The addressee’s attention and addressee-anchored proximity: 
The semantics of the Japanese demonstrative middle term so-

HIRATA, Miki

Abstract
In traditional studies, it is argued that a middle term in three-term demonstrative systems has two semantic features as its prototype; ‘Medial’ and ‘Proximal to addressee.’ The Japanese demonstrative middle term so- has been considered as encoding these two features. This paper focuses on the communicative function of demonstratives, which is the coordination of joint attentional focus as suggested by recent studies and reanalyzes the meaning of so- and its diachronic change by using natural conversation data and historical resources. Finally, this paper makes two claims: (a) the so-called ‘Medial’ so- is selected, in contrast to the Turkish şu and in a similar way to the Jahai ton, when the speaker considers that the addressee’s attention is already directed to the intended referent, (b) the prototypical meaning of so- is the information in relation to the addressee’s attention from which the feature ‘Proximal to addressee’ is derived from.

Keywords: Japanese demonstratives, Addressee’s attention, Natural conversation, Diachronic change

1. Introduction
The main goal of this paper is to suggest the process in which a middle term of the demonstrative system semanticizes the location of the addressee from the semantic feature in relation to the addressee’s attention, through the synchronic and diachronic analysis of the Japanese demonstrative so-. In traditional studies, it is argued that a middle term in three-term demonstrative systems has two semantic features as its prototype; ‘Medial’ being a short distance from the speaker, and ‘Proximal to addressee’ being near the addressee (Fillmore 1971, 1982, Anderson and Keenan 1985, Diessel 1999). These studies assumed that these two meanings might coexist independently. Meira (2003), however, suggested the tendency where the ‘Medial’ form conventionalizes the addressee as an anchor of a reference and derives ‘Proximal to addressee.’

In contrast, some recent studies introduce the status of the addressee’s attention as a new concept to describe the meaning of a middle term (Özyürek 1998, Küntay and Özyürek 2002, 2006, Burenhult 2003). Based on the detailed analysis of natural conversation data, they pointed out that some middle terms are used regardless of the spatial location of their referents. Küntay and Özyürek (2002) claimed that Turkish employs the middle term şu to direct the addressee’s attention to a new referent. Conversely, Burenhult (2003) suggested that the Jahai1 demonstrative form ton is used to refer to a referent which already has the addressee’s attention.

This paper makes two claims based on the analysis for the synchronic conversation data and diachronic literatures: (a) the Japanese demonstrative so- is selected, in contrast to the Turkish şu and in a similar way with the Jahai ton, when the speaker considers that the addressee’s attention is already directed

1 A Mon-Khmer language spoken in the northern Malay Peninsula and adjacent parts of southern Thailand.
to the intended referent, (b) the prototypical meaning of so- is the information in relation to the addressee’s attention from which the information on the addressee’s spatial location is derived from. These claims suggest that in demonstratives, the spatial information is just the secondary feature that is given in order to adjust the addressee’s attention and establish the joint attention to an intended referent.

2. Semantic features of middle terms

2.1 ‘Medial’ and ‘Proximal to addressee’

The traditional literature on demonstratives has maintained that demonstratives are a typical member of spatial deixis, encoding the relative distance contrast with respect to the position of the speaker and/or addressee (Fillmore 1971, 1982, Anderson and Keenan 1985, Diessel 1999). These studies assume two-term systems, indicating the contrast between proximal and distal relative to the speaker as the most standard system in demonstratives (Anderson and Keenan 1985: 281, Diessel 1999: 36). Demonstrative systems with three terms have a middle term in addition to this basic opposition proximal and distal term. Fillmore (1982: 49-50) suggested that middle terms represent two features as their semantic prototype: a short distance from the speaker and close to the addressee. Along with his description, Anderson and Keenan (1985) distinguished two types of three-term systems: distance-oriented and person-oriented (ibid. p. 282). In the first system, the middle term refers to the medial distance relative to the speaker, and in the latter, the middle term encodes the proximity to the addressee.

2.2 Semanticization of the addressee

While most previous studies treated ‘Medial’ and ‘Proximal to addressee’ as two independent semantic features, Meira (2003) pointed out that they may have continuity and suggested a possible path of evolution from a distance- to a person-oriented system. His suggestion is based on the analysis of the Tiriyô³ demonstrative system through fieldwork using a questionnaire. Tiriyô has a three-way distinction in its demonstrative pronoun system,³ ‘Proximal’, ‘Medial’, and ‘Distal’ all anchored to the speaker, and had been described as a typical distal-oriented system. But Meira (2003) found a tendency to favor ‘Medial’ terms over ‘Distal’ terms when the addressee is close to the referent. It suggests that the position of the addressee plays an important role in the choice of demonstrative forms, even in the so-called distance-oriented system.

Behind this tendency is the cognitive saliency of the addressee in the speech situation. Since the addressee is more salient than other positions in the conversational space, his/her position is a good candidate for a reference point and thus may have effects on the speaker’s choice of demonstrative forms, even in systems that lack addressee-anchored terms (Meira 2003: 9-10). This pragmatic salience of the addressee causes the middle term to evolve into an addressee-anchored term by semanticizing the position of the addressee. Meira (2003) considered that this tendency is not only in Tiriyô, rather, it is the cross-linguistic possible path of evolution from ‘Medial’ into ‘Proximal to addressee’ feature. Given the principle

---

2 A Cariban language spoken in the Amazon region.
3 According to Meira (2003: 4), Tiriyô has two additional forms, anaphora and a term that refers to invisible referents in the same formal class as the demonstratives.
of Occam’s razor, Meira’s (2003) analysis might be more plausible than those of previous studies: he reduced two semantic prototypes for the middle term to one and assumed the other is the result of semantic derivation from the prototypical feature.

