Title

Do people use gestures differently to disambiguate the meanings ofJapanese compounds?

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors

Kashiwadate, Kei
Yasuda, Tetsuya
Kobayashi, Harumi

Publication Date

2019

Peer reviewed

Do people use gestures differently to disambiguate the meanings of Japanese compounds?

Kei Kashiwadate (18rmd06@ms.dendai.ac.jp)
Graduate School of Tokyo Denki University
Ishizaka, Hatoyama-machi, Hiki-gun, Saitama 350-0394
Tetsuya Yasuda (t-yasuda@mail.dendai.ac.jp)
Tokyo Denki University
Harumi Kobayashi* (h-koba@mail.dendai.ac.jp)
Graduate School of Tokyo Denki University

Abstract

Spoken language often includes ambiguity in meaning. Compounds such as "green teacup" can be interpreted with two different meanings: "green colored teacup" and "cup for green tea." We can assume there are two different underlying syntactic structures. Phonetic aspects have been studied in the disambiguation process of such ambiguous phrases, but the roles of nonlinguistic information such as gestures have not been explored yet. We investigated whether people use gestures differently when they were asked to describe the meanings of Japanese compounds that can be interpreted as two different meanings. We found that the timing of gestures in relation to the target words of accompanying speech was different between right branching compounds and left branching compounds. Gestures seem to be used to suggest upcoming two words (adjective and noun) as a unit in branching. Gestures can be a useful means to disambiguate the meanings of compounds.

Keywords: Gestures; Disambiguation; Branching; Compounds

Introduction

A phrase consists of concatenation of words that are produced sequentially. It is known that compounds can be interpreted to have multiple meanings. For example, "green teacup" can be interpreted either as a green-colored teacup or as a cup of green tea. A phrase structure with the meaning "teacup for green tea" can be classified as left branching (LB); that is, "green" and "tea" are first grouped to "green tea" and then together play an adjective role in "cup." The phrase structure with the meaning of "green-colored teacup" can be classified as right branching (RB); that is, "tea" and "cup" are first grouped together, and the word "green" plays an adjective role in "teacup" (Figure 1). Because speech is produced sequentially, the surface structure does not have enough information to show the underlying syntactic structure. Therefore, phrases inevitably have ambiguity in meaning. Nevertheless, people usually seem to have little difficulty in discerning the meanings of such phrases. Humans may use some disambiguation cues to resolve ambiguities in such ambiguous structures.

Previous studies have focused on prosodic cues as a means of disambiguation (Ito, Arai, & Hirose, 2015; Hirose & Mazuka, 2015; Venditti, 1994). Native Japanese speakers prefer LB interpretation over RB interpretation for slightly simpler Japanese compound constructions and to make RB interpretation more accessible. A clear prosodic demarcation that raises the pitch range of the second word has been found effective (Ito, et al, 2015; Hirose & Mazuka, 2015; Venditti, 1994). However, the exact disambiguation cues are still unknown. In the present study, we focused on nonverbal cues, in particular, gestures that have not been examined yet in the disambiguation mechanism of syntactic structures.

Figure 1: Two different syntactic structures, left branching (left) and right branching (right), in the compound "green teacup"

Gestures play an important role in communication. Humans simultaneously use gestures and language to convey information to others. Gestures are usually produced slightly earlier than associated speech, and this can make the hearer anticipate the information in the upcoming speech (MacNeill, 1987). Iconic gestures (e.g., depicting objects by movement trajectories) and pointing gestures can reflect aspects of the speaker's nonlinguistic spatial representations (Majit, Bowerman, Kita, Haun & Levinson, 2004). Gestures can spontaneously accompany speech and make communication smooth (Kita & Saito, 2002). Representational gestures (i.e., iconic and deictic gestures) can express spatial contents or metaphorically express temporal concepts (Kita, 2009). Additionally, gestures express information even when it is difficult to express in language (Alibali, Evans, Hostetter, Ryan & Mainela-Arnold, 2009). Various functions are known about gestures, but the topic of whether gestures can contribute disambiguation mechanisms of syntactic structures has been largely unexplored.

