

Self-Split During Mental Simulations

Yusuke Tanaka

General Education Center, Fukuoka University

Abstract: When engaging in mental simulations, language comprehenders are more likely to adopt an internal perspective for sentences using a second-person pronoun and an external perspective for sentences using a third-person pronoun (Brunyé et al. 2009; Sato & Bergen 2013). The present study examines which perspective is adopted when sentences describe language comprehenders in the third-person. In an online sentence–picture verification experiment, participants read a Japanese sentence that used a second-person pronoun (i.e., *you*) or a noun phrase that refers to them (i.e., *the person reading this sentence*) and determined whether or not the picture matched the sentence. They were faster to determine sentences using *you* as matching action pictures (e.g., slicing an apple) taken from an internal perspective. In contrast, they were faster at judging sentences using *the person reading this sentence* as matching external-perspective pictures. These results demonstrated that, during mental simulations, language comprehenders not only subjectify another individual (Horton & Rapp 2003; Pickering et al. 2012) but also objectify themselves, which may be an effort to understand the speaker’s construal of the described situation (Langacker 2008).

Keywords: sentence comprehension, perspective adoption, self-split, mental simulation

1 Introduction

When individuals understand expressions, such as *You are slicing the tomato*, they adopt an internal perspective. Alternatively, they adopt an external perspective when understanding *He is slicing the tomato*. However, which perspective is adopted when sentences describe the language comprehender in the third-person? Resolving this question would help deepen our understanding of the role of the cognitive operation called *self-split* (Ikegami 2008) or *displacement* (Langacker 1985) in language comprehension.

1.1 Perspective-taking during language comprehension

With regard to language comprehension, scholars proposed that people understand language by performing a *mental simulation* of its content (Barsalou 1999; Zwaan 1999). Mental simulation is “the reenactment of

perceptual, motor, and introspective states acquired during experience with the world, body, and mind” (Barsalou 2008: 618). For instance, when individuals understand the meaning of the word *pen*, they reenact the perceptual states acquired while seeing pens or the motor states acquired while using them.

In support of this view of simulation, the literature has demonstrated that language comprehenders mentally simulate the shape of an object implied in a sentence (Hoeben Mannaert et al. 2019; Zwaan et al. 2002). For example, Zwaan et al. (2002) had participants read a sentence such as *The ranger saw the eagle in {the sky/its nest}* and determine whether or not the sentence mentioned a pictured object. They took longer to verify pictures when the shape of the pictured objects was incompatible with the corresponding described objects (e.g., an eagle outstretching its wings after reading that the eagle is in its nest) than when the depicted and implied shapes were compatible (e.g., an eagle folding its wings after reading that the eagle is in its nest). These results were interpreted as a reflection of the mental simulation of the content of sentences; participants were faster to verify a picture when it was similar to their mental simulations. Similar sentence–picture verification experiments have revealed that the color, visibility, size, and orientation of objects are simulated during language processing (Connell 2007; Connell & Lynott 2009; de Koning et al. 2017; Hoeben Mannaert et al. 2017; Stanfield & Zwaan 2001; Yaxley & Zwaan 2007; Zwaan & Pecher 2012).¹

Recent studies have revealed that perspective is also simulated (Brunyé et al. 2009; Pickering et al. 2012; Sato & Bergen 2013). The participants in Brunyé et al. (2009) read a sentence that describes a progressing action, such as *{I am/You are/He is} slicing the tomato*, and viewed a picture taken from an internal perspective (i.e., the perspective of the person performing the action) or an external perspective (i.e., the perspective of the person observing the action performed by another person), as shown in Figure 1.

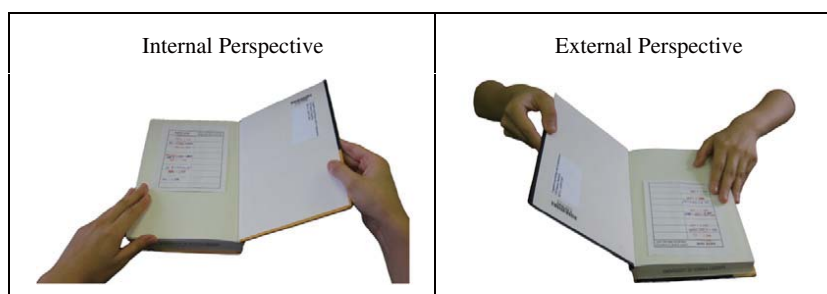


Figure 1. Internal versus external perspective (Sato & Bergen 2013: 364). In the internal-perspective picture, a person is reading a book. In the external-perspective picture, a person is observing someone else reading a book.

