F-4

A visual-world eye-tracking study on the cognitive processing for demonstratives in Japanese

Yusuke Sugaya

Our eye-tracking experiment with visual stimuli investigated the cognitive process giving rise to the utterance of Japanese demonstratives: "ko-," "a-," and "so-" series. The results mainly revealed the following: (i) The shared region of the speaker and the hearer is, by default, equally bisected into each of the areas of using "ko-" and "so-." (ii) This baseline can be affected by several factors related to semantic/pragmatic components involved in distance judgments, such as competitors, standards, and the hearer. (iii) Specifically, the use of "so-" evokes the processing of the position of and the view from a hearer, as participants focused more on the hearer when selecting "so-."

1 Introduction

Japanese demonstratives *ko-*, *a-*, and *so-* can be used to refer to a physical item in a speaker's view so as to introduce it in a linguistic context (e.g., "*ko*re wa watashi no hon desu" ≈ *this* is my book). This study focuses on such **spatial**, **deictic** usages of these expressions. The three-way system in Japanese and Korean seems to make the usage of them more complicated, unlike in English and Chinese which have just two options (i.e., **proximal/distal**): *this/that* and *zhè/nà*, respectively. Therefore, Japanese linguistics has long examined the principles of how to select among them in both deictic and anaphoric usage. However, previous studies on demonstratives in Japanese still cannot precisely predict their usages, as the latest artificial intelligence (AI) can never use these expressions exactly as the humans do, although it has a high capability of spatial recognition. If the thought and decision on demonstrative selection must be employed mentally, it is necessary to investigate the **mental processing** in some way for theoretical forecast. Thus, the current study seeks to answer the following general question: What psychological processing emerges in a speaker's brain to make an utterance including *ko-*, *a-*, or *so-* in the deictic usage?

2 Previous studies

Although the primitive or typical use of all demonstratives must be the spatial deixis, rather than conceptual and anaphoric references (i.e., the discourse deixis) (cf. Fillmore 1997), from the perspective of language acquisition and history (i.e., ontogeny and phylogeny), most Japanese indexical studies have been dedicated to the non-deictic use (e.g., Kuno 1973, 1992, Kuroda 1979, Kinsui and Takubo 1992) and the relation of deictic and non-deictic uses (e.g., Sakata 1971, Kinsui 1999, Hirata 2014, Tokimoto 2015).

Nevertheless, many theories have worked on how differently Japanese three deictic demonstratives are used (e.g., Sakuma 1951, Shoho 1981, Kamio 1990), many of them commonly claiming

that *ko-* refers to an object on the **speaker territory**, *so-* on the **hearer territory**, and *a-* on the outside of these territories. Moreover, some research has defined two kinds of speaker–hearer locations, so-called **opposition** (split) and **fusion** (unification) types. In the latter, a demonstrative is chosen, rather than in terms of these territories, but simply according to the distance from the speaker. In this type, there seems to be a lack of consensus as to whether the usage of *so* is used as **medial** (Yoshimoto 1992, Kinsui 1999, Tsutsumi 2012). Moreover, the way in which three demonstratives are composed appears to be in question; they consist of *ko–a* and *ko–so* alignments (Ogawa and Nozawa 2015) or only *so-* is diverged from the *ko–a* combination (Kinsui 1999).

Methodologically, however, few of these studies have adopted an experimental approach to make a *scientific* investigation into the cognitive process behind the usage of demonstratives. Traditionally, linguists have collected a small amount of language data or created a sentence and judged its acceptability based on their own intuition, on which linguistic theories have been constructed. All of those theories on indexicals are nonetheless concerned with, or reduced to the research on, the *conditions* as to how to use demonstratives *ko-, a-,* and *so-* properly. If so, they can be examined using experiments with visual stimuli presented before choosing an indexical.

