# Multimodal Dynamicity in Fictive Expressions: Exploring Co-speech Gestures in Spatial Descriptions

**Hinano Iida**
**(iida.hinano.i1@s.mail.nagoya-u.ac.jp)**
Department of English Linguistics,
Nagoya University

**Takanori Nanahara**
**(nanahara.takanori.r3@s.mail.nagoya-u.ac.jp)**
Department of Environmental Engineering and Architecture, Nagoya University

**Mai Mori (mori.mai.s7@s.mail.nagoya-u.ac.jp)**
Department of English Linguistics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya Aichi 464-8601 Japan

**Abstract**

Both fictive change and motion expressions are linked to dynamic conceptualization, a central concept in cognitive linguistics. However, it remains unclear whether producing these expressions involves the mental simulation of change or a dynamic perception of events—an area that invites further exploration. In this paper, we examine co-speech gestures in a spatial description task, exploring two main predictions: (1) If fictive expressions involve some form of dynamicity or simulation of change or motion, speakers will gesture more frequently than with factive expressions; and (2) If fictive expressions involve dynamicity or simulation, the gestures will reflect this imagery and may be more dynamic than static. The findings from this study support both predictions, suggesting that fictive expressions indeed involve a dynamic conceptualization or simulation of static spatial concepts.

**Keywords:** fictive change; fictive motion; dynamic conceptualization; co-speech gesture; perceptual simulation; spatial language; multimodal analysis; iconicity

## Introduction

### Fictive Expressions and Human Cognition

Certain expressions describe static states as if they were dynamic events in language. The following examples illustrate this phenomenon in both Japanese (1a) (2a) and English (1b) (2b).

(1) a. Sono heya wa maruku natte iru.
The room TOP round become ASP
'The room is round' (Matsumoto, 1996a)
b. A broken line (Langacker, 2008)

(2) a. Sono michi wa massugu hashitte iru
The road TOP straight run ASP
'The road runs straight' (Matsumoto, 1996b)
b. This fence goes from the plateau to the valley.

The expressions in (1) are called "fictive (also known as subjective) change" (Matsumoto, 1996a; Sweetser, 1996; Talmy, 1990, 1996) where the change predicate *maruku natte* 'become round' is used even though there is no change (1a). The examples in (2) are called "fictive (also known as subjective, non-actual, or abstract) motion" (e.g., Blomberg, 2014, 2015; Honda, 1994; Langacker, 1986, 1987; Matlock, 2004, 2010; Matsumoto, 1996bc; Talmy, 1983, 1996, 2000) where the motion verbs *run* or *go* are used even though there is no actual motion.

These expressions have garnered significant attention in cognitive linguistics because they reflect fundamental cognitive abilities or human capacities, or dynamic conceptualization (Langacker, 1997, 2005) in language. For instance, Kunihiro (1985) argues that fictive change and motion expressions are grounded in "resultative perception" —the human cognitive ability to perceive or simulate the current state of an object as the result of a change. Regarding fictive change, Matsumoto (1996a) argues that the cognitive basis of subjective change lies in the human tendency to conceptualize a given situation as a typical or ideal one, based on linguistic data. As for fictive motion, it has been suggested that these expressions involve a subjectively induced process of motion (e.g., Talmy, 1983).

Empirical studies on fictive motion have demonstrated that processing these expressions involves mentally simulated motion (Matlock, 2010; Matlock & Richardson, 2004; Richardson & Matlock, 2007). In their study, participants' eye movements were tracked as they viewed two-dimensional depictions of static spatial scenes while listening to either fictive motion sentences or non-fictive (i.e., actual) motion sentences. The results showed that the type of sentence influenced the participants' eye movements. In addition, Núñez (2006) observed fictive motion expressions in mathematics and found that these expressions tend to co-occur with moving gestures (see also Marghetis & Núñez, 2013; Alcaraz-Carrión, Alibali, & Valenzuela, 2022). However, the question of whether producing fictive change and motion expressions in general involves mentally simulated change or dynamic perception of events remains an area that requires further investigation.

### Gestures Tell Us What the Mind is Doing

Speakers' gestures reflect their experiences, and the way people gesture about an event or object is linked to how they simulate that event or object (Hostetter & Alibali, 2019). In other words, the form of a gesture is shaped by the nature of the underlying mental imagery. For example, speakers often use one hand to gesture when lifting small, light objects, rather than two hands (Beilock & Goldin-Meadow, 2010). Furthermore, when speakers perceive objects as heavier, they



In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* 

tend to gesture as if lifting them at a slower velocity, as though the objects were more difficult to lift (Pouw et al., 2020). Therefore, the way an individual simulates an event is reflected in the corresponding gesture.