2.3 From spatial location to the status of attention
In some recent studies, the validity of the “basicness” of spatial information in demonstrative semantics has been questioned. For example, Levinson (2004: 109) states that there is no a priori reason to suppose demonstratives are spatial in nature, and argued demonstratives are used to draw the addressee’s attention to some features of the surrounding environment in order to find the intended referent. Diessel (2006) focused on the importance of joint attention, the social behavior that provides a foundation for the verbal communication of humans to explain the function of demonstratives. In his definition, demonstratives serve the communicative function of coordinating the interlocutors’ joint focus of attention in addition to the space demarcation (ibid. p.469).

Along these lines, some recent studies of demonstrative usage in individual languages introduced the new concept, which may contribute more directly to the establishment of joint attention in order to describe the meaning of middle terms. The representative example is the analysis of the Turkish middle term şu in the series of studies by Özyürek and Küntay (Özyürek 1998, Küntay and Özyürek 2002, 2006).

Turkish has three terms in the demonstrative system: bu, şu, and o. Bu and o encode speaker-anchored proximal and distal respectively, whereas şu does not represent any spatial information and may refer to any objects in the conversational space. It does not fall into two prototypical categories presumed in traditional distance-based studies; thus, Turkish demonstratives had previously been considered to comprise an exceptional three-term system (Anderson and Keenan 1985: 285). Özyürek (1998) and Küntay and Özyürek (2002, 2006) provided a detailed investigation of videotaped conversation data to show şu is fully devoted to the communicative function of attention drawing. The natural conversation data shows that the presence or absence of the addressee’s attention on the referent is a crucial factor in determining the speaker’s choice of demonstrative forms (cf. Küntay and Özyürek 2002: 338). In cases where the addressee’s visual attention had already been directed to the referent, bu and o are used in relation to its spatial location. On the other hand, şu is selected to refer to objects to which the addressee’s visual attention is not directed. The following example (1) provided by Özyürek (1998) shows the distribution of şu in natural discourse. It is a part of spontaneous conversation in a ceramics class where a teacher and students discuss the ceramic objects placed around them.

(1) Oval object
(The teacher and students are standing around the table that contains the ceramic object of the focus of conversation. Then, a student points at the oval object that is far away from the table and proposes to put it on the ceramic object on the table.)

a. mesela hocam şu oval mesela
for example sir this/that oval for example

“sir for example, this/that oval for example”
b. şunun dis yuzeyine koyup ta
this/that-GEN outside surface-POSS-DAT put
“if you put on this’s outside surface”
c. ondan da olabilir
that-ABL also be
“it might be from that too”

(Özyürek 1998: 606-607, glosses and English translations are as in the original)

In (1a, c), the speaker uses şu and o respectively to refer to the same ‘oval object.’ Therefore, it is not the distance between the referential object and the participant that differentiates them. From the point of view of attentional contrast, in example (1), the speaker selected şu to introduce the referent that is not in the joint focus of attention in (1a), whereas once the referent got the addressee’s attention due to the reference of şu, used o to refer to the same object in (1c). Küntay and Özyürek (2006) described the meaning of this şu as the ‘absence of addressee’s attention.’

On the contrary, Burenhult (2003) argued that the Jahai demonstrative form ton is selected when the speaker is able to recognize that the addressee’s attention is present on the intended referent. The analysis of the task-oriented interaction between Jahai native speakers showed that ton refers to the object that is visually attended by the addressee or in the focus of joint attention previously established in the context. This feature is reflected in the tendency of ton to appear in the sequential final position in the discourse as well as its pragmatic function in referential situations: contrary to remaining forms in Jahai demonstratives, ton does not instruct interlocutors to divert their attention, rather, it signals that interlocutors should not divert their attention from the object in focus (Burenhult 2003: 377).

Interestingly, ton seems to refer not only to the object that has the addressee’s attention but also to the addressee-proximal object. Burenhult (2003) covered the usages of ton via the wide concept of “cognitive accessibility” to the addressee, but assumed the dominance of the concept in relation to the status of the addressee’s attention to the spatial information. His explanation of the fact that referents of ton are frequently addressee-proximal is that referents located in close proximity to the addressee are the ones that are most likely to be known or attended to by the addressee (ibid. p. 367).

In this paper, I will argue that the Japanese demonstrative so- also encodes the status of the addressee’s attention as well as Turkish şu and Jahai ton; and the spatial information of so-, proximity to the addressee, is the derivative information from it as Burenhult (2003) assumed. In the next section, I will re-analyze Japanese so-called ‘Medial’ so- using the natural conversation data and show that it indicates not spatial information but rather the ‘presence of addressee’s attention.’