3 Theory and hypotheses

The selection of demonstratives essentially involves **distance judgments**, that is, how close to or far from the speaker or the hearer an object (called a target here) is. Obviously, the distance is not actually measured in a physical manner, but varies psychologically according to different factors. Look at Figure 1, for example, and then consider which demonstrative you would choose among *ko-, a-,* and *so-* for the target, "[ko/a/so]-no cup wa kiiro (this/that cup is yellow)." Figure 2

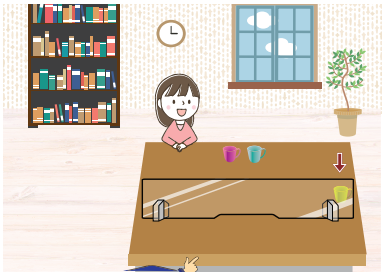

Figure 1: Choice of *ko/a/so* under a complex situation.

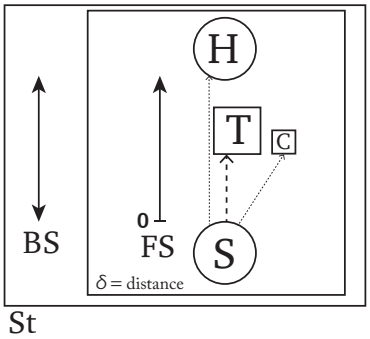

Figure 2: Structure of distance judgments.

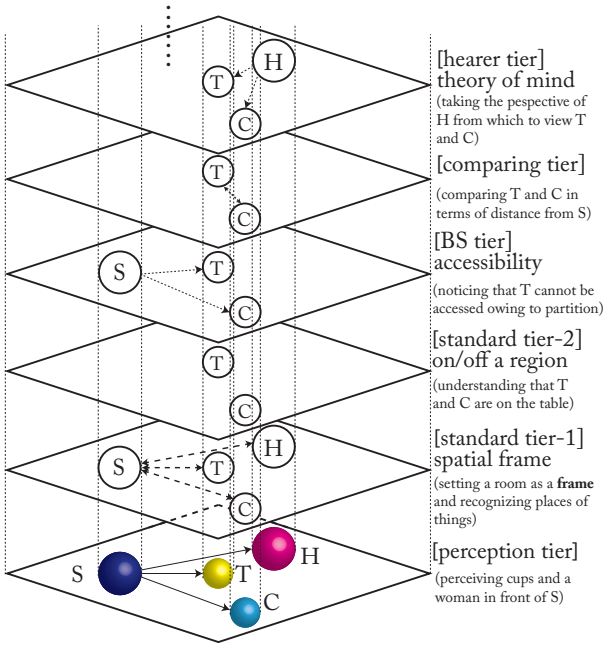

Figure 3: Multi-tiers model of demonstrative selection.

indicates the structure of distance judgments involved in the usage of demonstratives (cf. Sugaya 2015, 2020). This shows that, when a **speaker** (**S**) evaluates a **target** (**T**) in terms of distance (δ) as a **foregrounded scale** (**FS**), they may (i) compare the T with **competitors** (**Cs**), (ii) assume **backgrounded scales** (**BS**s) defining the FS, (iii) consider *qualitative* situations bringing about absolute **standards** (**St**s), and/or (iv) take the perspective from the **hearer** (**H**). Figure 3 is an applied model of explaining complex contexts, showing each processing individually with tiers divided. As tiers increase, the processing must be more complicated and so takes more time, in principal, before the utterance of a demonstrative.

After investigating the **baseline** of the usage of *ko-, a-,* and *so-*, this study tests each of the components and tiers. Developed from the literature, not only the territories of S (f_s) and H (f_h), but also the comparative distances of T from S (d_s) and H (d_h) must work out. Thus, our hypotheses are theoretically formalized as follows: ***ko-*; $(\mathbf{T} \ni f_s) \wedge (d_s < d_h)$, *so-*; $(\mathbf{T} \ni f_h) \wedge (d_s > d_h)$, *a-*; $(f_s \vee f_h) \notin \mathbf{T}$**. Put simply, the shared region (intersection) of S and H are **equally bisected**. Put differently, the (perpendicular) bisector of a line connecting the points of S and H works as separating the regions in which *ko-* and *so-* are used. In most cases, this ideal default can be affected by a number of factors. This study assumes that central ones among them would be related to the components and tiers presented above.