There is also significant evidence suggesting that the likelihood of a description being accompanied by gesture is influenced by the extent to which the description evokes thoughts of action or simulation (see Hostetter & Alibali, 2019). For instance, Hostetter, Alibali, and Bartholomew (2011) compared speakers' gesture rates during active mental rotation problem-solving and when describing the final state of the problem. They found that speakers gestured more frequently during the rotation task. In general, speakers tend to gesture more when their imagery is primarily motor-based or spatial, compared to when it is purely visual. In this sense, gestures partly result from how strongly individuals activate simulations while thinking about images.

Building on the evidence, the current paper explores co-speech gestures in a spatial description task, investigating two key predictions: (1) If fictive expressions involve some form of dynamicity or simulation of change or motion, speakers will gesture at a higher rate than with factive expressions; and (2) If fictive expressions involve dynamicity or simulation of change or motion, the gestures will reflect this imagery and may be dynamic rather than static. The results of this study support both predictions, indicating that both types of fictive expressions indeed involve a dynamic interpretation or simulation of static spatial concepts. These findings suggest that producing fictive change and motion involves the mental simulation of dynamicity.

## Methods

### Participants

Sixty-two university students participated in the experiment (male: 44, female: 16, no answer: 2, $M_{age}$ = 23.03, $SD$ = 1.90). They were all native speakers of Japanese. Each participant was paid 2,000 JPY.

### Materials

A total of eleven images were presented: two for practice and nine for the main task. Figure 1 shows the presented images. The images consisted of three types: 2D abstract diagrams, photographs of an architecture and floor plans of an architecture. Two of the 2D abstract diagrams (pr-1 and 2) were used for practice. Each image was printed on a separate sheet of A3 paper. The diagram (pr-1) was used in Kunihiro (1985). The 2D abstract diagrams ab-1 to ab-3 were created by the authors and modeled after Kunihiro's (1985) figures. The architectures in the photographs ph-1 to ph-3 were selected as having simple geometric shapes, i.e., triangles for ph-1, rectangles for ph-2, and a cylinder for ph-3, on their facades. The floor plans pl-1 to 3 were selected or created to have different design characteristics from each other: the pl-1 has isolated rectangular boxes connected by curves, the pl-2 has unified but complicated rectangular geometry, and the architecture of pl-3 has an organic shape.

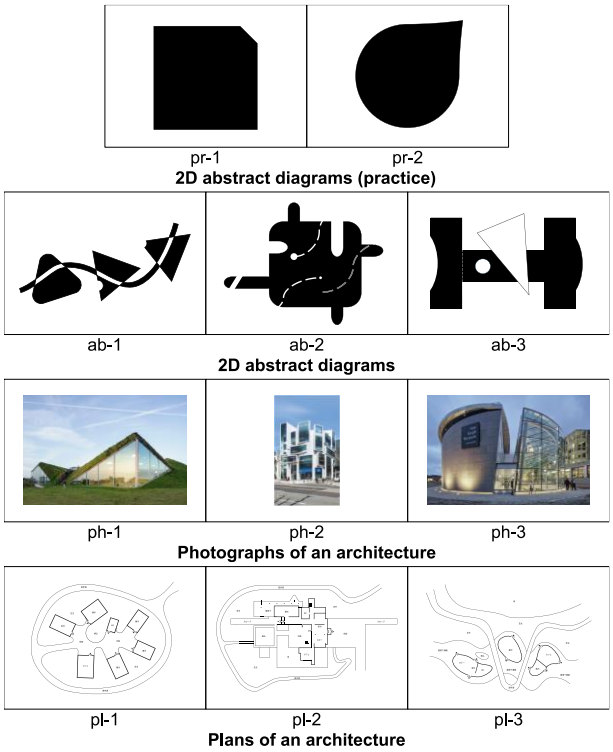


Figure 1: All images used in the present study

All photographs were used with permission from their respective photographers.[1] The pl-1 is a floor plan originally created by the authors. The pl-2 is a floor plan of Sorol Art Museum by Meier Partners, and the pl-3 is a floor plan of Lujiatan Wetland Park Commercial Service Centre by MUDA Architects, partly changed by the authors.