They ascertained whether the described action was performed in the picture. The results revealed that when the participants read a sentence using either a first- or second-person pronoun, they were faster to verify

¹ Mental simulation is not limited to the visual modality (Bergen & Wheeler 2005, 2010; Brunyé et al. 2010; Glenberg & Kaschak 2002; Kaschak et al. 2006).

pictures taken from an internal perspective.² By contrast, they verified external-perspective pictures faster when they read a sentence using a third-person pronoun.³ These results demonstrated that language comprehenders adopt an internal perspective when processing a sentence about their own action and an external perspective when processing a sentence about the action of another person.

1.2 Self-split in cognitive linguistics

As previously discussed, scholars demonstrated that language comprehenders adopt different perspectives (internal or external) based on the subject of a sentence when engaging in mental simulation. However, cognitive linguistics assumes that individuals can subjectify another person through *self-projection* and objectify themselves through *self-split* or *displacement* (Ikegami 2008; Langacker 1985). Thus, taking an internal perspective for sentences using a third-person expression and an external perspective for sentences using a first- or second-person expression would be possible.

In fact, studies have found that the perspective of another person can be adopted during mental simulations. Pickering et al. (2012) conducted an experiment in which participants alternated between acting as a speaker and an addressee. When acting as the addressee, they tended to adopt an internal perspective for sentences involving *I* and an external perspective for sentences involving *you*. The participants in Horton & Rapp (2003) read stories where a certain object was blocked or visible from the point of view of the protagonist; they then answered a question about the object. The participants were slowest to respond to the question that asked about occluded information. These results suggested that language comprehenders can adopt the perspective of another person (i.e., the perspective of the speaker or protagonist) during mental simulations.

As demonstrated by these studies, individuals can subjectify another person during mental simulations. However, whether or not they objectify themselves while engaging in mental simulations remains unknown. As previously cited, such self-objectification is termed *self-split* (Ikegami 2008) or *displacement* (Langacker 1985).⁴ This cognitive operation has been used to explain various linguistic phenomena such as the semantic contrast between (1a) and (1b).

- (1) a. The person uttering this sentence doesn't really know.
- b. I don't really know.

(Langacker 1985: 126)

² When a short discourse about the actor was presented prior to sentences using a first-person pronoun, they derived faster responses to external-perspective pictures (Brunyé et al. 2009). With single sentences, therefore, ambiguity with regard to the actor leads language comprehenders to adopt an internal perspective.

³ These results were replicated by Sato & Bergen (2013), who used Japanese sentences, and Pickering et al. (2012), who made a confederate speaker read aloud sentences instead of displaying them on a computer screen.

⁴ Strictly speaking, the definitions of self-split and displacement are not entirely compatible; but this study considers them as an identical cognitive operation that enables people to objectify themselves.

Although (1a) and (1b) present the same conceptual content, their meanings differ due to varying degrees of objectivity with which the speaker is conceptualized: through self-split or displacement, the speaker is conceptualized more objectively in (1a) than in (1b).

Notably, describing oneself in the third-person does not necessarily indicate that the speaker objectively construes themselves. For example, consider the contrast between the English sentence (2a) and its Japanese counterpart (2b). Note that *daddy* in (2a) and *otosan* in (2b) refer to the speakers.

- (2) a. Daddy wants you to sit still.
 b. *Otosan wa Mariko ni jitto suwatte-ite hoshii-nda.*
 father TOP Mariko DAT still sit-STAT want-ASSERT
 ‘Father wants Mariko [= you] to sit still.’