Previous studies on interpretation of compounds of possibly different branching structures have showed that people prefer a certain branching over other branching when two (or more than two) different branchings are possible (e.g., Ito et al., 2015). In our study (accepted) on Japanese participants' interpretation of Adjective_1 + Noun_1 + Adjective_2 + Noun_2 compounds, we found that some adjectives are interpreted more dominantly than other adjectives for certain nouns. For example, "long" can be a

typical adjective for "tail," but an atypical one for "cat." It is possible that "long cat" may mean that the cat's body is long, but this sounds somewhat strange. The typicality of the adjective + noun combination may affect the predominant interpretation.

The present study investigated whether gestures are used as a clue to resolve ambiguities in branching structures of Japanese compounds. We examined whether the productions of participants' gestures differ in the case of compounds of either LB or RB. Our prediction was that participants might make gestures with different timings when verbally producing the compounds of LB or RB.

We also examined whether people may exaggerate their gestures by taking more time for relevant gestures when they are aware that more than one interpretation is possible for ambiguous compounds. To examine this exaggeration aspect, we compared a one-picture condition (Alone condition) that denoted either LB or RB meaning and a two-picture condition (side-by-side condition) that denoted both LB and RB meanings side by side so that people could more easily notice the different interpretations.

Further, we also examined another source for possible exaggeration, the combination of nouns and adjectives. We decided to compare the two adjectives "big" and "long." The adjective "big" can be typically applied to "cat" (big cat) or "tail" (big tail), whereas the adjective "long" can be typically applied to only "tail" (long tail) and not "cat" (??long cat). We expected the participants to feel less ambiguity when they interpreted phrases with "long" rather than phrases with "big," so the participants would take less time for "long" condition and the timing of the gestures may also be different between the "big" and "long" conditions.

Method

Participants

Sixteen Japanese monolingual students who spoke Japanese as a first language participated (M age = 21.6, SD = 1.32; 1 female). This study was approved by the ethics committee of the participants' university.

Stimuli

A total of 32 slides were prepared using Adobe Illustrator. Sixteen slides were prepared for the side-by-side picture condition, and another 16 slides were prepared for the alone picture condition. In the side-by-side condition, two comparable objects were drawn side by side in each slide (Figure 2). One was the object (animal) according to LB interpretation, and the other was the object (animal) according to RB interpretation. The slide in the side-by-side condition consisted of one target phrase on the top, two illustrations (i.e., LB and RB interpretations) in the middle, and explanatory notes for each illustration on the bottom (Figure 3). A compound had two possible interpretations: an LB interpretation and RB interpretation. For example, [Kuroi] [Sippono] [Ookina] [Neko] in Japanese (i.e., [Black] [Tailed] [Big] [Cat]) can be interpreted either as "a big cat with black tail" (LB) or as "a black cat with a big tail" (RB)[1]. The difference of meaning can be explained as follows: in the case of LB, the "tailed" branch connects to "black" branch. In the case of RB, the "tailed" branch connects to the "big cat" branch (Figure 4). The position of the LB object and RB object in each slide was counterbalanced. In the alone condition, there was only one object of either LB interpretation or RB interpretation on each slide. There were eight side-by-side slides and eight alone slides. The slide in the alone condition consisted of one target phrase on the top, one illustration in the middle, and an explanatory note for the illustration on the bottom.

In the stimulus compounds, 16 slides included the adjective "big," and another 16 slides included the adjective "long."

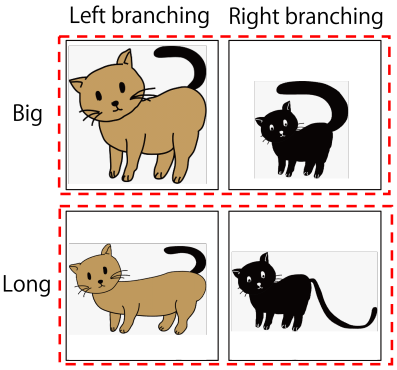

Figure 2: An example of object sets

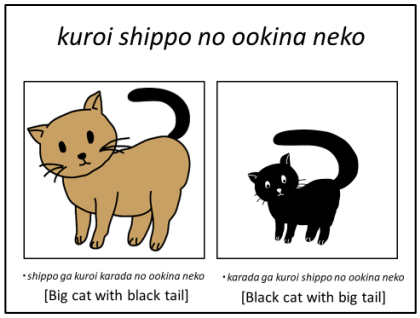

Figure 3: An example of side-by-side slide

1

Japanese	|*kuroi*	|*shippo*	|*no*	|*ookina*	|*neko*
Word class	|[adjective]	|[noun]	|[particle]	|[adjective]	|[noun]
English	|black	|tail		| big	|cat