### Instructions

The experimenter gave verbal instructions[2] to the participants. They were asked to freely describe each of the images presented one at a time for two minutes. They were told that they could use gestures and that their descriptions did not need to be understandable to others.

### Procedure

Each participant's experiment was conducted in a small room with an experimenter who was unfamiliar to them.

Prior to the experiment, participants read and signed a consent form. For practice, pr-1 and pr-2 were presented one by one, and participants were asked to describe them. They were allowed to stop describing at any time before reaching the two-minute mark during the practice phase. After the

[1] ph-1 and ph-3 are photographs by ©Ronald Tilleman, and ph-2 is by ©Kyungsub Shin.

[2] The whole instruction is available on our OSF page.

practice tasks, ab-1 to ab-3, ph-1 to ph-3, and pl-1 to pl-3 were presented. The order of presentation was randomized within each type of image across participants, but the sequence of image types remained fixed as follows: 2D abstract diagrams, followed by photographs of architecture, and then architectural plans. This order was chosen to prevent participants from describing elements other than the shape and geometric information of the architecture, such as the presence of people, weather, materials, etc.

Two cameras were used to record the participants during both the practice and the main task. These cameras were fixed on tripods and recorded from the top-right and top-left angles, respectively. Ethics approval for the project was granted by the Graduate School of Environmental Studies' Ethics Committee, Nagoya University (NU_ENV_2024-12).

## Analysis

**Data** Videos were segmented by image for each participant. The start point of each video segment was defined as the moment when the experimenter placed the image on the desk. For the practice phase, the endpoint of the video segments was marked just before the participant signaled the end of their description with phrases such as "that's all", while for the main phase, the endpoint was the moment immediately before the experimenter said "stop".

The average length of the video segments was 41.40 seconds (*SD* = 26.26) for the practice phase and 136.32 seconds (*SD* = 24.09) for the main phase. All speech produced by participants was transcribed, including fillers and rephrasing.

**Extraction of Spatial Descriptions** Participants' utterances describing objects in the images, particularly their spatial components—such as their existence (e.g., *A square is in the center*), quantity (e.g., *Three triangles are on the printed page*), position (e.g., *A curve is on the three triangles*), form (e.g., *the square is rounded*), similes (e.g., *this is like a tree*), and geometric measures (e.g., size, length, height, and width)—were extracted from their entire speech only during the main phase. These utterances were considered as units of subject-predicate constructions (e.g., *this road runs from south to north*) or modifier-NP constructions (e.g., *a rounded triangle*), which are the focus of previous research on fictive expressions in cognitive linguistics, as in (3) and (4) (see also the examples in the introduction section).

(3) Sankakkei ga mittsu narande iru
Triangles SBJ three array ASP
'Three triangles are arrayed'

(4) Eru-ji ni haichis-are- ta tatemono
L-shaped DAT place-PASS PST building
'A building placed in L-shape'

Utterances describing the shape or size of the paper, or color of the figures in the images, were excluded. For the photographs of architecture, descriptions of spatial elements unrelated to the architecture pictured in the center of the images (e.g., a road in front of the building, surrounding buildings, or sky), functions (e.g., museum or art gallery), and materials were also excluded. Additionally, regarding the floor plans, any expressions reflecting the user's perspective were excluded. For instance, an utterance like *I can see the reception to the right immediately after entering the museum from the entrance* was excluded because this type of utterance can be interpreted as describing a mentally imagined, egocentric field of view rather than the printed spatial details of the floor plan. The total number of utterances was 4,921.

**Annotation of Utterances and Gestures** All utterances were annotated by tagging their type of linguistic expression and the co-occurring gestures. The annotation of linguistic expressions was performed by two native Japanese linguists, who did not refer to the videos during the process. Each utterance was classified as "factive", "fictive change", or "fictive motion". The annotation was done based on units of subject-predicate constructions or modifier-NP constructions. When the utterance contained change of state verbs (Matsumoto, 1996a), it was tagged as "fictive change". When the utterance contained motion verbs (Tanaka & Matsumoto, 1997), it was tagged as "fictive motion". For example, the utterance *The walking trail **run**s from the northeast* was tagged as "fictive motion". Below are actual examples of each linguistic type: factive, fictive change, fictive motion.