(Hirose 2000: 1635)

Japanese predicates referring to private psychological processes, such as *hoshii*, can co-occur with first-person subject nouns but not with second- or third-person subject nouns (Ikegami 2008). Thus, the fact that (2b) is a natural sentence indicates that the third-person subject noun *otosan* in (2b) functions as a first-person noun. Therefore, speakers of Japanese do not objectively construe themselves even when they refer to themselves using expressions such as *otosan* in contrast to speakers of English.⁵

As previously discussed, referring to oneself in the third-person is seemingly, although not necessarily, based on the self-split or displacement ability. However, it remains unknown whether or not language comprehenders objectify themselves when sentences refer to themselves in the third-person. The experiment conducted in this study intends to address this question. If language comprehenders objectify themselves when reading sentences that describe them in the third-person, then they would be faster to verify external-perspective pictures after reading such sentences.

2 Experiment

The objective of the experiment was to determine whether language comprehenders adopt an internal or external perspective when understanding sentences that describe them in the third-person.

2.1 Participants

Twenty-seven graduate students at Fukuoka University (mean age = 20.44 years) and 40 people recruited through crowdsourcing (mean age = 41.73 years) participated for monetary compensation. Five participants who failed to input their crowdsourcing platform ID in the experiment platform were excluded because their data cannot be linked to them. The study also excluded three participants who failed to complete the experiment, and six participants who were left-handed; thus, analyses are based on the results of 53 participants.

⁵ Ihara (2017) also claims that expressions in Japanese, such as *okaasan* (mom), do not function as the third-person when referring to oneself.

2.2 Materials

The study produced pictures of 48 events taken from the internal and external perspectives for a total of 96 pictures (Figure 2).



Figure 2. Examples of pictures used in the experiment.

In addition, the study formulated 40 Japanese sentences that describe an action such as those in (3a) and (3b). The subject of the sentences was *anata* (you) or *kono bun o yondeiru hito* (the person reading this sentence). Half of the sentences correctly describe corresponding pictures, whereas the other half do not; they describe an action, object, or both incorrectly (e.g., tearing a leaf after reading (3a)). The study also formulated eight ungrammatical sentences such as (3c). They are ungrammatical, because intransitive verbs, such as *nemuru* (sleep), have an accusative object.

- (3) a. *Ima anata wa syokupan o chigitteiru tokoro da.*
Now you TOP bread ACC tearing NMLZ BE
'Now you are tearing a piece of bread.'
- b. *Ima kono bun o yondeiru hito wa ringo o kitteiru tokoro da.*
Now this sentence ACC reading person TOP apple ACC slicing NMLZ BE
'Now the person reading this sentence is slicing the apple.'
- c. *Ima kono bun o nemutteiru hito wa memo o totteiru tokoro da.*
Now this sentence ACC sleeping person TOP note ACC taking NMLZ BE
'Now the person sleeping this sentence is taking notes.'

2.3 Procedure

The experiment was conducted online using PsychoPy (Peirce 2007) and the pavlovia.org platform (<https://pavlovia.org>). After providing consent to the treatment of data obtained during the experiment, the participants watched a video that explained the procedures of the experiment.

Each trial began with a fixation cross for 500 ms followed by the presentation of a sentence for 4,000 ms. After the sentence was removed, another fixation cross was displayed for 500 ms followed by a picture (Figure 3). The participants were asked to decide as quickly and accurately as possible if the pictured event matched the described event by pressing the “m” key, which indicated “a match,” or the “v” key, which indicated “a

mismatch,” regardless of the pictured perspectives. The participants can perform this task by reading only the verb phrases and do not have to read entire sentences. To prevent the participants from using such a strategy, they were instructed to press the “v” key if they found the sentence unnatural regardless of whether the pictured event matched the described event. In this case, the participants did not have to press the key quickly.