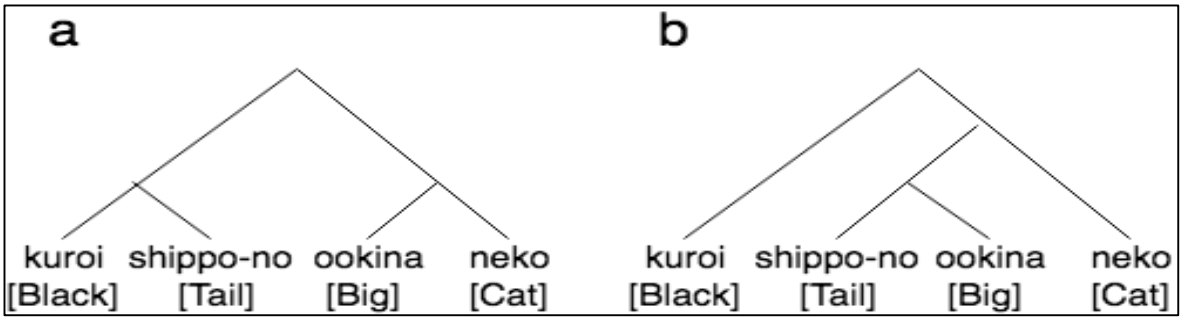

Figure 4: Branching structure. (a): LB interpretation; (b): RB interpretation

Procedure

The participants were divided into two experimental groups: alone slide group and side-by-side slide group. Each group looked at eight slides on a computer monitor.

After filling in the consent form, the participants were seated in front of a monitor (Figure 5).

The participants took part in one practice trial to be familiarized with the task, and then the experiment was started. On each trial, a fixation cross appeared in the center of the monitor. After the cross was fixated for one second, a slide appeared for 10 seconds. Then, only the top phrase was displayed. At that moment, participants were asked to make gestures to describe the presented picture while verbally producing the phrase (Figure 6).

Participants' gestures and utterances participants were recorded by a video recorder (Microsoft LifeCam). In the side-by-side slides, one of the two objects was presented with a surrounding red frame, and the participants were asked to describe the indicated object.

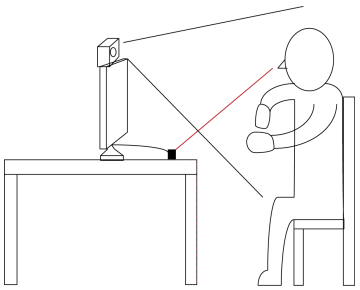

Figure 5: Experimental layout

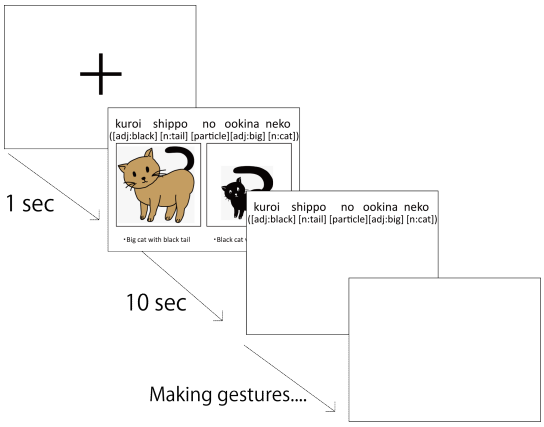

Figure 6: Flowchart of slides

Coding

We annotated utterances and gestures using ELAN 2017 (Version 5.1). The timing and duration of each gestures were recorded (Figure 7). We used the coding scheme modified version of Kita, Gijn, and Hulst's (2014) gesture coding. In this scheme, a gesture consists of a preparatory movement, followed by a stroke, and then finally, a finishing movement. We recorded the time of onset and end of each gesture stroke to determine the timing and duration of each gesture. The third word was the critical adjective "big" or "long" that was grouped with either the second word (e.g., "tail") or the final word (e.g., "cat").