3. Re-analysis of the Japanese demonstrative so-
In the traditional studies of demonstratives, the Japanese demonstrative has been described as prototypical person-oriented system with a proximal ko-, distal a-, and middle term so- indicating the proximity to the addressee (Fillmore 1982, Anderson and Keenan 1985, Diessel 1999). However, Japanese-specific demonstrative studies have pointed out that so- has two-fold meanings, encoding ‘Medial’ as well as
‘Proximal to addressee’ (Hattori 1968, Imai 2003, etc.). Imai (2003) demonstrated this is the case by conducting spatial-focused experiments involving Japanese native speakers: informants were asked to refer to objects placed at different locations on the table. He called the Japanese type of demonstrative system the ‘dual-anchor system’ since middle terms can indicate medial distance from the speaker and proximity to the addressee.

In contrast to this space-based approach, Özyürek and Kita (cited in Levinson 2004) provided an alternative analysis of so-based on close investigation of videotaped conversational data. Levinson (2004) summarized their claim as follows:

For Turkish the correct analysis seems to be that şu presumes lack of joint attention and is used to draw the attention of the addressee to a referent in the context […]. A similar story can be told for Japanese: so has two functions – one simply to indicate that the referent is close to addressee, the other (as with Turkish şu) to draw the addressee’s attention to a new referent. This latter usage is pre-empted by ko when the referent is very close to speaker, and by a when far from both speaker and addressee. A primary opposition here involves not proximity to speaker vs. addressee, but rather shared vs. non-shared attentional focus. (Levinson 2004: 110, underlines added)

The present study shares the same view: the so-called ‘Medial’ so-in fact encodes the concept directly related to the opposition, whether or not the attentional focus to the intended referent is shared, as well as the Turkish şu. However, unlike Özyürek and Kita’s claim, it does not seem to be the case that so- and şu serve the same function to “draw the addressee’s attention to a new referent.” This is due to the asymmetric distribution they show with respect to the status of the addressee’s attention in natural conversation. In the next section, I will demonstrate the distribution of so-in the discourse by using the natural conversation data and argue that so-is relevant to the object that already has the addressee’s attention.

3.1 The distribution of so-in natural conversation

Conversational data were collected from spontaneous interaction between two Japanese native speakers who had a previously established friendship. All conversations were videotaped and transcribed to investigate the physical and interactional context factors underlying the speakers’ choice of demonstratives.4

From the data, it was revealed that in the process of establishing joint attention, so-is frequently selected to refer to the object, which is indicated by other demonstrative forms in the initiation of the series of conversation. I provide example (2) to illustrate this usage of so-. The conversation in (2) took place in the observatory. Participants A and B sat in chairs across from the window and talked on the landscape below. While they were talking, A found the building in which their mutual friend C was living and tried to explain this to B.

4 The following data are quoted from Hirata (2014). Hirata (2014) provides detailed analysis on the usage of so-using conversational data.
(2) Zebra-patterned building

a A: (pointing gesture)  
asoko  
that place  
look diagonal  
forest  
exist  
right?

“Look over there. There is a forest diagonally across, right?”

b B:  
(moves the face according to A’s pointing)

c A:  
soko  
no  
ura!

“Behind that place”

shiro kuro  
no  
zebra-gara  
no  
biru  
zebra-gara  
no  
black-white  
zebra-patterned  
biru  
zebra-patterned  
building  
“black and white, zebra-patterned building. Zebra-patterned one”

d B:  
edore  
dore?  
(pointing gesture)  
ano  
tettoo  
ga  
chikai  
yatsu?

“Huh? Which one? The one close to that iron tower?”

e A: 
(turns to look at the side of B and then turns back to the front and makes a pointing gesture)  
sore  
sore!

“That one, that one over there!”

shiro to kuro  
ja  
a  
guree  
no  
massugu  
mee  
white  
black  
black  
not  
gray  
P  
P  
straight  
ahead  
“Black and white, wait no it’s not black it’s gray. That one right ahead of you.”

f B:  
ap  
are  
kan  

“Um, it might be that one.”

g A:  
(moves the face towards B)  
soo  
soo  
right  
right

“Right, right.”

(turns back to the front)  
nezumi-iro  
no  
yatsu  
mouse-color  
thing  
“The gray one.”

---

5 All the examples cited in this paper are composed of three parts: on each first line, there is a Romanized original Japanese transcript with the participants’ body movements in parentheses, and beneath those are phrase-by-phrase glosses. Finally, a rough English translation is added after each utterance. Those target forms are marked in bold. The abbreviations for the Japanese glossing are:

JD: Judgmental, P: Particle, NOM: Nominalizer.

6 Each Japanese demonstrative is formed by two kinds of features: deictic features  
deo-,  
sao-,  
a-,  
and qualitative features which classify the referent (see Diessel 1999: 35). In addition to basic pronominal forms  
kore/sore/are  ‘this/that/that one’, Japanese demonstrative system includes pronominal forms for location  
koko/soko/asoko  ‘this/that/place,’ for direction in formal usage  
kochira/sochira/achira  ‘this/that/that way’ and informal usage  
kocchi/socchi/acchi  ‘this/that/that way’; demonstrative determiners  
kono/sono/ano  ‘this/that/that’; adnominal forms for characteristics of an object  
konna/sonna/anna  ‘this/that/kind of’; and adverbial forms  
koo/soo/aa  ‘in this/that/that manner.’
At the beginning of the conversation, A indicates the space where the intended building is located with the *a*-form, as in (2a). In (2c), the adjacent space is referred to with the *so*-form and in the closing part of the conversation, as in (2e,g), the intended building is indicated using the demonstrative pronoun “sore”, which was referred to as “are” in B’s previous turn, (2f). Traditional space-based studies cannot explain how this term-switches between the *a*- and *so*-forms.

