Who to Blame? The Underlying Representation of Japanese Sentences with Unspecified Agents of Blamable Acts

赤嶺　奨†，大嶺　明李‡, 小波津　豪‡，新国　佳祐§，里　麻奈美‡
Sho Akamine, Akari Omine, Tsuyoshi Kohatsu, Keiyu Niikuni, Manami Sato
†カリフォルニア州立大学フレズノ校，‡沖縄国際大学，§新潟青陵大学
California State University, Fresno, Okinawa International University, Niigata Seiryo University
shoakamine1@gmail.com

Abstract
First-person pronouns elicit the agent's perspective (Brunyé, Ditman, Mahoney, Augustyn, & Taylor, 2009). In some cases, however, sentences contain no explicit agent, as in "unintentional" sentences (e.g., *Kabin-ga wareta* 'The vase broke') and null subject sentences (e.g., *ø Kabin-o watta* 'ø broke the vase'). In fact, Japanese speakers prefer non-agentive expressions, using them much more frequently than English speakers in describing events that equally allow agentive (e.g., 'I dropped the keys') and non-agentive (e.g., 'The keys dropped') descriptions (Choi, 2009; Teramura, 1976). This study examined how native Japanese speakers comprehend and construe the agents of unintentional and intentional events in sentences with unspecified agents of blamable acts. Reaction times on a sentence-picture matching task support that listeners flexibly adopt an agent's or observer's perspective given explicit grammatical pronouns ("I" or "the other") in Japanese. Moreover, the results suggest that Japanese speakers consider another person to be the agent of negative events.

Keywords — Japanese sentence comprehension, event representations, intentionality, perspective-taking

1. Introduction

Patterns in language, even in the local linguistic environment (e.g., a conversation, a paragraph), facilitate the decision of whether we construe someone as being the agent of an event and as the person to pay attention to [1]. Furthermore, linguistic and contextual framing also plays a functional role in the judgement of who is to blame and to what extent they should be blamed and punished. For instance, people attribute more blame and financial liability to the subject after comprehending agentive descriptions (e.g., *Timberlake ripped the costume*) than after comprehending non-agentive descriptions (e.g., *The costume ripped*) [1].

Understanding and producing literal as well as metaphorical language evokes perceptual simulation, which entails the reactivation of perceptual symbols extracted from experience and stored in memory [2][3][4][5][6]. A body of recent work demonstrates that these simulations incorporate an appropriate visual perspective [7][8], even when, in a language where the subject pronouns are frequently omitted when the subject is deducible from the context (e.g., Japanese), the subject pronouns are missing [9]. For example, first-person pronouns (e.g., "I") elicit the agent's perspective, showing that linguistic cues can affect a listener's perspective-taking [8]. In some cases, however, sentences contain no explicit agent, as in "unintentional" sentences (Type 4 below) and null subject sentences (Type 3 below). In fact, Japanese speakers often prefer non-agentive expressions, using them much more frequently than English speakers in describing events that equally allow agentive (e.g., 'I dropped the keys') and non-agentive (e.g., 'The keys dropped') descriptions [10] [11]. However, Japanese speakers seem to be more likely to use agentive expressions (Types 1–3) when describing intentional events [12], suggesting that Japanese speakers pay attention to whether the events are caused intentionally and mark the intentionality of events (intentional vs. accidental) using either agentive or non-agentive expressions. Thus, this study draws on simulation research that has established perspective adoption as an indicator of listeners' construals of agentivity.

2. Experimental Methodology

2.1. Participants

Twenty-nine (22 female and 7 male) right-handed native speakers of Japanese were recruited from Okinawa International University. The mean age of the participants was 20.5 (*SD* = 1.2), ranging from 18 to 23. Each participant received a ¥1000 (approximately $10) gift card for participation.