4 Eye-tracking experiments

To prove these hypotheses, we conducted an eye-tracking experiment with drawings based on the **visual-world paradigm** (e.g., Tanenhaus et al. 1995). This is a particular procedure or method, which assumes that the location at which a participant gazes in an image, is concerned with the place in their mental processing. This methodology could, therefore, determine what element should be processed for a certain language phenomenon.

4.1 Methods

Participants

A total of 11 (four female, eight male) participants, aged 19–43 (M = 24.9, SD = 7.4), were recruited on our university campus. All of them were either undergraduate or graduate students and native speakers of Japanese.

Materials: Drawings as visual stimuli

To test the hypotheses, 82 images were prepared, excluding the filler items. The abstract schemas, based on which all concrete drawings were created, are presented in Figure 4. The schemas b-1 and b-2, and their instances, examined the hypothesis concerning the *baseline* of the usage of *ko-, a-,* and *so-*. In addition, the schemas c-1 to c-9, s-1 to s-8, and h-1 to h-8 were prepared for testing the effects of *competitor*, *standard*, and *hearer*, respectively. Additionally, schemas t-1 and t-2 surveyed the influence of **familiarity** of T on indexical selection—S would tend to choose a proximal form *ko-* to refer to a psychologically close item (cf. Lakoff (1972) for the affective *this*).

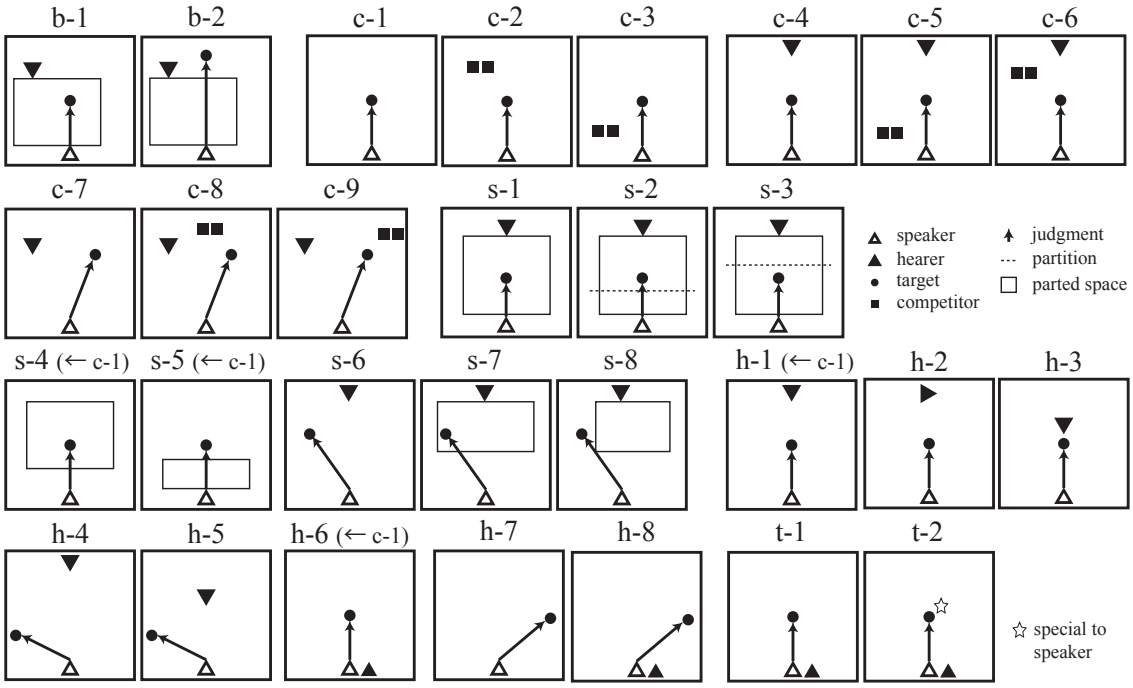

Figure 4: Schemas of visual stimuli indicating different conditions.