This study proposes that this type of *so*- is selected according to the status of the addressee’s attention with respect to the intended referent as well as *şu*. However, if *şu* encodes the ‘absence of addressee’s attention’ as Küntay and Özyürek (2006) described, *so*- represents the ‘presence of addressee’s attention’ on the contrary. The comparison between Turkish and Japanese conversation data shows the asymmetrical distribution of these two parties. While *şu* is used typically in the initiation of the conversation to draw the addressee’s visual attention to a new referent as showed in (1), *so*- in (2) refers to the referent that is introduced by linguistic antecedents such as *a*- and/or pointing gestures and thus is recognizable for the speaker that the addressee previously directed his/her visual attention to it. This contrasting distribution in the discourse shows that *so*- is the form selected when the speaker is assured that the addressee’s attention
is present on the intended referent in contrast to ū.

Note that this so- is not always used to indicate objects that are already referred to by other terms. If the context provides the evidence to determine the speaker’s assurance that the referent has the addressee’s visual attention, (s)he is able to indicate to the referent using so- without any previous references. In example (3), conversation participants C and D are given a picture of a model building and asked to construct this building collaboratively with blocks on the desk. The model picture is projected on the wall at a distance of around 5 meters from the desk where C and D are (see Figure 5). While D said to C that white blocks are not enough to construct the building in the picture, C explained that they do not have to construct the same colored building with the model by quoting the instruction given to them by the organizer (author).

(3) Model picture

a. C: iro issho ja naku temo ii tte itteta mon
   color same P not even if okay P said NOM
   “She said we don’t have to use the same colored blocks as the model.”

b. D: (looks up at the screen) a sooka a soo nano ka (looks down to blocks at hand)
   oh I see oh I see NOM P
   “Oh, I see, is that so?”

c. C: itteta yo
   said P
   “Yeah she said so.”

d. D: (looks up the screen) sooka
   I see
   “Got it.”

e. C: (gives a quick look at D and drops her eyes to the blocks) nanka sore o mite
   kinda that one P look
   “Look at it,”
   (after B turns his gaze to the blocks on the desk) ppoku sure ba ii n da yo
   like do if okay NOM JD P
   “and make it like that.”

7 In (3f), B indicates the referent with ‘are’, even though A is obviously directing her attention to it and referred to it with ‘sore’ in the previous turn, as in (3e). B’s choice of ‘are’ is seemingly a counterexample to my analysis, but this is not; at the end of B’s utterance in (3f), there is a sentence-final particle ‘kana’, which expresses that the utterance is a part of internal speech. Therefore, it may be that this utterance is not addressee-oriented but similar to self-talk to organize input. The fact that B did not select so- but are- in the non-addressee-oriented utterance supports the claim of this paper; so- is fully devoted to the communicative function of attention drawing, similar to ū.
Contrary to (2), in (3), there are no linguistic antecedents, pointing gestures, or body movements to indicate the intended referent before the reference with “sore” in (3e). Instead, in (3e), C turns her face to D and checks the direction of D’s gaze before the reference (see Figure 4). In addition, contrary to (2) where there are a lot of candidate referents in the interlocutors’ visual field, in the speech situation of (3e), the model picture projected on the wall is the only candidate for the referent in their visual field, as shown in Figure 5. It may be that the fact that the addressee is directing his attention to the referent as well as the saliency of the referent itself in the speech situation allows the speaker to use “sore” without any previous references in (3e).

Similarly, the example (4) also shows a case of direct reference with so- without any antecedents.

(4) English classroom

(E and F sit down at a table across from each other in a classroom.)

a. E: ashita eigo aru?
   tomorrow English exist
   “Do you have an English class tomorrow?”

b. F: (nodding) un aru
    yeah exist
    “Yeah, I do.”

c. E: eigo no kyooshitsu doko?
    English P classroom where
    “Where is the classroom?”

d. F: (keeps her gaze on E’s face and points behind her) sore
    that place
    “There.”
In situations in which the speaker and addressee are facing each other and the referent is located behind the speaker, as in (4), so- is frequently selected instead of ko- or a-. This so- does not represent the proximity to the addressee; in the above situation, the referent is obviously closer to the speaker than the addressee.

In terms of the interlocutor’s attention state, in the above situation, the speaker can monitor the direction of the addressee’s attention constantly as they are facing each other. Moreover, in such face-to-face situations, the referent behind the speaker may inevitably come within the addressee’s visual field. Thus, in (4), speaker F does not draw the interlocutor’s visual attention to the referent by demonstrative terms as in (3a), nor does she turn her face to check the direction of the interlocutor’s gaze, as in (3e, g) and (4e). This face-to-face situation has been given as a typical situation where the so-called ‘Medial’ so-, which is considered to represent the ‘presence of addressee’s attention’ in this study, appears in the previous studies on Japanese demonstratives (Hattori 1968 and so on.) This may occur because in such situations it is easy for the speaker to monitor the direction of the addressee’s attention, thus so-, which denotes the status of attention, tends to be selected.