2.2. Materials

2.2.1. Auditory Materials

The auditory materials comprised four types of Japanese sentences (JS) manipulating the intentionality (intransitive vs.

transitive verbs) and the subject pronoun of sentences with transitive verbs (explicit "I" or "the other" vs. null subjects), as shown below:

Type 1: **Intentional JS with "I"**:
Watashi-ga kabin-o watta 'I broke the vase'
Type 2: **Intentional JS with "the other"**:
Aite-ga kabin-o watta 'The other broke the vase'
Type 3: **Intentional JS with null subject**:
ø Kabin-o watta 'ø broke the vase'
Type 4: **Unintentional JS**:
Kabin-ga wareta 'The vase broke'

The auditory materials were created by recording a 25-year-old female native Japanese speaker reciting each sentence. The target items used 32 pairs of intransitive and transitive verbs (e.g., *wareru* 'to get broken' vs. *waru* 'to break'; *yogoreru* 'to get dirty' vs. *yogosu* 'to dirty'; *nureru* 'to get wet' vs. *nurasu* 'to wet'); all verbs described events that negatively affected an object's status. All participants heard 32 target sentences (four sentences in each of eight conditions); the correct response for all target trials was YES. They also heard 32 filler sentences (three Type 1, three Type 2, two Type 3, and 24 Type 4); the correct response for all filler trials was NO.

Hence, participants encountered equal numbers of unintentional and intentional sentences, as well as approximately equal numbers of the three different types of intentional sentences ("I," "the other," and null subjects). Because the agent responsible for the object's resultative status is explicitly stated in Types 1 and 2, while no agent information is linguistically provided in Types 3 and 4, the experiment assesses which person (i.e., "I" or "the other") Japanese language comprehenders consider to be the most plausible candidate to blame for the described negative event.

2.2.2. Picture Materials

The pictures depicted the two people (i.e., I and the other) sitting across a table and two objects located on the right side and the left side of the screen (Figure 1). The location of the target object was counterbalanced.

Figure 1. A sample of picture materials. The target object *kabin* 'vase' is on the right.

2.3. Procedure

Each participant completed a sentence-picture matching task. They were instructed to place the index finger of their dominant hand (i.e., right hand) on the center of the keyboard (i.e., "J" key) and move the finger to press keys marked either YES (O) or NO (X). The YES was placed on the third key to the right (i.e., ";" key) from the center, and the NO was placed on the third key to the left (i.e., "F" key). In addition, they were instructed to respond as quickly and accurately as possible; we collected their responses on whether the auditory sentence correctly described the pictured event, and analyzed their accuracy and reaction times (RTs). They first saw a fixation cross for 600 ms, followed by an image of an event involving two people and two objects (Figure 1). After viewing the picture for 1000 ms, they heard a simple Japanese sentence and judged whether it correctly described the depicted event by pressing either YES or NO. The picture was presented until participants pressed one of the keys (Figure 2).

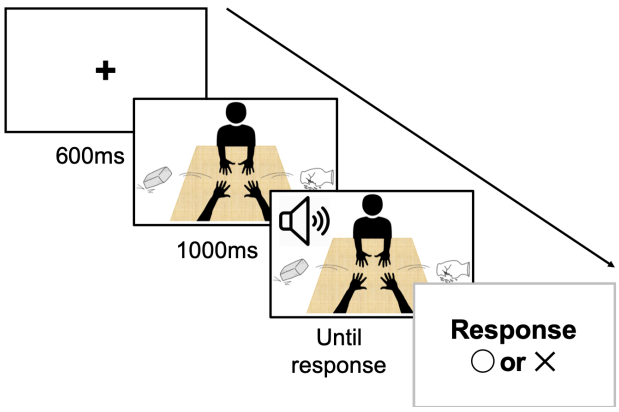

Figure 2. Sequence of events for trials.

2.4. Experimental Design

There were eight conditions: 4 (sentence type) x 2 (location of target object). Because the YES response key was on the right side while the NO key was on the left, counterbalancing the object locations created location-match and location-mismatch conditions. For example, when the agent is explicitly stated as "I" (Type 1), the right-handed participants would be expected to perceive the event by adopting the "I" perspective, and to simulate the event as occurring on the right side of the screen, as in 'I broke the vase' for Figure 1, since they would be more likely to use their dominant hand (i.e., right hand) to cause the event. This is called a location-match condition because the location of the target object and the location of the YES key are both on the right.