Apparatus and procedure

The Participants' eye movements were recorded by a screen-based eye-tracker (Tobii Technology's Tobii Pro Nano), with a sampling rate of 60 Hz, mounted on a 15.6-inch mobile monitor. Tobii Pro Lab, working as both a presenter and an analyzer, manipulated the entire experiment, set the areas of interest (AOIs) and times of interest (TOIs), and recorded the response and its duration, as well as gaze data during the trials.

Figure 5 illustrates the general procedure used in the experiment. Participants underwent a calibration process and several rehearsals before the experiment. In each of six sections, after situations or contextual information were clarified, trials comprising three steps, (i) options presented, (ii) fixation, and (iii) a drawing presented, were repeated.

4.2 Results and discussion

First, the response data (all participants' choices of demonstratives) on the stimuli instantiated from the schemas b-1 and b-2 supported the hypothesis on the baseline that territories of use of *ko-* and *so-* are divided by the bisector of a line connecting S and H. Also, *a-* is adopted to refer to a T outside the shared territories (this is concerned with St). Figure 6 and Table 1 present this in intuitional and statistical ways, respectively. Figure 6 presents the proportion of each demonstrative on different spots; *ko-, a-*, and *so-* appeared to be largely assigned for each of three distinct regions. Moreover, Table 1 reports the results of multiple logistic regression analysis, which measured the effects of distances from S and H and the positional relation of a T to a parted space (as independent variables) on choices of demonstratives *ko/a/so* (as dependent variables). The results provided a few important insights; (i) As for the *ko-/so-* pair, the distance from both S and H must be significant for the selection (S: $p < .001$, H: $p < .05$). (ii) As for the *ko-/a-* pair, rather than the distance from

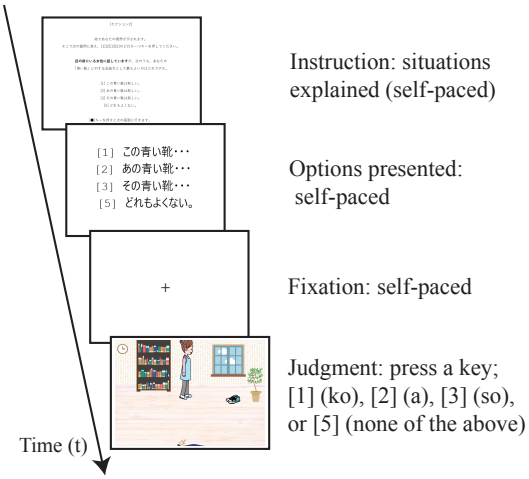

Figure 5: Examples of the trial sequence.

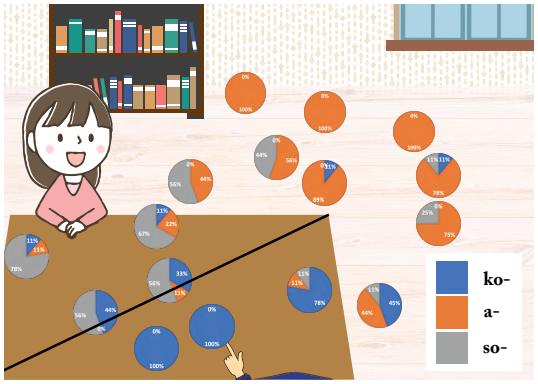

Figure 6: On the baseline of the uses of *ko-/a-/so-*.

		estimate value	standard error	odds ratio	95%CI	p
ko- and *so-*	intercept	-1.739	1.700	0.176	[0.005, 5.262]	
	distance from S	2.823	1.027	16.828	[2.664, 162.057]	**<.001**
	distance from H	-1.481	0.773	0.227	[0.039, 0.923]	**<.05**
ko- and *a-*	intercept	-4.828	1.716	0.008	[0.000, 0.195]	<.01
	distance from S	2.982	0.895	19.729	[3.935, 149.612]	**<.001**
	distance from H	-0.204	0.764	0.816	[0.154, 3.499]	
	on/around table	2.494	1.370	12.106	[1.041, 271.806]	**<.05**

Table 1: Results of the multiple logistic regression analysis.

H, the absolute distinction concerning whether a T is put on a table influenced the selection of *ko-/a-*. (iii) In both cases, the distance from S is primarily significant to choose any indexical.