3.2 The use of so- in linguistic discourse

As an antithesis to traditional space-based studies, recent studies on demonstratives in individual languages introduced in section 2.3, such as Özyürek or Burenhult, have been engaged in the close observation of the context of natural conversation. Thus, their focus has been on the exophoric use of demonstratives with the reference to concrete entities in the surrounding environment. Although they note the concept of the addressee’s attention may be relevant not only to the reference in physical situations but also to the reference to linguistic elements in the discourse (Burenhult 2003: 367, Küntay and Özyürek 2006: 319), the empirical studies on it have not been conducted.

Furthermore, in the theoretical studies, it is already argued that the communicative function of demonstratives can be extended from the physical world to linguistic discourse. Diessel (2006: 481) pointed out that the discourse use of demonstratives involves the same psychological mechanism with their exophoric use, and argued that the addressee’s visual attention on a concrete entity in the physical world will be extended to the attention on a linguistic element in the universe of linguistic discourse.

Within this framework, so- and şu, which show the contrastive distribution regarding the status of the addressee’s attention to the referent in the physical situation, are supposed to be used contrastively even
in the linguistic discourse. I will examine this supposition by comparing the discourse deictic use of so- and şu as following. The example sentences are quoted from Hayashi (1989), who described the difference between the Japanese and Turkish demonstrative system.8

(5) Reference to what follows
a. Hasana {bunu / şuun} şöyleyeceğim, “Anneni meraklandırma.”
   Hasan-DAT this one-ACC / that one-ACC I will say your mother-ACC not make worry
b. Hasan ni { *soo / koo } iu tsumori desu. “kimi no hahaoya o shinpai saseru na”
   Hasan P like that / like this say going to JD you P mother P worry make P
   “I will say this to Hasan, don’t make your mother worry.”

(6) Reference to what precedes
a. Bugün ders yokmuş ama, {bunu / *şunu} hiç bilmiyordum.
   today class seems no but this one-ACC / that one-ACC never I didn’t know
b. kyoo jugyoo ga nai soo da ga, watashi wa {sore / kore} o makkaku shiranakatta.
   today ckass P no seems JD but I P that one / this one P never didn’t know
   “Though they say we have no class today, I didn’t know it at all.”
   (Hayashi 1989: 110, glosses and English translations added9)

Table 1. Distribution of so- and şu in linguistic discourse

<table>
<thead>
<tr>
<th></th>
<th>Turkish</th>
<th>Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference to what follows</td>
<td>*şu / bu</td>
<td>so- / ko-</td>
</tr>
<tr>
<td>Reference to what precedes</td>
<td>şu / bu</td>
<td>*so- / ko-</td>
</tr>
</tbody>
</table>

As shown in (5a), şu, which represents the ‘presence of addressee’s attention’ in the physical speech situation, refers to the linguistic elements that have not been introduced into the discourse at the speech time and thus are impossible to be seen as attention-directed. In contrast, the choice of so- is not grammatically correct in this context: so- is perfectly suitable for the reference to linguistic elements that have already been introduced into the previous discourse, that is, in the context where joint attention had already been achieved to the intended referent as in (6b). Şu is not allowed to be used in this context (see (6a)).

As illustrated above, the asymmetric distribution of so- and şu is maintained in linguistic texts as well.10 This fact supports my claim: the Japanese demonstrative so- is selected, in contrast to Turkish şu,

8 In previous studies, it has been generally accepted that demonstratives have three distinct usages in addition to the exophoric use; the discourse deictic use, anaphoric use, and recognitional use (Diessel 1999, Levinson 2004, etc.). I used the discourse deictic use to examine the distribution of so- and şu in the linguistic discourse since only this use may refer to the discourse that precedes as well as to the discourse that follows in the Japanese demonstrative system.
9 Three Turkish native speakers judged the acceptability of the use of bu and şu in (5a) and (6a).
10 In their anaphoric use, so- and şu show the asymmetric distribution; so- is frequently used as anaphora, whereas şu
when the speaker assumes that the addressee’s attention is already directed to the intended referent. This argument also shows that in terms of the addressee’s attention, the integrated account of the exophoric and non-exophoric use will be possible.

3.3 The presence of the addressee’s attention

According to the above results, it is revealed that so- has the use that represents the ‘presence of addressee’s attention.’ Its frequent anaphoric use in the linguistic discourse is considered as one of the manifestations of the concept of the ‘presence of addressee’s attention.’ Based on the above discussion, I define the contextual environment in which the term for ‘presence of addressee’s attention’ is selected as follows:

(7) Descriptive characterization of the ‘presence of addressee’s attention’

The term for ‘presence of addressee’s attention’ is used to inform the addressee that s/he is already directing his/her attention to the intended referent at the time of speech, or his/her attention was directed to the intended referent before the time of speech in order to establish joint attention.

