expect that Korean sentences should produce a mismatch advantage. This is unlike canonical English sentences, which, due to their verb occurring earlier in the sentence, would be expected to produce *match* advantages, as has been found in other studies investigating language-action interaction effects (e.g., Glenberg & Kaschak, 2002).

(3) Ne-nun Andy-hanthey ku pizza-lul *kenneyesse*.
You-TOPAndy-DAT the pizza-ACC handed.
'You *handed* Andy the pizza.' (English translation)

## Results

Among data from 32 participants and 20 critical items, results from two participants and one item were excluded due to low mean accuracy (more than 3 standard deviations below the participants' and items' overall means of 97% and 96%, respectively). The means for accurate responses from each of 30 participants and 19 items all fell within 3 standard deviations from the overall means for participants and items, respectively, and none were therefore eliminated as outliers.

The dependent measure was Reaching Time, the time it took participants from release of the yellow buttons until press of the pink or green buttons, indicating that the sentence was meaningful. We first eliminated all incorrect responses. Correct responses to critical items were then winsorized using 3 standard deviations from each participant's mean as a cut-off and submitted to the subsequent inferential statistical tests. Mean reaction times in each condition are shown in Figure 2.

We performed two two-way repeated measures ANOVAs (one by participants and one by items) to look at effects of the two independent variables – Sentence-Type with Hand-Number. We found one significant main effect, that was unrelated to our hypothesis; bimanual responses were on average slower than unimanual responses, regardless of the Sentence-Type manipulation: $F_1(1,29)=20.21$, $p<.001$; $F_2(1,18)=27.18$, $p<.001$. More importantly, however, we also observed an interaction between Hand-Number and Sentence-Type that was significant both by participants $F_1(1,29)=7.60$, $p=0.01$, and by items, $F_2(1,18)=6.60$, $p=0.02$. This effect, as seen in Figure 2, seems to have been driven by slower two-handed responses in reaction to sentences

about object transfer to high status recipients than low status recipients, and one-handed responses that were slightly slower after sentences about transferring objects to low status recipients than high status recipients.


Figure 2: Response times as a product of high status and low status sentences, when manual responses were made with one hand or two hands, show a significant interaction between Hand-Number and Sentence-Type

Subsequent pairwise comparisons indicated that this interaction effect was driven mainly by the differences in bimanual responses. That is, the mean reaction time to respond with two hands was significantly slower when the manual action was preceded by high status sentences than low status sentences ($t_1(29)=2.05$, $p=.02$; $t_2(18)=2.11$, $p=.027$). However, looking at the one-handed responses revealed no significant effect of Sentence-Type on Reaching time.

## Discussion

The results reported above indicate that Korean speakers' manual responses to indicate sentence meaningfulness were significantly influenced by the social context mentioned in the sentences they were processing. People were significantly slowed down in making bimanual actions after they heard sentences describing object transfer to someone of higher social status, as compared with sentences about people of equal or lower status. In contrast, unimanual responses were slower with the low status sentences than the high status sentences, although this numerical difference did not approach statistical significance.

This result is yet another piece of evidence in line with previous findings, showing that people engage their motor systems while processing language about interacting with objects and moving their bodies more generally. The study of how exactly we extract meaning from utterances is still in its infancy. And yet, the discovery that modality-specific systems are automatically engaged during the process suggests one part of the puzzle – it could be that motor (and perceptual) simulation plays a functional role in language understanding by allowing the comprehender to construct an experience that is in some ways like what it would be to experience a distal described scene. This simulation may do more than merely create a subjective experience akin to what the comprehender would experience when confronted with the described scene; it might also facilitate inference or be used to generate predictions or interpolate implied but implicit elements of the scene.

However, the findings we've reported here is different from previous work. The current study's findings suggest that the motor activation comprehenders engage during language processing is tailored not only to the objects described but also to the culturally appropriate motor actions that one would perform in the described social context. This suggests that motor simulation isn't merely activated by specific words. Instead, it reflects a computation of socially appropriate action, which must take into account not merely low-level physical properties of a mentioned object, but social variables like age and position. On some accounts, people perform mental simulation to create representations of what it would be like if they were immersed in the described experience (Zwaan, 2004). The finding that people take social variables into account when constructing motor simulations is coherent with this account.