Furthermore, the baseline could be affected by different factors related to the components of distance judgments. The effects of these components on the choice of *ko/a/so* were generally proved by the data of both eye movements and responses. Regarding the former, we calculated the ratio of the time gazing at the AOIs of each component to the entire time of each trial. As for the latter, as shown in Table 2, we not only used the analysis of association (between qualitative data, or LHS and RHS) but also computed the probabilities of change in participants' selection and the square test.

In terms of competitors, first, participants focused on Cs longer than the seemingly irrelevant items (distractors) over the trials of schemas c-1 to c-9, as each component was focused more in the following order: T (M = 0.337, SD = 0.173), C (M = **0.191**, SD = 0.158), H (M = 0.055, SD = 0.110), S (M = 0.037, SD = 0.061), and distractors (M = 0.016, SD = 0.036). However, there was no clear association rule according to the response data, and the appearance of Cs did not have significant effects on participants' selection of demonstratives. Second, the standard seemed to work in a significant way in selecting indexicals (specifically, in the case of a partition on a table). In terms of gaze duration, differences were not determined according to the different schemas from *s*-1 to *s*-3; however, the response data indicated that the significant impact of the St made lots of

			association analysis			probabilities of alternation	square test
component	LHS	RHS	support	confidence	lift	0-1	p
competitor	-	-	-	-	-	0.104	
standard	partition	*ko-/so-*	0.889	0.889	1	0.833	$<.05$
hearer	opposite	*so-/a-*	0.972	0.972	1	0.201	$<.05$
	face-to-face	*ko(a)-/so-*	0.981	0.981	1	0.535	$<.05$
	unification	*ko(a)-/so-*	0.990	0.990	1	0.279	
target	relation to S	*ko-/a-*	0.907	0.907	1	0.397	$<.01$

Table 2: Effects of each semantic/pragmatic component on demonstrative selection.

participants choose another indexical for the target that was in precisely the same locus.

Above all, the role of the hearer is concerned with one of the most relevant issues to resolve on Japanese indexicals, as emphasized in prior studies (e.g., Ogawa 2008, Hirata 2013). Both kinds of data confirmed the important role of H during the use of *so-*. As a result of the eye-tracking data of all trials including an H, the use of *so-* overwhelmingly called for the processing of (or focusing on) H—T (M = 0.334, SD = 0.184), H (M = **0.208**, SD = 0.191)), compared to *ko-* (T (M = 0.412, SD = 0.165), H (M = **0.073**, SD = 0.109), and *a-* (T (M = 0.420, SD = 0.225), H (M = **0.073**, SD = 0.127)). Furthermore, Table 2 suggests not only the strong association between the H (in different positions and directions) and the selection of *ko/a/so*, but also a strong probability of changing participants' responses. In addition, results on the images of the schemas t-1 and t-2, presenting a familiar target with a speaker, verified that psychological distance between S–T should be involved in indexical expressions.

5 Conclusion

Based on the literature, this study suggests a new theory on Japanese demonstratives from the perspective of distance judgments and examined some aspects of the theory via an eye-tracking experiment. Although additional tests are obviously required to strengthen it, the structural theory covering the flexibility in the use of demonstratives was largely demonstrated in an objective, scientific manner.

References

Fillmore, Charles J. 1997. *Lectures on Deixis*. Stanford: CSLI Publications.

Hirata, Miki. 2013. From distance to joint attention—the analysis of 'medial' so- in Japanese demonstratives using the concept of addressee's attention—[kyori kara kyodo tyusi e: kikite no tyui wo motiita tyukyorisizi no sokei no saibunseki]. *Papers from the 13rd National Conference of the Japanese Cognitive Linguistics Association* 13:461–472.

Hirata, Miki. 2014. An integrative analysis of deictic and non-deictic uses of Japanese demonstrative so-

using the concept of attention [tyui gainen-wo motiita sokei-no tyokuzi youhou to hi-tyokuzi youhou no touituteki bunseki]. *Gengo Kenkyu* (146):83–108.