4. Diachronic analysis of so-

In the previous section, I argued that the so-called ‘Medial’ so- should be interpreted as a term that expresses the ‘presence of addressee’s attention,’ in contrast to the Turkish şu. The remaining issue relates to another use of so-: ‘Proximal to addressee.’ As mentioned in section 2, Meira (2003) suggested the possible path of evolution from ‘Medial’ into ‘Proximal to addressee’ feature based on field work using a questionnaire with Tiriyó native speakers. Burenhult (2003: 367) posited that the ‘Proximal to addressee’ is the derivative use from the concept of the ‘presence of addressee’s attention,’ since the referents in close proximity to the addressee are the ones that are most likely to be attended to by the addressee. Although they suppose different prototypical meanings for middle terms, they share the view that the spatial information in relation to the addressee is secondary information. Meira (2003: 10) stated that “it would be interesting to check, for the language families whose history is better known, if there is any evidence of this kind of diachronic change.” In this section, I will present the very analysis that Meira (2003) stated: the examination on the diachronic semantic change of so- based on historical literature and previous references.

4.1 Ca. 8th century

It appears that originally, the Japanese demonstrative had a two-term system with ko- and so-. The most ancient Japanese literature we can now refer to is the poetry anthologies compiled around the 8th century. In poetry written from the beginning to the middle part of the 8th century, we cannot find any examples that include the demonstrative ka-, which is the previous form of the current a- (hereinafter called ka(a)-). Man’yoshu (Collection of Ten Thousand Poems), which was compiled during the latter half of the 8th century, includes the use of ka(a)-, but the amount and use of the inflectional forms are very limited compared to ko- or so- in the same literature. Therefore, diachronic studies have considered that the

never refers to linguistic antecedents. Besides, as with so-, the Jahai demonstrative ton also has anaphoric use in the discourse (Burenhult 2003: 366-367).
Japanese demonstrative system was originally constructed from two terms: *ko* and *so*. Thus, around the 8th century, *ka(a)* had just emerged and was still not well-developed.

Therefore, before the development of *ka(a)*, what contrast had been shown by *ko* and *so*? Before *ka(a)* became the well-established term, the Japanese demonstrative system seemed not to have included any spatial information of the referent. Old literature shows that all perceptible objects in the speech situation had been referred to by *ko*, regardless of their locations at that time (Li 2002: 256, Okazaki 2010: 88). *So*, on the other hand, had the following three main usages ((8a–c)) and exophoric usage ((8d)).

(8) Usages of *so* around the 8th century

a. Anaphoric use

\[
\text{waga yado ni hana so sakitaru, so o miredo, kokoro mo yukazu…}
\]

my house P flower P blooming that one P even if (I) see heart P not lighten

“Flowers are blooming in my garden. Even though I saw those, my heart doesn’t lighten.”

(Man’yoshu 366)

b. Discourse deictic use

\[
\text{wagi moko ni koitsutsuoreba harusame no sore mo shiru goto yamazu furitutu}
\]

my girl P when missing spring rain P that one P know as not stop raining

“When I’m missing my girl, spring rain keeps raining as if it knows it.”

(Man’yoshu 1933)

c. Recognitional use

\[
\text{Nubatama no sono yo no ume o tawasurete orazu kinikeri…}
\]

pitch-black P that night P plum blossom P forget not break have come

“The pitch-dark night, in that night I have come back without breaking the branch of the plum blossom”

(Man’yoshu 466)

d. Exophoric use

\[
\text{sode furaba mitsubeki kagiri wa wa aredo}
\]

sleeve if wave can see as long as I P though be

“Though I’m seeing you waving the sleeve as long as I can see,”

\[
\text{sono matsu ga eda ni kakurainikeri}
\]

that pine P branch P have been hidden

“Your figure went out of sight behind that branch of the pine tree.”

(Man’yoshu 2485)

11 In the following examples, I focus on only demonstrative pronouns and demonstrative determiners among all forms outlined in the note 6. In fact, Okazaki (2010) pointed out that the semantic development of adverbial forms in Japanese demonstratives is behind, as they follow the development of other forms. Although she raises an interesting fact in analyzing the historical change of Japanese demonstratives, this paper will not take into account the ways in which their development is the same.
Almost all examples of *so*- in poetry around the 8th century refer to linguistic items in the previous discourse, as in (8a) and (8b). As shown in the previous section, the currently used *so*- also has these usages. As for the recognitional use in (8c), that is, the demonstrative’s function to activate specific shared knowledge between interlocutors (Diessel 1999: 105), it is assigned to *ka(a)*- in the current Japanese demonstrative system. With the development of *ka(a)*-, *so*- gradually lost this usage. Taking these usages into account, diachronic studies on the Japanese demonstrative system defined *so*- in the 8th century as the term that indicated something imperceptible, such as linguistic items or elements in shared knowledge, in contrast to *ko*- (Li 2002, Okazaki 2010). However, though very few, we can see *so*- referring to visible entities in the surrounding situation as in (8d). The entities referred to by *so*- seem not to be obviously close to the speaker. Based on this observation, Li (2002: 163-166) hypothesized that a part of the visible non-proximal entities was not recognized as being within the perceptible realm, and thus indirectly referred to by using *so*-.

At the same time, *ka(a)*- was on the way to becoming a non-proximal term, which is used for objects that are located at a far distance from the speaker (Hashimoto 1966).

(9) Exophoric use of *ka(a)*- around the 8th century

okibe yori michikuru shio no iya mashi ni wa ga omou kimi ga mifune kamo *kare*
offshore from come tide P more increase P I P think you P craft P **that one**

“As the tide comes in from offshore, I’m missing you more and more. **That one** might be your craft.”