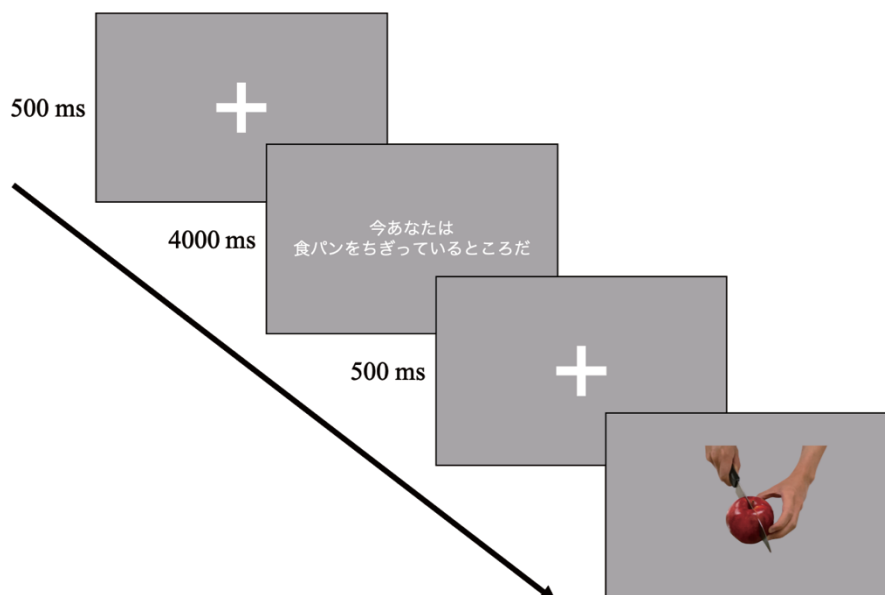


Figure 3. Example of an experimental trial sequence.

Half of the critical and filler items were in the second-person condition, and the other half were in the third-person condition. As a result, the participants encountered an equal number of second- and third-person sentences for the “yes” and “no” trials in the experiment session. For each participant, each sentence appeared in only one of the two perspectives, and the sentences were presented in random order.

3 Results

Three participants were excluded from further analysis because they provided a “yes” response to more than two ungrammatical sentences that correctly describe pictured actions. The first six trials were discarded to reduce the practice effect for each participant. In addition, the study omitted all trials with incorrect responses (4.6%) and with results less than or more than 2.5 standard deviations from the mean of the responses for each participant (2.1%). This step yielded the results presented graphically in Figure 4.

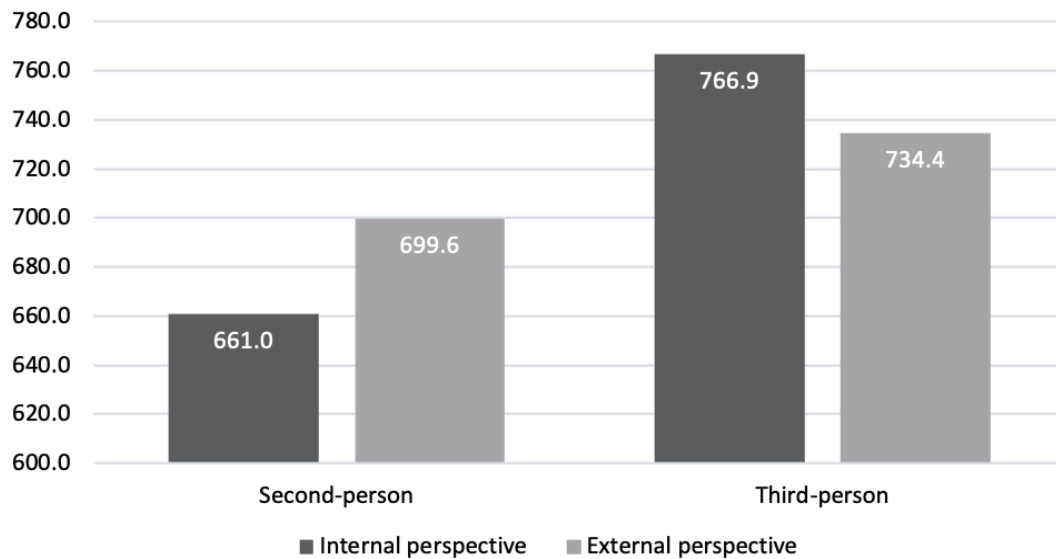


Figure 4. Mean reaction time (in milliseconds) for the verification of pictures that display the main effect of the person and the interaction effect between the person and perspective.