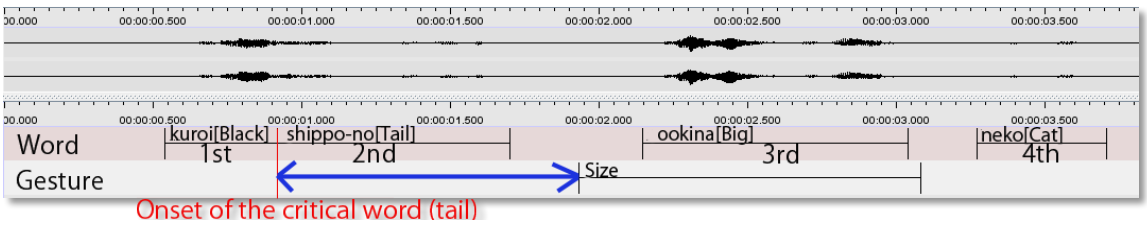

Figure 7: An example of the annotation using ELAN

Results

Timing of the first gesture

To find out whether the onset of gestures differed between LB and RB, between the alone slide and side-by-side slide, and between the adjectives "big" and "long," the difference in time (seconds) between the onset of the critical word "shippo (tail)" on the branching point and the onset of the first gesture in each slide in the participant's performance was measured and taken as a dependent measure. A 2 (Slide: Alone, Side-by-side) × 2 (Adjective: Big, Long) × 2 (Branching: LB, RB) three-way ANOVA was performed on the measure. Slide (Alone, Side-by-side) was the between-participants variable. Adjective (Big, Long) and branching (LB, RB) were the within-participants variables.

There was a marginally significant effect of branching ($F(1,14) = 3.6925$, $p = .075$, $\eta^2 = 0.0956$). This meant that the onset of the first gesture was earlier in RB (M= -0.55) than in LB (M= -0.07). Furthermore, there was a significant Slide × Adjective × Branching interaction ($F(1,14) = 6.3258$, $p < .05$). To explore the significant Slide × Adjective × Branching interaction, the simple interaction effects of Slide, Adjective, and Branching within each condition were calculated (Figure 8).

The simple main effect of branching in the "long" condition was marginally significant ($F(1,14) = 3.1574$, p = .097, $\eta^2 = 0.0842$). This meant that when the third word was "long," the onset of the first gesture tended to be earlier in RB (M= -0.49) than in LB (M= -0.08).

There was a simple Slide × Branching interaction in the "long condition" ($F(1,14) = 4.9359$, $p < .01$). Simple-simple main effects of Slide and Branching within the "long" condition were calculated. There were simple-simple main effects of branching ($F(1,7) = 7.0215$, $p < .05$) and slide ($F(1,14) = 4.4872$, $p = .052$, $\eta^2 = 0.2427$). The simple-simple main effect of slide was marginally significant. It meant that when the third word was "long" and the slide was "alone," the onset of the first gesture was earlier in RB (M= -0.88) than in LB (M= 0.04). When the third word was "long" and branching was RB, the onset of the first gesture was more behind when slide was side by side (M= -0.11) than when slide was alone (M= -0.88).

Total duration of the gestures

Using the video recordings, we calculated the total duration of gestures (seconds) produced in each slide. We predicted that the duration of gestures was different between LB and RB.

A 2 (Slide: Alone, Side-by-side) × 2 (Adjective: Big, Long) × 2(Branching: LB, RB) three-way ANOVA was performed on the total duration of gestures.