(5) Sore wa zentaiteki ni maru i
It TOP general COP round NPST
'it is generally round' (factive)

(6) **Bundans-** are ta enchuu
**cut off** PASS PST cylinder
'A cut cylinder' (fictive change)

(7) Tensen ga **hait-**te i ru
dottedline NOM **enter** CONJ be NPST
'There are three dashed lines' (fictive motion)

After the linguistic annotation, the co-occurring hand gestures[3] were annotated by referring to the video recordings, but without consulting the tags for linguistic expressions. This was done by the authors who had not annotated the linguistic expressions. Since this study focused on spatial content, only representational gestures (i.e., iconic and deictic gestures) were tagged (Kita, 2009; McNeill, 1992), while beat gestures were excluded.

Each gesture was classified as either "dynamic" or "static" (Marghetis & Núñez, 2013). While every gesture involves some physical movement of the hands and fingers, certain gestures can be interpreted as "static". Static gestures were

[3] We also observed iconic co-speech head gestures (e.g., de Ruiter, 2000), but they are not investigated here.

defined as those where the hand or finger points to a single point or area without moving from the target, or when the shape is represented with the hands or fingers by forming them at one time and without changing or moving the original shape afterwards. Dynamic gestures, on the other hand, were defined as those that involve tracing shapes with fingers, drawing outlines in the air, or tilting or shifting hands to indicate directions. Figure 2 illustrates examples of each type of gesture. We will discuss the specific examples after presenting the quantitative results in the result section.

When any gesture did not occur with an utterance, it was tagged "none". Additionally, when the participant's hands were not captured by cameras (e.g., when their hands were under the desk, during an utterance), it was tagged "invisible". Since this study focuses on the relation between linguistic expression and co-occurring gestures, the utterances tagged "invisible" were excluded from the analysis.

**Statistical Analysis** All statistical analyses reported in this paper were conducted using R version 4.4.0 (R Core Team, 2024). For the current analysis, we employed generalized linear mixed models, a statistical approach designed for the analysis of binary response variables. These models were constructed using the lme4 package (Bates et al., 2015), to assess the binary response variable (occur or not, dynamic or static), with participants and items included as random effects. The complete dataset and R scripts used for the analysis, excluding audio and video files, are available at: https://osf.io/ekvmy/

## Results

### Result for Prediction 1

**Frequency of Co-speech Gestures by Linguistic Expression Type** Figure 3 shows the rates of utterances with and without co-speech gestures grouped by linguistic expression types. For factive expressions, 926 utterances were accompanied by gestures and 952 were not. For fictive change expression, 1,257 occurred with gestures and 701 did not. The number of fictive motion utterances with and without gestures was 302 and 138, respectively.

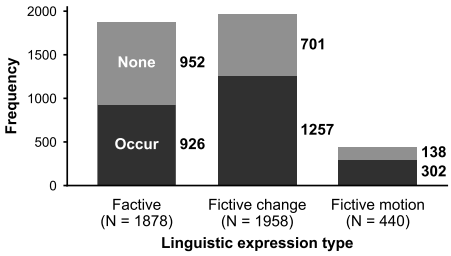


Figure 3: Gesture occurrence frequency by linguistic expression type

We constructed a generalized linear mixed model to predict the occurrence of gestures (whether they occur or not) based on fixed effects: linguistic expression type (factive, fictive motion, or fictive change), along with random effects for stimuli and participants. The model revealed a statistically reliable difference: compared to fictive expressions, co-speech gestures occurred more frequently with fictive motion expressions ($b$ = 0.85, $SE$ = 0.93, $z$ = 9.23, $p$ < .001) and with fictive change expressions ($b$ = 1.13, $SE$ = 0.16, $z$ = 7.19, $p$ < .001). When the reference level was changed to fictive motion expressions, no statistically significant difference was found between fictive motion and fictive change ($b$ = 0.28, $SE$ = 0.16, $z$ = 1.78, $p$ > .05). These results suggest that fictive change and motion expressions, which are assumed to involve dynamic conceptualization, are more frequently accompanied by co-speech gestures, consistent with prediction 1.

### Results for Prediction 2

**Gesture Type by Linguistic Expression Type** Figure 4 compares the frequency of dynamic and static gestures within each linguistic expression type. The total number of utterances with gestures was 2,485, consisting of 926 factive expressions, 1,257 fictive change expressions, and 302 fictive motion expressions. For the factive expressions, 557 were accompanied by dynamic gestures and 369 by static gestures. Of the fictive change expressions, 1,092 were accompanied by dynamic gestures, and 165 were accompanied by static gestures. For the fictive motion expressions, 286 were accompanied by dynamic gestures, while 16 accompanied by static gestures.