This study used R (R Core Team 2022) and lme4 (Bates et al. 2015) to perform a linear mixed-effects analysis (Baayen et al. 2008). For the fixed effects, the researcher entered the person (second-/third-person) and perspective (internal/external perspective) with the interaction term into the model. In addition, the study adopted the maximal random-effects structure justified by the design (Barr et al. 2013). The study used intercepts for subjects, items, and images, as well as by-subject for the effects of the person, perspective, and their interaction. A by-item random slope for the effect of the perspective and a by-image random slope for the effect of the person were also used. The study did not enter a by-item random slope for the effect of the person and a by-image random slope for the effect of the perspective, because they were unsuitable for this design, that is, each item and image have one and only one person and perspective, respectively. Finally, p -values were obtained using the likelihood ratio tests of the full model with the effect in question against the model without the effect in question.

Table 1 presents the results of the linear mixed-effects analysis. The results revealed that the main effect of the person was significant, whereas that of the perspective was non-significant ($p = .023$ and $p = .756$, respectively). The interaction effect between the two was significant ($p = .026$), which demonstrates that the participants were faster to verify internal-perspective pictures when reading sentences using *you*, while they verified external-perspective pictures faster when reading sentences using *the person reading this sentence*.

Table 1. Results of the linear mixed-effects analysis for reaction times.

	Estimate	Std. Error	t -value	p -value
Person	70.529	30.137	2.340	0.023
Perspective	-4.132	13.543	-0.305	0.756
Person \times Perspective	-66.805	29.163	-2.291	0.026

4 Discussion

The results revealed that language comprehenders were faster to verify pictures when sentences used a second-person pronoun. In addition, they verified internal-perspective pictures and external-perspective pictures faster when sentences used a second-person pronoun and a third-person noun phrase, respectively. This section discusses the reasons that underlie these tendencies.

First, I will discuss the reason for why sentences using a second-person pronoun derived faster responses to internal-perspective pictures, whereas those describing language comprehenders in the third-person enabled them to verify external-perspective pictures more rapidly. These results demonstrate that individuals adopt an internal perspective when comprehending sentences using a second-person pronoun, as presented in previous studies (Brunyé et al. 2009; Pickering et al. 2012; Sato & Bergen 2013). The results also suggest that language comprehenders objectify themselves when sentences describe them in the third-person, which may be due to the effort to understand the speaker's *construal* of the described situation.

In cognitive grammar, linguistic meaning is supposed to consist of not only conceptual content but also a particular manner of construing such a content (Langacker 2008). For example, consider the following sentences: *The glass is half-empty* and *The glass is half-full*. These sentences present the same conceptual content in that they describe the same glass that contains water that occupies only half of its volume. However, their meanings differ, because the glass is construed differently; attention is directed to the volume occupied by the void in *The glass is half-empty* and to the volume occupied by the liquid in *The glass is half-full*. Similarly, sentences beginning with *you* and *the person reading this sentence* differ based not on their conceptual contents but on the construals imposed on the contents or, specifically, the vantage point adopted by the hearer, as diagrammed in Figure 5.

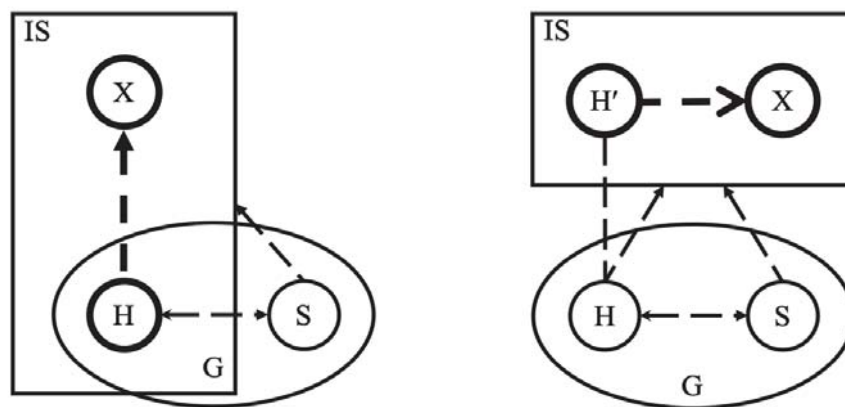


Figure 5. Conceptual structures of *You are V-ing X* and *The person reading this sentence is V-ing X*.