While the finding reported above highlights the importance of social knowledge in language comprehension, the difference we observed between the high and low status sentences could be due to either or both of two differences between them. As we discussed above, the sentences differed both in terms of the social character of mentioned recipients, as well as in the presence or absence of honorific markers. Honorifics—also known as indexical politeness forms—are grammatical and lexical markers in the Korean language that systematically encode the speaker's socio-culturally appropriate regard towards the addressee and/or the referent (Sohn, 1999). In our stimuli, the High Status sentences, as exemplified in the example (1), indicated that the status of the recipient was higher than that of the subject (*you*) by employing honorific markers in three places within the sentence, since that is how native Korean speakers would most naturally express this. The marker *-nim* 'sir/madam' on the recipient noun indicates deference to the high status referent. The marker *-kkey* '(honorable) to' is a case marker used for a high status recipient. The verb root *tuli-*, which is the humble form of the plain verb *cwu-* 'give', is used to indicate deference to the high status recipient. To any native Korean speaker, the presence of these honorific markers clearly indicates that the subject of the sentence (*you*) has a lower status than the mentioned recipient. When these honorifics are dropped, Korean speakers naturally understand that the status of the sentential subject is at least equal or even higher to the mentioned recipient. For instance, dropping honorifics is a common way for Korean students to make fun of their teachers behind their back.

Due to these two kinds of differences, when we compare effects of these two sentence types (1) and (2), we cannot tell if the status of the recipients or the presence/absence of honorifics is responsible for the differences in hand

responses. We believe it is critical to determine whether the social status of the recipients by itself produces the interaction effect, or whether the presence of honorifics is important as well.

To investigate this question, we are currently collecting data for a follow-up study. In that study, we closely matched the current experiment's design and materials, but included a third Sentence-Type condition, one with sentences using high status recipients but no honorific markers. If these sentences behave like the high status sentences in this first experiment, that will suggest that it is the status of the recipient, and not the presence of honorifics, that produces the effect we've observed. But if we find that these new sentences behave like low-status sentences, then this will lead us to conclude that honorifics present in a sentence are a critical factor affecting the motor routines people simulate. This in turn will tell us a little bit about how world knowledge and linguistic cues affect the motor simulations comprehenders construct.

In the Method section, we presented the reasoning why we expected a mismatch-advantage rather than a match-advantage. We found such an effect, and the next step is for us to investigate whether the proposed explanation is correct. We intend to do this through another study, currently under design, in which we change the structure of the critical Korean sentences, so that several words appear at the end of the sentence, after the transfer-action verb. When we thereby increase the interval between the action verb and the subsequent motor response, our explanation would predict that we should find a reversal in the direction of the effect: from a mismatch- to a match-advantage.

The final aspect of the findings reported here that may be of interest is the result of the follow-up pairwise t-tests reported in the Results section. The significant mismatch advantage we observed appears to have been driven more by the bimanual responses than the unimanual responses. We'd like to offer two possible explanations for this fact.

First, in situations of actual transfer, people use the right hand regardless of the status of the recipient; two-handed actions use both hands by definition, and one-handed actions conventionally use the right hand, as left hand transfer is regarded as disrespectful in Korean culture. As a result, both the high-status and low-status sentences should engage motor circuitry involved in the use of the right hand. And the two sentence types might thus interfere to the same extent with subsequent one-hand actions. By contrast, two-handed transfer actions are more similar to – and thus are more interfered with following – sentences about high-status actions, which also use two hands.