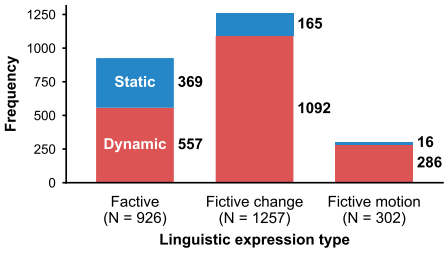


Figure 4: Gesture type frequency by linguistic expression type

We constructed another generalized linear mixed model to predict gesture type (static vs. dynamic) from linguistic expression type, along with random effects for stimuli and participants. The model revealed that the linguistic expressions (fictive change and fictive motion) significantly increased the likelihood of dynamic gesture occurrence ($b$ = 1.62, $SE$ = 0.12, $z$ = 13.3, $p$ < .001; $b$ = 2.75, $SE$ = 0.28, $z$ = 10.0, $p$ < .001). This indicates that dynamic gestures are more likely to occur when these expressions were used. There was also a statistically reliable difference between fictive change and fictive motion ($b$ = 1.13, $SE$ = 0.28, $z$ = 4.06, $p$ < .001). These results indicate that dynamic gestures are more likely to occur when fictive motion expressions were used.

**Examples of Co-speech Gestures** Figure 2 displays actual examples of observed gestures categorized by their type and the associated linguistic expression type. The left column illustrates static gestures, while the right column shows dynamic gestures. Colored lines in the dynamic gestures represent the movement, with red lines marking the starting point and blue lines indicating the endpoint. Black arrows depict the movement paths of the corresponding body parts. Below each illustration, the associated phrases are provided in both Japanese and English (translated). Parts of the utterance corresponding to the gesture are written in bold font. Bold font in the static gestures' column start when the participant shaped their hands and end when they started changing the shape to quit or change gestures. Those in dynamic gestures' column associate with the motion of each gesture (from red to blue). "[ges]" indicates the gesture occurred during a pause.

**Static Gestures** Several different types of gestures can be observed among the static gestures (E.g., 1 to 9), depending on the body parts participants used. First, in E.g. 1, 3, 4 and 9, participants used a single forearm and hand. Second, in E.g. 2, 5, 6 and 8, participants used both forearms and hands to depict the shape of the figure. Third, in E.g. 7, participants used the tip of their index finger to point to the image.

With factive expressions, static gestures were used to describe the shapes or positions of figures in the image. For example, in E.g. 1, the participant used a gesture to iconically indicate the width 'thick' while describing the size and shape of figures such as 'rectangle' or 'square' in E.g. 2 and 3. Additionally, we observed other instances of static gestures with factive expressions that describe positions or spatial relations, such as *the cafeteria is next to the lake*.

With fictive change expressions, static gestures were used to represent the resultative states of change. In E.g. 4, the gesture describes the resultative state or shape 'circle' by forming a rounded corner of the triangle, using the finger or hand, which is similar to E.g. 6 (depicting a walking trail). In E.g. 5, the gesture iconically represents the resultative state (i.e., shape) of *tsunagaru* 'combine' using both hands.

Similarly, with fictive motion expressions, static gestures were used to refer to the end points of motion or the direction