In Figure 5, S is the speaker, H is the hearer, H' is the split hearer, and X is the other participant. The box labeled IS represents the immediate scope, which is characterized as the general locus of attention. The heavy-

line circles and arrow indicate the profiled entities (i.e., things and their relationship designated by the expression). The circle labeled G represents the ground, which includes the speaker and the hearer, the speech event in which they participate, and their immediate circumstances. The dashed arrows indicate that the speaker and hearer apprehend one another in the context of a speech event, as well as directing their attention to the profiled entities in the immediate scope. The line that connects the hearer and the split hearer indicates their referential identity (Langacker 2007, 2008).

In the left diagram, the hearer plays a dual role as a subject and an object of conception, being referred to using a second-person pronoun. In the right diagram, by contrast, the hearer functions solely as a subject of conception, while the split hearer produced by self-split serves as an object of conception, being designated by a third-person expression. This difference results in the different perspective-taking of the hearer, that is, they adopt an internal perspective in the left diagram (i.e., internal to the IS-region) and an external perspective in the right diagram (i.e., external to the IS-region).

Since linguistic meaning involves conceptual content and the construal imposed on that content, the hearer needs to mentally simulate not only what the situation would look like but also the manner in which it is construed by the speaker. For this reason, the hearer adopts an external perspective when processing sentences describing them in the third-person as the speaker does.

Second, I will discuss the reason why the main effect of the person was significant. The simplest possibility is that the participants took longer to engage in mental simulations when the sentences used *the person reading this sentence* because they have more words than sentences including *you*. Another possibility is that self-split delayed the enactment of mental simulation. Scholars suggest that speakers of Japanese prefer event-internal vantage points in contrast to speakers of other languages such as English (Ikegami 2008; Uehara 2006). For example, Compare (4a) and (4b).

- (4) a. I heard the wind.
 b. *Kaze no oto ga kikoeteita wa.*
 wind GEN sound NOM was heard CONF
 ‘The sound of the wind was being heard.’

(Ikegami 2008: 236)

In contrast to (4a), (4b) does not mention the speaker, because they are internal to the described event and cannot perceive themselves. In (4a), by contrast, the split speaker external to the event observes and mentions themselves. Especially for speakers of Japanese, therefore, objectifying themselves requires an additional cognitive load, which delays the enactment of mental simulation. Unfortunately, the results of the experiment cannot pinpoint which possibility is more plausible. If the second one is valid, then the main effect of the person also demonstrates that language comprehenders objectify themselves when processing sentences that describe them in the third-person.

The results of the experiment do not demonstrate that language comprehenders never adopt an internal perspective when sentences describe them in the third-person. The preferred perspective adopted during mental

simulations varies depending on factors such as sense of agency⁶ or empathic engagement (Brunyé et al. 2016; Niikuni et al. 2021). Therefore, certain individuals would adopt an internal perspective even if they are described in the third-person.

5 Conclusion

To assess if language comprehenders adopt an external perspective when sentences describe them in the third-person, the study conducted a sentence–picture verification experiment. The participants were faster to determine sentences using a second-person pronoun as matching action pictures taken from an internal perspective and more rapidly determined sentences that describe them in the third-person as matching external-perspective pictures. These results demonstrated that, during mental simulations, individuals not only subjectify another person but also objectify themselves, probably to understand the speaker’s construal of the described situation.

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–660.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Bergen, B. K., & Wheeler, K. B. (2005). Sentence understanding engages motor processes. In *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society*.
- Bergen, B. K., & Wheeler, K. B. (2010). Grammatical aspect and mental simulation. *Brain and Language*, 112, 150–158.
- Brunyé, T. T., Ditman, T., Giles, G. E., Holmes, A., & Taylor, H. A. (2016). Mentally simulating narrative perspective is not universal or necessary for language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 1592–1605.
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Augustyn, J. S., & Taylor, H. A. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, 20, 27–32.
- Brunyé, T. T., Ditman, T., Mahoney, C. R., Walters, E. K., & Taylor, H. A. (2010). You heard it here first: Readers mentally simulate described sounds. *Acta Psychologica*, 135, 209–215.
- Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, 102, 476–485.