(*Man’yoshu* 4045)

*Ka(a)*- is assumed to derive from *ko*- by vowel alternation (Hashimoto 1966). In diachronic studies, it is generally accepted that the derivation of *ka(a)*- defined the realm of non-proximity anchored to the speaker in the speech situation so that the addressee could attend to the intended referent more effectively using this specific information. According to Li (2002: 142), it is conceivable that the exophoric *so*-., as in (8d), and *ka(a)*- were interchangeable at that time. However, since number of the examples of the exophoric *so*- is very limited even compared to the developing *ka(a)*-, it is not clear when the speaker recognized entities in the surrounding situation as imperceptible and selected *so*- to refer to them, instead of direct-referential terms *ko*- or *ka(a)*-.

4.2 Ca. late 10th to 11th century

From around the 9th century onwards, a large body of court literature was being produced in Japan. The most notables are *Genji Monogatari* (The Tale of Genji) and *Makura no Soshi* (The Pillow Book), both compiled around the late 10th to 11th centuries. These literary works provide more examples of the exophoric *so*- than previous poetry. In this era, in addition to the original exophoric usages, the indirect reference to non-proximal entities, as in (8d) or (12), we can observe that *so*- indicates addressee-proximal referents, as in (13).

(12) Original exophoric use of *so*-
(Interlocutors are together on one cart.)

kaku iu tokoro wa Akinobu no ason no ie narikeru.
in this way say place P Akinobu P couturier P house JD

“Where we talked in this way was in front of the courtier Akinobu’s house.”

’soko mo iza min’ to iite, kuruma yosete orinu.
that place P let’s see P say and cart pull over and get out

“(She) said ‘Let’s see there, too’ and pulled the cart over and got out.”

(13) So- represents ‘Proximal to addressee’

(Two people are playing a game of Go. They are facing each other across the Go board. One says to the other:)

‘machitamae ya. Soko wa Ji ni koso arame, kono watari no Koo o koso’
please wait P that place P Ji P P JD this around P Koo P P

“Please wait, that place is “Ji”(draw), isn’t it? (Let’s count) the “Koo”(conflicting) area around here”

(Genji Monogatari, Utsusemi vol.2)

Note, however, that ka(a)- also indicated the location of close proximity to the addressee in the above literary works. It means that the location of the addressee was not grammaticalized into semantic features, yet in the Japanese demonstrative system of that time and thus the addressee’s realm, it had been integrated into the speaker-anchoring non-proximal area (Li 2002: 216-217, Okazaki 2010: 95).

(14) Ka- represents ‘Proximal to addressee’

(Father found a paper with poetry at his daughter’s feet.)

kore wa ikanaru mono domo zo to mikokoro odorokarete,
this one P what thing pl P P mind surprised and

“He was surprised and thought ‘what happened with this?’”

‘kare wa tare ga zo. keshiki kotonaru mono no sama kana. tamae. …’
that one P who P P look unusual thing P appearance P give me

“Who’s is it? It seems strange. Give it to me.”

(Genji Monogatari, Hahakigi vol.2)

As the above examples show, in this era, both middle and distal terms were used when the referent was relatively far from the speaker and the addressee was close. This fact suggests that the Japanese demonstrative system between the 10th and 11th centuries was one the ways of “semanticizing” the position of the addressee (Meira 2003: 10) as well as the synchrony of the Tiriyó demonstrative system.

4.3 Ca. 16th century
It is widely accepted that the ‘Proximal to addressee’ usage had been incorporated into the Japanese demonstrative system around the 16th century (Okazaki 2010 etc.). To illustrate the process, the war chronicle, *Heike Monogatari* (The Tale of the Heike) is often cited as an example. It was compiled in the 13th century and for centuries afterward various differing versions were created. They provide good for examining the semantic changes in demonstrative terms. For instance, according to Okazaki (2010: 99), the referent that the addressee is holding is referred to by *are* in the 14th century version (as in (15)), whereas it was referred to by *sore* in the 16th century version (as in (16)).

(15) Reference to the addressee’s realm in the 14th century

Futokoro yori shiroi nuno ni tsutsumudaru sharekoobe o hitotsu toridasu.

bosom P white cloth P wrapped skull P one take out

“(He) takes out a skull wrapped in white cloth from his bosom.”

Hyooenosuke ‘are wa ika ni’ to notamaeba,
Hyooenosuke that one P what P P when say

“When Hyooenosuke said ‘What is that?’”

‘kore koso wadono no chichi, ko-Samanokoono-tono no koobe yo’
this one P you P father dead-Samanokoono-Mr. P skull P

“(He said) ‘this one is the skull of your dead father, Mr. Samanokoono.’”

(16) Reference to the addressee’s realm in the 16th century

shiroi nuno de tsutsunda sharekoobe o hitotsu toridaitareba,
white cloth P wrapped skull P one when took out

“When (he) took out a skull wrapped in white cloth,”

Yorimoto sore wa nan zo to towaruruni,
Yorimoto that one P what P P as ask

“Yoritomo asked ‘What is that?’”

kore koso onmi no chichi Samanokoono-tono no koobe degozare.
this one P you P father Samanokoono-Mr. P head JD

“(He said) ‘this one is the skull of your father, Mr. Samanokoono.’”