A second possible explanation for the larger effect in two-handed responses than one-handed responses is that it could result from differences in the degree to which the two actions are routinized. Perhaps reaching actions to press buttons with one hand are so common as to be routinized to the point where people perform them at floor (i.e., shortest possible reaction times). This would leave little room for effects of previously heard sentences on reaching behavior. In contrast, using both hands is a more difficult task, particularly when participants are trying to maximize both speed and accuracy. It could be that greater difficulty in bimanual action provided room for the critical Sentence-Type factor to more differentially influence reaction times. Indeed, the significantly faster responses in the one-hand condition, as compared with the two-hand condition, are compatible with this floor account. We expect our follow-up studies will further illuminate the extent to which the ACE paradigm can tease apart the action execution resulting from a mental computation of the socially appropriate action, versus an action execution based on a routinized motor command.

## Conclusion

Is language-driven motor simulation sensitive to not only physical but also social constraints on action? The current study suggests that it is. We've presented three new findings. First, the motor simulation that people engage during language comprehension includes the number of hands one would use to perform a described action. Second, cultural-specific rules for motor control enter into these mental simulations. For Korean people, it is customary to tailor the number of hands one uses to transfer objects to the relative social status of the self and the object recipient. This social knowledge about proper actions is engaged in the form of active motor representations of the number of hands appropriate in each social context, even when people merely hear language describing those events. Finally, the direction of action-sentence interactions is different in Korean and English – arguably due to differences in word order. These findings clearly indicate that language understanding is a constructive process that broadly engages heterogeneous cognitive systems – it uses our understanding of physical actions, and even the social conventions surrounding those actions.

These results raise a host of potentially productive follow-up questions. Do people engage socially constrained motor knowledge during other non-motor tasks, such as object perception or recall? Do people raised biculturally, such that they have two distinct systems of motor conventions, simulate different actions depending on the cultural context or the language of an utterance? And is a social constraint on action the type of thing that can be learned late in life – do people introduced to a culture and language as adults also simulate socially appropriate actions? Clearly, there is much more to know about how social constraints on action are learned and what other aspects of cognition they affect.

## Acknowledgments

The authors are grateful to Chae Eun Kim for assisting the stimuli recording, Bodo Winter for helpful suggestions on this work, and to Korean students at the University of Hawai‘i who participated in the experiment.

## References


Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577-609.

Bergen, B., & Wheeler, K. (2005). Sentence Understanding Engages Motor Processes. In *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society.*

Bergen, B. (2007). Experimental methods for simulation semantics. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, and M. J. Spivey (eds.) *Methods in Cognitive Linguistics: Ithaca*. John Benjamins.

Bergen, B., Lau, A., Narayan, S., Stojanovic, D., & Wheeler, K. (In press). Body part representations in verbal semantics. *Memory and Cognition*.

Bergen, B., & Wheeler, K. (In press). Grammatical aspect and mental simulation. *Brain and Language*.

Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition*, *106*, 27-58

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355–363.

Connell, L. (2007). Representing object color in language comprehension. *Cognition*, *102*, 476–485.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding action in language. *Psychonomic Bulletin & Review*, *9(3)*, 558–565.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M. E., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*, B79–B89.

Kaschak, M. P., & Borreggine, K. L. (2008). Temporal dynamics of the action-sentence compatibility effect. *Quarterly Journal of Experimental Psychology*, *61*, 883-895.

Mauss, M. (1934). Les Techniques du corps, *Journal de Psychologie 32*, 3–4. Reprinted in Mauss, Sociologie et anthropologie, 1936, Paris: PUF.

Sohn, H.- M. (1999). *The Korean Language*. New York: NY. Cambridge University Press.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*, 153–156.

Tettamanti, M., Buccino, G., Saccuman, M., Gallese, V., Danna, M., Scifo, P., Fazio, F. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*, 273–281.

Zwaan, R.A., Stanfield, R., & Yaxley, R. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological science, 13(2)*, 168–171.

Zwaan, R. A. (2004). The immersed experiencer: toward an embodied theory of language comprehension. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation*. New York: Academic Press.

Zwaan, R.A., Madden, C.J., Yaxley, R.H., & Aveyard, M.E. (2004). Moving words: Dynamic mental representations in language comprehension. *Cognitive Science*, *28*, 611–619.