⁶ The feeling of control over actions and their consequences (Moore 2016).

- Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin and Review*, 16, 573–577.
- de Koning, B. B., Wassenburg, S. I., Bos, L. T., & Van der Schoot, M. (2017). Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse Processes*, 54, 493–503.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, 9, 558–565.
- Hirose, Y. (2000). Public and private self as two aspects of the speaker: A contrastive study of Japanese and English. *Journal of Pragmatics*, 32, 1623–1656.
- Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation? *Memory and Cognition*, 45, 974–982.
- Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2019). How are mental simulations updated across sentences? *Memory and Cognition*, 47, 1201–1214.
- Horton, W. S., & Rapp, D. N. (2003). Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension. *Psychonomic Bulletin and Review*, 10, 104–110.
- Ihara, T. (2017). A critical study of the viewing arrangement on subjectivity by R. W. Langacker. *Gakuen*, 918, 100–119. [Written in Japanese]
- Ikegami, Y. (2008). Subjective construal as a ‘fashion of speaking’ in Japanese. In M. Á. Gómez González, J. L. Mackenzie, & E. M. González Álvarez (eds.), *Current Trends in Contrastive Linguistics: Functional and Cognitive Perspectives* (pp. 227–250). Amsterdam/Philadelphia: John Benjamins.
- Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science*, 30, 733–744.
- Langacker, R. W. (1985). Observations and speculations on subjectivity. In J. Haiman (ed.), *Iconicity in Syntax* (pp. 109–150). Amsterdam/Philadelphia: John Benjamins.
- Langacker, R. W. (2007). Constructing the meanings of personal pronouns. In G. Radden, K.-M. Köpcke, T. Berg, & P. Siemund (eds.), *Aspects of Meaning Construction* (pp. 171–187). Amsterdam/Philadelphia: John Benjamins.
- Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, 7, 1272.
- Niikuni, K., Sato, M., & Muramoto, T. (2021). Individual differences in sense of agency and perspective adoption in comprehending Japanese null-subject sentences. *The Japanese Journal of Psychology*, 92, 89–99. [Written in Japanese]
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Pickering, M. J., McLean, J. F., & Gambi, C. (2012). Do addressees adopt the perspective of the speaker? *Acta Psychologica*, 141, 261–269.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

- Sato, M., & Bergen, B. K. (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, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153–156.
- Uehara, S. (2006). Internal state predicates in Japanese: A cognitive approach. In J. Luchjenbroers (ed.), *Cognitive Linguistics Investigations: Across Languages, Fields and Philosophical Boundaries* (pp. 271–291). Amsterdam/Philadelphia: John Benjamins.
- Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, 105, 229–236.
- Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, 28, 81–88.
- Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLOS ONE*, 7, e51382.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13, 168–171.

〈要旨〉

メンタル・シミュレーションにおける自己分裂

田中悠介（福岡大学）

人は言語を、その内容を心内でシミュレーションすることで理解している (Barsalou 1999; Zwaan 1999)。そのようなシミュレーション（メンタル・シミュレーション）の際に取得される視点は主語の人称によって異なり，二人称の場合は「行為者視点」，三人称の場合は「観察者視点」が取得される (Brunyé et al. 2009; Sato & Bergen 2013)。また，他者の視点が取得されることもある (Horton & Rapp 2003; Pickering et al. 2012)。しかし，自己を客観視する視点 (Ikegami 2008; Langacker 1985) が取得される場合があるかについては不明である。これを明らかにするために，Sato & Bergen (2013) と同様の手法を用いた実験をオンラインで実施した。実験参加者は，文とそれに続いて呈示される画像の内容が一致しているかどうかを，可能な限り速く正確に判断するよう指示された。実験の結果，二人称代名詞（あなた）を使った文は行為者視点の画像に対する反応を促進する一方で，実験参加者を三人称で指示する表現（この文を読んでいる人）を使った文は観察者視点の画像に対する反応を促進することが明らかになった。この結果は，メンタル・シミュレーションの際には他者の視点だけでなく，自己を客観視する視点が取得される場合もあることを示唆している。