Such examples show that between the 14th and 16th centuries, the exclusion of the *ka(a)*- term had begun in reference to the addressee’s realm and the ‘Proximal to addressee’ usage had been grammaticalized gradually as the semantic feature of *so-* . Along with this semanticization, the reference domain of *ka(a)*- had become narrower, and finally it became reduced to the ‘Distal’ area, except for the addressee’s realm.
4.4 From the addressee’s attention to the addressee’s realm

In this section, I will outline so-’s semantic changes from the 8th to the 16th centuries. The only remaining issue is the characterization of the exophoric use of so-, which has been observed since the 8th century. As indicated above, the Japanese demonstrative system in the 8th century has been described as a two-term system, where ko- remarked perceptible entities in the surrounding situation and so- conducted indirect references through linguistic discourse or shared knowledge. However, this description does not appear to adequately account for the exophoric use of so-, as in (8d). As noted above, Li (2002: 163-166) posited that the speaker selected so- to refer to visible entities when its distance led him/her to recognize it as almost imperceptible. But this account based on the distance of the referent is not consistent with the definition of ko-, which referred all perceptible entities in the speech situation nor with the account that the derivation of ka(a)- led to the emergence of spatial information in the Japanese demonstrative system. Instead of distance, then, what did determine the speaker’s choice of indirect reference to the physical entity through use of the term so-? This question remains unanswered mainly because of the limitation of examples of the exophoric use of so- in 8th-century literature.

In this paper, I will posit that the original exophoric use of so-, which has been observed continually since the 8th century, represented the concept of the ‘presence of addressee’s attention’, which was defined in section 3.3. This being so, we can provide an alternative account for the speaker’s use of so- to refer to physical entities in the 8th century: the speaker selected so- when (s)he recognized that its referent already had the addressee’s attention at the time of speech in order to make his/her commitment to the referent weaker and signal that the addressee was very close to the solution of his/her reference. The empirical evidence for my claim is the fact that so- at that time had the anaphoric function as its main usage. As discussed in section 3.2, the anaphoric function is considered to be a manifestation of the ‘presence of addressee’s attention’ in linguistic discourse.

If this is correct, why did the poetry anthologies written around the 8th century provide very few examples of the exophoric use of so-, even though they are frequently observed in current natural discourse data? It may be presumed that so- which is used in relation to the status of the addressee’s attention, was much more conversation-oriented than other demonstrative terms and thus seldom appeared in written texts. As shown in section 3.1, the choice of so- requires the speaker to monitor whether or not the addressee’s visual attention is already directed to the intended referent. Therefore, compared to, for instance ka(a)-, which is selected based on the spatial location of the referent, so- seems more grounded in interactional factors and requires the physical presence of the addressee. The fact that the exophoric use of so- began to increase in literary works from the 10th century onward may support my claim: unlike poems, these literary works include face-to-face conversational situations, although they are not natural. All exophoric use of so- appeared in the interlocutors’ utterances (typically marked with quotation marks) in situations such as (13) and never appeared in monologues. The reason why the exophoric so- was extremely limited in poetry

12 Around that time, the recognitional use had disappeared from so- and had been incorporated into ka(a)-.
antilogies is assumed to be due to the monologic nature of the poetry.

5. Conclusion
To summarize, this paper has analyzed the meaning of the Japanese middle term so- and its diachronic change based on the natural conversation data as well as historical resources. From this analysis, the following three conclusions can be drawn:

(i) The Japanese middle term so- has two meanings: one indicates the proximity to the addressee, the other represents the ‘presence of addressee’s attention’ and refers to the object that already has the addressee’s visual attention.
(ii) The concept of the addressee’s attentional contrast can be extended from the physical world to linguistic discourse. Here, so- serves an anaphoric function and refers to the linguistic elements that have already introduced into the discourse.
(iii) The prototypical meaning of so- is the ‘presence of addressee’s attention’, which has been observed since the 8th century. The spatial information related to the addressee is considered to be the derivative meaning of it.

This paper argued that two semantic features of so-, the ‘presence of addressee’s attention’ and ‘Proximal to the addressee,’ have diachronic continuity. However, they are two distinct features synchronically: in the references based on the existence of the addressee’s attention, the joint attention is established by gaze following, whereas in the references to objects close to the addressee, the speaker remarks the spatial position of the referent by deictic projection, where the anchor of the reference is projected onto the addressee. Therefore, this paper claims that the current Japanese demonstrative system is a three-term system with four semantic features.

As mentioned in section 2.3, some recent studies of demonstratives usage in individual languages have challenged the presumed primariness of spatial distinction in traditional analyses and introduced more interactional functions of demonstratives, such as the status of the addressee’s attention. Diessel (2006) focused on the importance of the communicative function of demonstratives to establish joint attention as well as the space demarcation, and provided theoretical grounding for the above studies. Their analysis suggests that spatial information may be secondary information in the demonstrative system, which has been considered as a typical member of spatial deixis. This paper has shown the diachronic process in which the Japanese demonstrative so- derived spatial information from the concept related to the addressee’s attention. It may provide empirical data to support the analyses of the above studies.

References


Hattori, Shiro (1968) Kore, sore, are to this, that. In: *Eigo kiso goi no kenkyu.* Tokyo: English language education council.

Hayashi, Toru (1989) Toruko-go no susume 3 ‘kore, sore, are’ are-kore. (An invitation to Turkish 3: this and that about ‘kore, sore, are’. *Gengo* 18 (1): 96-101.