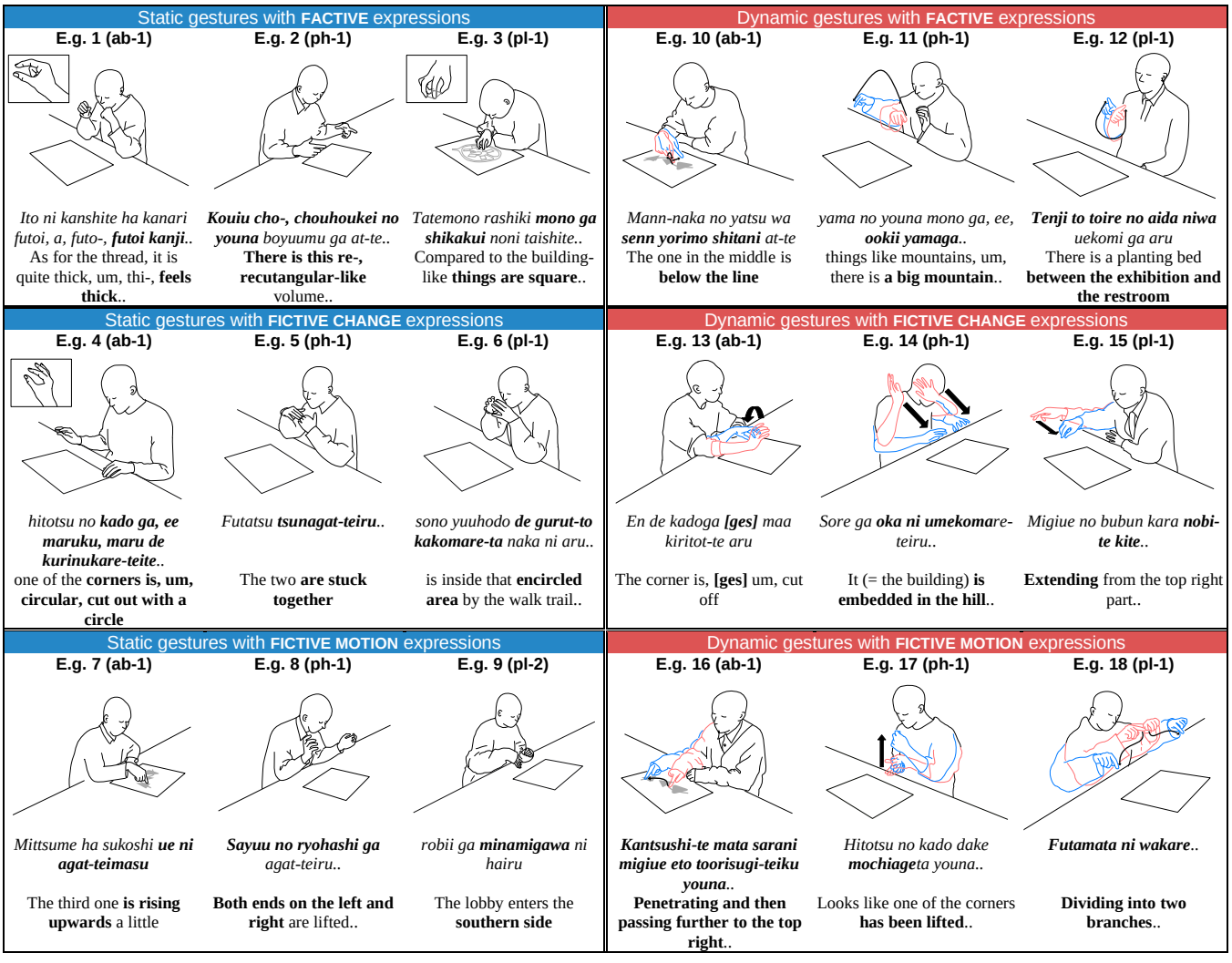


Figure 2: Examples of observed co-speech gestures

of motion. In E.g. 7, 8, and 9, participants indicated the direction of movements or the final position of the referent. E.g. 7 featured a pointing gesture, while the participant in E.g. 9 used all four fingers to indicate the location or direction. The participant in E.g. 8 used both hands to indicate the state of the referent. These gestures co-occurred with terms referring to position or orientation, such as *ue* 'top', *sayuu* 'left and right', and *minami* 'south', and maintained their form until the end of the fictive expressions.

**Dynamic Gestures** Different body parts are also used in dynamic gestures, as shown in E.g. 10 to E.g. 18. First, in E.g. 13 and E.g. 15, participants used a single forearm and hand. Second, in E.g. 14 and E.g. 17, participants used both forearms and hands to depict (fictive) motion or processes related to the figure. Third, in E.g. 16 and E.g. 18, participants used the tip of their index finger to describe (fictive) motion or the shape of the figure. The scale of the movement of gestures also varied, primarily depending on the participants. For instance, E.g. 13 involved smaller gestures compared to E.g. 18.

With factive expressions, dynamic gestures were used to indicate position or direction (as in E.g. 10) and shape (as in E.g. 11 and E.g. 12), similarly to static gestures. Other dynamic gestures, particularly those using the index finger, often depicted or traced the shapes or contours of parts of the image or indicated relative positions. Another example of dynamic gestures with factive expressions involved drawing a small circle repeatedly to point to an area corresponding to the utterances, frequently accompanied by a deictic expression (e.g., *kono* 'this' or *koko* 'here').