In the location-mismatch condition, the target object would appear on the left side with the same sentence (i.e., the location of the target object and the location of the YES key mismatch). When the agent is explicitly stated as "the other" (Type 2), participants would be expected to simulate an event in which the person across the table is the agent, and the target object would be on the left (the expected event location if "the other person" used his/her right hand); thus, Figure 1 and 'the other broke the vase' also create a location-mismatch condition.

2.5. Predictions

We predicted faster reaction times in the location-match condition than in the location-mismatch condition when the agent is clearly stated as in Types 1 and 2. The Types 1 and 2 results provided a baseline for understanding the Types 3 and 4 results, allowing us to access whom the participants construe as the agent of the event in the absence of explicit agent information. If they construe themselves as responsible, and hence simulate the event from the "I"-perspective, RTs should be faster when the target object appears on the right side than when it appears on the left side. Likewise, if they construe "the other person" as responsible for the event, RTs should be faster when the target object appears on the left side.

2.6. Analysis

No subjects or items were excluded. Exceedingly fast or slow responses (i.e., below 250 ms or above 3000 ms), incorrect responses, and responses more than 2.5 *SD* above or below the mean RT for each participant were removed. The rest of the data were modeled using linear mixed-effects model (LME), with random intercepts for participants and items. All the analyses were conducted in R [13] using the *lme4* package [14].

3. Results

No significant difference of accuracy was observed across the eight conditions (4 sentence types x 2 target object locations). Mean RTs for sentence-picture verification demonstrated: (a) a significant main effect of sentence types (Intentional–"I," Intentional–"the other," Intentional–null, and Unintentional) [χ^2 (3) = 32.0, $p = 5.37*10^{-7}$]; (b) a significant main effect of object locations (left vs. right) [χ^2 (1) = 9.3, p = .002]; and (c) a significant interaction effect between sentence types and object locations [χ^2 (3) = 36.4, $p = 6.19*10^{-8}$]. More specifically, treating the Intentional–"I" condition as the reference level, we observed (marginally) significant interactions between (a) the Intentional–"the other" condition and object location [β = 128.8, *SE* = 71.6, *t* = 1.80, p = .073], (b) the Intentional–null condition and object location [β = 163.1, *SE* = 67.7, *t* = 2.41, p = .016], and (c) the Unintentional condition and object location [β = 118.7, *SE* = 67.7, *t* = 1.75, p = .08] (Figure 3; Table 1).

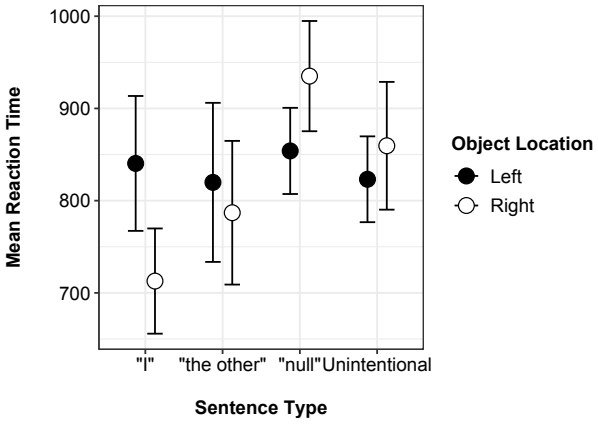

Figure 3. Mean reaction time for the sentence types and object locations. Error bars indicate standard errors.

