Projection and grammar: notes on the ‘action-projecting’ use of the distal demonstrative *are* in Japanese

Makoto Hayashi

Department of East Asian Languages and Cultures, University of Illinois at Urbana-Champaign, 2090 Foreign Languages Building, 707 South Mathews Avenue, Urbana, IL 61801, USA

Received 26 September 2003; received in revised form 9 May 2004; accepted 28 May 2004

Abstract

Projection refers to the feature of human conduct that prefigures possible trajectories of how an action (or a sequence of actions) might develop in the next moment, and which thereby allows interactants to negotiate and accomplish coordinated action in the subsequent course of interaction. The present study explores the mechanism of projection with a focus on the relationship between projection and grammar. To that end, this study examines a particular turn-constructional practice involving the ‘action-projecting’ use of the distal demonstrative pronoun *are* (‘that one; that thing’) in Japanese. The point of departure for the study is the observation made by previous research that, due to some typological features of Japanese grammar, the projection of unfolding turn-shape and action-type is achieved relatively late in the course of a turn in Japanese. Through a detailed examination of relevant data from naturally-occurring Japanese conversations among peers as well as from TV talk shows, this study shows that Japanese speakers can nonetheless achieve early projection of forthcoming action through the use of the turn-constructional practice involving ‘action-projecting’ *are* and thereby compensate for delayed projectability. Issues arising from the analytic section of this study include the question of whether and how various projection techniques are grammaticized into grammatical constructions in different languages.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Projection; Grammar; Conversation Analysis; Demonstrative; Turn construction; Japanese
1. Introduction

The present study explores the relationship between grammar and ‘projection.’ By projection, I mean the property of human action by which an individual action (or a part of it) foreshadows what comes next in the temporal unfolding of interaction. Projection allows interactants to systematically recognize what action is being performed currently and what range of subsequent actions will be made relevant next before the entire course of the current action has been fully disclosed. It thus provides a central resource for interactants to coordinate their conduct vis-à-vis one another’s in prospect of what is happening now and what is going to happen next, and thereby negotiate joint courses of action-in-progress. Therefore, investigating mechanisms by which projection is accomplished contributes to our understanding of crucial aspects of how humans organize coordinated behavior in social interaction.

The phenomenon of projection, i.e., knowing what someone else is going to say and do, involves a wide range of multidisciplinary interests (cf. Levinson, 2003). Understanding the mechanism of ‘mind-reading’ in everyday cognition has clear relevance to the theory of mind, whereas uncovering the resources for ‘look-ahead’ in action for achieving behavioral coordination contributes to the theory of action and interaction. That projection may be based on culturally shared behavioral expectations speaks to the relevance of projection to anthropological inquiry. The role of language in projection and the role of projection in linguistic communication suggest that understanding projection can be fundamental to understanding the foundations of human language and perhaps its evolution. Exploring how projection works thus constitutes a significant area of interdisciplinary research. This study aims to contribute to our further understanding of this important and intriguing phenomenon by focusing its investigation on the interrelationship between projection and grammatical structure.

Whereas a variety of features of human conduct are relevant to the mechanism of projection (see the discussion in Section 2.1), with regard to linguistic conduct in talk-in-interaction, one key resource for the projectability of what an unfolding utterance will eventually look like, and what action it will eventually implement, is the grammatical structure of the language in which the interaction is conducted. This is so because grammar provides interactants with shared ways of organizing the ordering of elements in the unfolding course of a turn-in-progress. It is at least plausible, then, that typologically divergent grammatical resources in different languages might provide varying ways for their speakers to organize the projection of the eventual shapes of unfolding turns. In fact, recent studies (e.g., Fox et al., 1996; Tanaka, 1999, 2000, 2001a,b; Hayashi, 1999, 2003a, 2004) have suggested that, due to the postpositional, predicate-final structure of Japanese grammar, the projection of the type of turn being produced (and the type of action being performed by that turn) tends to be made relatively late in the developmental course of a turn-in-progress in Japanese, compared to how the projection of turn-shape is typically achieved in English (cf. Section 2.2). The present study builds on this line of research, and aims to contribute to a further understanding of the relationship between grammar and projection by closely examining one particular turn-constructional format recurrently observed in Japanese conversation. Through an explication of sequential and interactional motivations for the use of this turn-format in
specific interactional contexts, I will argue that the turn-design in question serves as a means to counteract the above-mentioned ‘delayed projectability’ (Tanaka, 1999) of turn-type/action-type in Japanese.