Kamio, Akio. 1990. *The theory of territory of information: A Functional Analysis of Language [Zyouhou no nawabari riron: Gengo no Kinou-teki Bunseki]*. Tokyo: Taishukan.

Kinsui, Satoshi. 1999. On the relation between the deictic use and the non-deictic use of the Japanese demonstratives [nihongo-no sizisi niokeru tyokuziyouhou to hi-tyokuziyouhou no kankei nitsuite]. *Natural Language Processing [Shizen Gengo Shori]*. 6(4):67–91.

Kinsui, Satoshi and Yukinori Takubo. 1992. Japanese demonstratives seen from a dynamic model of discourse management [danwa kanri riron kara mita nihongo no shizishi]. In S. Kinsui and Y. Takubo, eds., *Studies on the Japanese Language: Demonstratives [Nihongo Kenkyu Siryoushu: Shizishi]*, pages 123–149. Hitsuji Shobo.

Kuno, Susumu. 1973. ko, so, a. In *Japanese grammar research [Nihon Bunpo Kenkyu]*, pages 185–190. Tokyo: Taishukan.

Kuno, Susumu. 1992. ko, so, a. In S. Kinsui and Y. Takubo, eds., *Studies on the Japanese Language Demonstratives [Nihongo Kenkyu Siryoushu: Shizishi]*, pages 69–73. Tokyo: Hitsuji Shobo.

Kuroda, Shige-yuki. 1979. On (ko,) so, a [(ko) so a ni tuite]. In *Japanese and English*, pages 41–59. Tokyo: Kuroshio shuppan.

Lakoff, Robin. 1972. Remarks on 'this' and 'that'. *Proceedings of the Chicago Linguistics Society* 9:229–246.

Ogawa, Noriko. 2008. A cognitive research on Japanese demonstratives: On the treatment of the listener in so-series [nihongo shizishi no ninchiteki kenkyu: So-kei shizishi niokeru kikite no itizuke saikou]. *Papers in Linguistic Science [Gengo Kagau Ronshu]* 14:57–88.

Ogawa, Noriko and Hazime Nozawa. 2015. Cognitive foundation of demonstratives and selection principles [shizishi no ninchi teki kiban to sentaku genri]. In M. Yamanashi, ed., *Studies in Cognitive linguistics [Ninchi Gengogaku Ronko]*, vol. 12, pages 115–166. Tokyo: Hitsuji Shobo.

Sakata, Yukiko. 1971. On the functions of demonstratives ko, so, a [shizishi ko so a no kinou nituite]. *Tokyo University of Foreign Studies* 21:125–138.

Sakuma, Kanae. 1951. *Modern Japanese Expression and Usage [Gendai Nihongo no Hyougen to Gohou]*. Tokyo: Kouseikaku.

Shoho, Isamu. 1981. The system of ko/so/a [ko, so, a no taikei]. In *Japanese demonstratives [Nihongo no Sizisi]*, pages 51–122. Tokyo: National Institute for Japanese Language and Linguistics.

Sugaya, Yusuke. 2015. Inside the mental mechanism of evaluation processing. *Papers in Linguistic Science* 21: 153–182.

Sugaya, Yusuke. 2020. *Constructing Evaluations: The Meaning-Making Process of Adjectives*. Ph.D. thesis, Kyoto University.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217):1632–1634.

Tokimoto, Shingo. 2015. An experimental description of deictic and nondeictic usages of Japanese indexicals, ko-, so-, and a-forms [ko-so-a siyou kozinsa no zikkenteki kijutu]. *Mathematical Linguistics* 30(1):1–13.

Tsutsumi, Ryoichi. 2012. *An Overall Study of the Demonstratives of Modern Japanese [Gendai Nihongo Shizishi no Sougouteki Kenkyu]*. Tokyo: Koko Shupan.

Yoshimoto, Kei. 1992. The system of Japanese demonstratives ko/so/a [nihongo shizishi ko/so/a no taikei]. In S. Kinsui and Y. Takubo, eds., *Studies on the Japanese Language: Demonstratives [Nihongo Kenkyu Siryoushu: Shizishi]*, pages 105–122. Tokyo: Hitsuji Shobo.