4. Discussion

This study's findings support the conclusion of previous work [8] that listeners flexibly adopt an agent's or observer's perspective given explicit grammatical pronouns ("I" or "the other") in

Table 1. Reaction Time for Sentence Type and Object Location

Sentence Type	Example	Object Location	Mean (*SD*)	95% CI *LL*	95% CI *UL*
Type 1 Intentional (I)	*Watashi-ga kabin-o watta* 'I broke the vase'	Left	840.37 (380.0)	690.04	990.70
		Right	712.81 (301.8)	595.77	829.85
Type 2 Intentional (the other)	*Aite-ga kabin-o watta* 'The other broke the vase'	Left	819.83 (431.2)	641.84	997.82
		Right	786.91 (397.2)	626.48	947.35
Type 3 Intentional (null subject)	*ø Kabin-o watta* 'ø broke the vase'	Left	853.93 (247.2)	758.06	949.80
		Right	935.04 (316.3)	812.41	1057.67
Type 4 Unintentional	*Kabin-ga wareta* 'The vase broke'	Left	823.19 (246.2)	727.72	918.66
		Right	859.51 (366.5)	717.38	1001.64

Japanese. More interestingly, the study reports a new finding: Japanese speakers preferred the observer's perspective when the agent was not explicitly mentioned (i.e., unintentional events and intentional events with null subjects). This result suggests that Japanese speakers construe the underlying agent of events causing negative results for an object's status as another person rather than as themselves, and that they cognitively represent such described events from the observer's perspective. Although Japanese speakers seem to prefer not to express the agent of negative events, choosing non-agentive sentences [10] and showing attenuation of attention to agents in their non-agentive linguistic descriptions of events causing negative change of status to objects [15], they in fact assume that the person responsible for such events is the other person, not themselves, even when both are equally plausible candidates.

Acknowledgements

This research was supported by a Grant-in-Aid for Scientific Research (B) (PI: Manami Sato, # 19H01263).

References

[1] Fausey, C. M., & Boroditsky, L., (2010) "Subtle linguistic cues influence perceived blame and financial liability", Psychonomic Bulletin & Review, Vol. 17, No. 5, pp. 644-650
[2] Barsalou, L. W., (1999) "Perceptual symbol systems", Behavioral and Brain Sciences, Vol. 22, pp. 577-660
[3] Barsalou, L. W., (2008) "Grounded cognition", Annual Review of Psychology, Vol. 59, pp. 617-645
[4] Bergen, B. K., (2015) "Embodiment", In Dabrowska, E., & Divjak, D., (Eds.) "Handbook of Cognitive Linguistics", Walter de Gruyter GmbH, pp. 10-23
[5] Wilson, N. L., & Gibbs, R. W., Jr., (2007) "Real and imagined body movement primes metaphor comprehension", Cognitive Science, Vol. 31, No. 4, pp. 721-731
[6] Sato, M., Schafer, A., & Bergen, B. K., (2015) "Metaphor priming in sentence production: Concrete pictures affect abstract language production", Acta Psychologica, Vol. 156, pp. 136-142
[7] Borghi, A. M., Glenberg, A. M., & Kaschak, M. P., (2004) "Putting words in perspective", Memory and Cognition, Vol. 32, No. 6, pp. 863-873
[8] 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, Vol. 20 No. 1, pp. 27-32
[9] 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, Vol. 127, No. 3, pp. 361-374
[10] Choi, S., (2009) "Typological differences in syntactic expressions of path and causation", In Gathercole, V., (Ed.) "Routes to language: Studies in honor of Melissa Bowerman", Psychology Press , pp. 169-194
[11] Teramura, H., (1976) "The syntax and semantics of the noun modification part 2", In "Nihongo Nihonbunka, No.5", Osaka University of Foreign Studies
[12] Fausey, C. M., Long, B. L., & Boroditsky, L., (2009) "The role of language in eye-witness memory: Remembering who did it in English and Japanese" Proceedings of the Annual Meeting of the Cognitive Science Society, Vol. 31, pp. 2426-2431
[13] R Core Team., (2020) "R: A language and environment for statistical computing", R Foundation for Statistical Computing
[14] Bates, D., Maechler, M., Bolker, B., & Walker, S., (2015) "Fitting linear mixed-effects models using lme4", Journal of Statistical Software, Vol. 67, No. 1, pp. 1-48
[15] Fausey, C. M., Long, B. L., Inamori, A., & Boroditsky, L., (2010) "Constructing agency: The role of language", Frontiers in Psychology, Vol. 1