With fictive change expressions, dynamic gestures were used to iconically demonstrate the process of (fictive) change or action related to the figure. For example, in E.g. 13, participants made a gesture to depict the change in the shape of the corner of a triangle, which is expressed by the predicate, specifically, cutting off the edge. In E.g. 14, participants described a building and performed a gesture to convey the action of embedding it in the ground, as expressed by the predicate. In E.g. 15, when the participant described the walking trail, the gesture occurred alongside the phrase *nobite-kite* 'extend and come', representing the change. The participant gradually closed their thumb and index finger while moving their hand along the walking trail. Another example can be seen in the gestures for the architecture shown in ph-3 in Figure 1. The building was often described as *enchu wo naname ni kit-ta tatemono* 'a building that a cylinder cut diagonally', and a dynamic gesture accompanied this description, where participants swung their hand down in front of their chest, as if using their hand as a blade.

With fictive motion expressions, dynamic gestures were used to iconically represent the (fictive) motion expressed by language. For example, in E.g. 17, the participant shaped their left hand as if grabbing something and raised it as though they were actually lifting something small. In descriptions of architectural plans, many participants used dynamic, iconic gestures to represent the walking trail surrounding the building blocks, moving their index finger along the trail in the air, similar to the gesture in E.g. 18, which involved fictive motion expressions.

## General Discussion

In this paper, we explored the frequency and dynamicity of gestures in fictive change and motion expressions within a spatial description experiment. Two main findings emerged: (1) fictive expressions were more frequently accompanied by gestures compared to factive expressions; (2) co-speech gestures in fictive change expressions tended to be more dynamic than those in factive expressions, and gestures in fictive motion expressions tended to be more dynamic than those in fictive change expressions. These results suggest that fictive change and motion expressions indeed involve a dynamic conceptualization or simulation of static spatial concepts, as discussed in cognitive linguistics.

Regarding the second finding, which indicates that fictive change expressions tend to be more static than fictive motion expressions, this may support Matsumoto's (1996a) claim that fictive change describes the final stage of a change, with all intermediate stages being irrelevant. We observed examples of static gestures co-occurring with fictive change expressions (e.g., E.g. 4 to 6 in Figure 2), whereas fictive motion typically involves continuous motion of an object.

Future research should extend the current study by addressing several important issues. First, the categorization of gestures into dynamic and static could be further refined to fully capture the dynamic conceptualization in fictive expressions. Although we observed many examples of iconic gestures, the dynamic gesture category also included pointing-like gestures with factive expressions, which may warrant a more detailed distinction. Additionally, we did not consider the potential influence of syntactic structures, such as subject-predicate constructions and modifier-NP constructions (Lavelli & Majorano, 2016), or their positions within discourse (McNeill, 1997), which could affect the frequency of gestures. Furthermore, we could not delve into the individual differences in frequency and dynamicity of co-speech gestures.

Additionally, the current study focused solely on fictive change and motion expressions related to spatial descriptions. However, these expressions are not limited to spatial contexts and can extend to purely visual content, such as color, which is less likely to be gestured (Hostetter & Alibali, 2019). A different approach may be needed to address this limitation. The experimental setting was also restricted, as participants spoke to themselves without a communicative need. It may also be beneficial to control the experimental settings by using minimal pairs of sentences describing the same situation, both with and without fictive expressions. Last but not least, a comprehensive comparison of gestural behavior between actual and fictive expressions, along with cross-linguistic studies (Matsumoto, 1996a), would provide valuable insights into dynamic conceptualization in fictive expressions.

## Acknowledgments

We thank Kimi Akita and the three anonymous reviewers of CogSci 2025 for their thought-provoking comments on earlier versions of this paper. In particular, the suggestions regarding previous research on fictive motion and gesture in mathematics greatly improved the positioning of this study within the existing literature. Any remaining shortcomings are our own. This work was supported by the Graduate Program for Lifestyle Revolution based on Transdisciplinary Mobility Innovation.

## References

Alcaraz-Carrión, D., Alibali, M. W., & Valenzuela, J. (2022). Adding and subtracting by hand: Metaphorical representations of arithmetic in spontaneous co-speech gestures. *Acta Psychologica, 228*, 103624.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Green, P., & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, *12*(1), 2.

Beilock, S. L. & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science*, *21*(11), 1605-1610.

Blomberg, J. (2014). *Motion in language and experience: actual and non-actual motion in Swedish, French and Thai* Doctoral dissertation, The Faculties of Humanities and Theology.

Blomberg, J. (2015). The expression of non-actual motion in Swedish, French and Thai. *Cognitive Linguistics*, *26*(4), 657-696.