There was a marginally significant effect of slide ($F(1,14) = 3.935$, $p = .067$, $\eta^2 = 0.0512$). The total duration time of gestures was longer when the slide was alone (M=2.16) than when the slide was side by side (M=1.82). There was also a significant Slide ×Adjective × Branching interaction ($F(1,14) = 15.9588$, $p < .01$). To explore the significant Slide × Adjective × Branching interaction, simple interaction effects of slide, adjective, and branching within each condition were calculated.

There was a simple main effect of Slide in the "big" condition ($F(1,14) = 5.6515$, $p < .05$) and adjective in the alone condition ($F(1,7) = 4.8931$, $p = .062$, $\eta^2 = 0.0171$). The simple main effect of the adjective in the alone condition was marginally significant. When the third word was "big," the total duration time of gestures was longer when the slide was alone (M= 2.27) than when the slide was side-by-side (M= 1.70). When the slide was alone, the total duration time of gestures was longer when the third word was "big" (M=2.27) than when the third word was "long" (M=2.06).

There were simple Adjective × Branching interactions in the alone condition ($F(1,7) = 6.9771$, $p < .05$, $\eta^2 = 0.0368$) and in the side-by-side condition ($F(1,7) = 9.1492$, $p < .05$, $\eta^2 = 0.1196$). The simple-simple main effects of adjective and branching within the alone condition and side-by-side condition were calculated. There were simple-simple main effects of the branching in the alone condition ($F(1,7) = 33.2153$, $p < .001$, $\eta^2 = 0.1386$) and in the side-by-side condition ($F(1,7) = 7.6180$, $p < .05$, $\eta^2 = 0.2495$). When the slide was alone and branching was LB, the total duration time of gestures was longer when the third word was "big" (M= 2.39) than when the third word was "long" (M= 1.89). When the slide was side by side and the third word was "long," the total duration time of the gestures was longer in LB (M= 2.32) than in RB (M= 1.55).

There was a simple Slide × Branching interaction in the "long" condition ($F(1,14) = 5.0550$, $p < .05$). The simple-simple main effect of slide and branching within the "long" condition was calculated. There was a simple-simple main effect of branching that was marginally significant ($F(1,14) = 4.0335$, $p < .05$). When the third word was "long," and branching was RB, the total duration time of the gestures was longer when the slide was alone (M= 2.23) than when the slide was side by side (M= 1.55).

There was a simple Slide × Adjective interaction in the LB condition ($F(1,14) = 7.8598$, $p < .05$). The simple-simple main effect of slide and adjective within the LB condition was calculated. There was a simple-simple main effect of slide ($F(1,14) = 5.9902$, $p < .05$). When branching was LB and the third word was "big," the total duration time of the gestures was longer when the slide was alone (M= 2.39) than when the slide was side by side (M = 1.60).

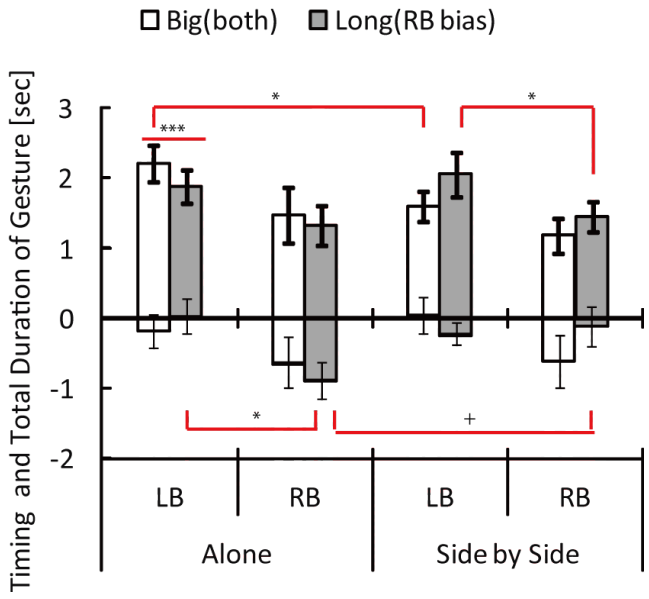

Figure 8: The start time point, the end time point, and the duration of the gestures in each condition. Zero denotes the starting time point of the target word "shippo (tail)" in accompanying speech. The thin error bars denote standard errors in the onset of gesture. The thick error bars denote standard errors in the duration of gesture.

Discussion

The most important finding was that the onset of the first gesture in relation to the critical word (the second word "tail") was earlier in the RB condition than in the alone slide. This means that the participants started a gesture earlier when they wanted to mean "a black cat with a big tail" (RB) than they wanted to mean "a big cat with a black tail" (LB). We interpreted this result as follows: in RB, the meaning "the tail is big" is important to convey, so participants started the gesture earlier to easily bring "tail" and "big" together as words that belonged to the same branch. Usually, the participants' gesture in this RB slide involved extending and moving their arm horizontally to describe a long big tail. Thus, the participants seemed to emphasize "big tail" by this gesture. In contrast, the gesture used in LB slide involved moving both hands widely up and down to describe the shape of the big cat. Thus, the participants seemed to emphasize "big cat." The point is that the critical adjective "big" must be grouped with either "tail" (RB) or "cat" (LB). As for the duration data, we found that the duration of the gestures was shorter in the side-by-side slides than in the alone slides. The participants might have thought that certain gestures would be enough for disambiguation when an alternative picture was explicitly presented. In the side-by-side/long/RB condition, participants might not have thought that disambiguation was necessary because "long cat" (LB interpretation) sounded too atypical compared with "long tail" (RB interpretation).

In conclusion, using Japanese compounds, we found that the timing and duration of gestures in relation to the target words in accompanying speech were different between RB compounds and LB compounds. Gestures seemed to be used to suggest two upcoming words in speech (adjective and noun) as a unit in branching. Gestures can be a useful means to disambiguate the meaning of compounds.

References

Alibali, M., Evans, J., Hostetter, A., Ryan, K. & Mainela-Arnold, E. (2009). Gesture – speech integration in narrative: Are children less redundant than adults? Gesture 9 (3), 290 – 311. doi: 10.1075/gest.9.3.02ali

Hirose, Y. & Mazuka, R. (2015). Predictive processing of novel compounds: Evidence from Japanese. Cognition, 136, 350–358. doi: 10.1016/j.cognition.2014.11.033

Ito, K., Arai, M., & Hirose, Y. (2015). The interpretation of phrase-medial prosodic prominence in Japanese: Is it sensitive to visual and discourse context? Language, Cognition and Neuroscience , 30, 167-196. doi: 10.1080/01690965.2013.864778

Kita, S. (2002). Jesuchaa: Kangaeru karada [Gesture: The body that thinks]. Tokyo: Kaneko Shobo.

Kita, S. (2009). Cross-cultural variation of speech accompanying gesture: A review. Language and Cognitive Processes, 24 (2), 145167. doi: 10.1080/01690960802586188

Kita, S., & Sito H. (2002). Jesuchaa・Koui・Imi [Gesture, Action, and Meaning]. Tokyo: Kyoritu Shuppan.

Kita, S., Gijn, I., & van der Hulst, H. (2014). The non-linguistic status of the symmetry condition in signed languages: Evidence from a comparison of signs and speech-accompanying representational gestures. Sign Language & Linguistics, 17, 209-232. doi: 10.1075/sll.17.2.04kit

MacNeill, D. (1987). Psycholinguistics: A New Approach. Inc. New York, New York, U.S.A: Harper & Row.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M. & Levinson, S. C. L. (2004). Can language restructure cognition? The case of space. Trends in Cognitive Sciences, 8, 108-114. doi: 10.1016/j.tics.2004.01.003

Venditti, J. J. (1994). The influence of syntax on prosodic structure in Japanese. In J. J. Venditti (Ed.), Papers from the Linguistics laboratory, Ohio State working papers in Linguistics, 44 (pp. 191–223). Ohio State University, Department of Linguistics, Columbus OH.

Acknowledgments

We would like to thank all our participants. We would like to thank Professor Sotaro Kita of the University of Warwick for his invaluable suggestions. We would also like to thank Editage (www.editage.jp) for English language editing. This study was supported by JSPS/MEXT KAKEN JP17H06382 (H.K.) and JP16K04318 (H.K.)