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

Honda, A. (1994). From spatial cognition to semantic structure: The role of subjective motion in cognition and language. *English Linguistics*, *11*, 197–219.

Hostetter, A., Alibali, M., & Bartholomew, A. (2011). Gesture during mental rotation. *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 1448-1453). Austin, TX: Cognitive Science Society.

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review*, *26*(3), 721-752.

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, *24*(2), 145-167.

Kunihiro, T. (1985). Ninchi to gengohyogen [Cognition and linguistic expressions]. *Gengo Kenkyu*, *1985*(88), 1-19.

Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive Science*, *10*(1), 1-40.

Langacker, R. W. (1987). *Foundations of cognitive grammar*. Redwood City, CA: Stanford University Press.

Langacker, R. W. (1997). A dynamic account of grammatical function. In J. Bybee, J. Haiman, & S. A. Thompson (Eds.), *Essays on language function and language type dedicated to T. Givon*, Amsterdam and Philadelphia: John Benjamins

Langacker, R. W. (2005). Dynamicity, fictivity, and scanning: the imaginative basis of logic and linguistic meaning. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking*, Cambridge: Cambridge University Press.

Lavelli, M. & Majorano, M. (2016). Spontaneous gesture production and lexical abilities in children with specific language impairment in a naming task. *Journal of Speech, Language, and Hearing Research*, *59*(4), 784-796.

Marghetis, T., & Núñez, R. (2013). The motion behind the symbols: A vital role for dynamism in the conceptualization of limits and continuity in expert mathematics. *Topics in cognitive science*, *5*(2), 299-316.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, *32*(8), 1389-1400.

Matlock, T. (2010). Abstract motion is no longer abstract. *Language and Cognition*, *2*(2), 243-260. Cambridge Core.

Matlock, T. & Richardson, D. C. (2004). Do eye movements go with fictive motion? *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 909-914). Mahwa, NJ: Lawrence Erlbaum Associates.

Matsumoto, Y. (1996a). Subjective-change expressions in japanese and their cognitive and linguistic bases. In G. Fauconnier & E. Sweetser (Eds.), *Spaces, Worlds, and Grammar*. Chicago: University of Chicago Press.

Matsumoto, Y. (1996b). How abstract is subjective motion? A comparison of access path expressions and coverage path expressions. In A. Goldberg, (Ed.), *Conceptual Structure, Discourse, and Language*, Stanford: CSLI Publications.

Matsumoto, Y. (1996c). Subjective motion and the English and Japanese verbs. *Cognitive Linguistics*, *7*, 183-226.

McNeill, D. (1992). Hand and mind. In T. Sebeok & J. Umiker-Sebeok (Eds.), *Advances in visual semiotics: the semiotic web 1992-93*, Berlin, Boston: De Gruyter Mouton.

McNeill, D. (1997). Growth points cross-linguistically. In Nuyts, J. & Pederson, E. (Eds.), *Language and conceptualization. Language, culture and cognition*, New York: Cambridge University Press.

Núñez, R. (2006). Do real numbers really move? Language, thought, and gesture: The embodied cognitive foundations of mathematics. *18 unconventional essays on the nature of mathematics* (pp. 160-181). New York, NY: Springer New York.

Pouw, W., Wassenburg, S. I., Hostetter, A. B., de Koning, B. B., & Paas, F. (2020). Does gesture strengthen sensorimotor knowledge of objects? The case of the size-weight illusion. *Psychological Research*, *84*(4), 966-980.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, *102*(1), 129-138.

Sweetser, E. (1996). Changes in figures and changes in grounds: A note on change predicates, mental spaces and scalar norms. *Cognitive Studies*, *3*(3), 75-86.

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial Orientation: Theory, Research, and Application*, New York: Springer US.

Talmy, L. (1990). Fictive motion and change in language and cognition. *Paper presented at the Conference of the International Pragmatics Association*.

Talmy, L. (1996). Fictive motion in language and "ception". In P. Bloom, M. F. Garrett, L. Nadel, & M. A. Peterson (Eds.), *Language and Space*. Cambridge: MIT Press.

Talmy, L. (2000). *Toward a cognitive semantics: Concept structuring systems*. Cambridge: MIT Press.

Tanaka, S. & Matsumoto, Y. (1997). *Kuukan to idoo no hyoogen* [Expressions of space and motion], Tokyo: Kenkyuusha.