The turn-format to be explored in this study involves what I call the ‘action-projecting’ use of the distal demonstrative pronoun are (‘that one’ or ‘that thing’). Fragment (1) provides an example of this turn-format. The demonstrative in question appears in line 1 and is in bold face (see appendix for the notational conventions used in the transcripts).

(1) [TG 10] (Akira, an employee at a gas company, is explaining to his friends that the material used for underground gas pipes is changing from metal to plastic.)

1  →  Akira: sono: saikin  are  na n desu yo
       uhm  recently  that. one  CP  N  CP  FP
         “Uh: m, recently, (it)’s been are (=that thing).”

2  o'ano::?: (0.7) GAsu kan  aru  ja nai desu ka:.
       uhm  gas  pipe  exist  CP  not  CP  Q
         “Uh::m’ (0.7) You know there are gas pipes, right?”

3  (. ) >are  zenbu ima< purasuchikku  ni  naritsutsu  aru
       that  all  now  plastic  PT  is.becoming  exist
         “(. ) They’ve all been changing to plastic pipes now=”

4  n desu yo.=DONdon.= =TEtsu kara.
       N  CP  FP  MIM(steadily)  metal  from
         “=one  after  another=from  metal.”

While I wish to leave a detailed discussion of the workings of the action-projecting use of are to Section 3, here let me just point out two notable features of the turn-constructional practice exemplified by Akira’s utterance in the fragment above. First, when Akira produces are in line 1, its referent is not specified yet, in the sense that it is not found either in the prior linguistic context (as in the case of anaphoric uses of demonstratives) or in the surrounding physical setting (as in the case of spatial-deictic uses of demonstratives). Rather, the speaker projects a prospective specification of the referent in the subsequent talk.¹ Indeed, after completing the sentential utterance containing the demonstrative in question in line 1, Akira goes on to present a proposition (are zenbu ima purasuchikku ni naritsutsu aru ‘they [=gas pipes] have been all changing to plastic pipes now’ in line 3) which is the referent indexed by are in line 1.

¹ In other words, are works cataphorically. As will be discussed in Section 3.1, however, prototypical demonstratives used for cataphoric reference in Japanese are ‘proximal’ ones, such as kore (‘this one’) and koo (‘like this’) (cf. Shoho, 1981; Kinsui and Takubo, 1992), and in that regard, the use of a ‘distal’ demonstrative for cataphoric reference is rather peculiar.
Second, the demonstrative *are* in line 1 is embedded in a sentence that consists mostly of copulas and particles, which have little lexical semantic content. The string *na n desu yo*, which follows *are*, is an expression that only indicates that some kind of ‘explanation’ is being offered. Together with the demonstrative whose referent has not been specified yet, the sentential utterance in line 1, i.e., *saikin are nan desu yo* (‘Recently, it’s been that thing’), merely adumbrates or foreshadows some sort of explanation to be offered about something that is happening ‘recently,’ while postponing the specification of what that ‘something’ is. Thus, this utterance not only projects the subsequent specification of the referent of the demonstrative *are* as discussed above, but also projects that specification as part of a particular type of action, i.e., explaining-in-progress. Indeed, in lines 3 and 4, the speaker provides the specification of the referent of *are*, which offers the core of the explanation foreshadowed in line 1.

In what follows, I will first review past literature on projection, with a particular focus on the body of research that has suggested delayed projectability of turn-shape/action-type in Japanese due to some typological features of Japanese grammar. I will then present a detailed analysis of how, in specific sequential/interactional environments, the turn-format in question is mobilized as a means to facilitate an early projection of action-type being performed by a turn-in-progress. This analysis shows that, while Japanese grammar may tend to organize turns in such a way as to place action-indicating elements towards the end of unfolding turns, this does not mean that speakers of Japanese have no means to project early in the turn what action is being performed by the current turn. The study concludes with a discussion of implications for future research, especially the importance of exploring the relationship between typologically divergent grammatical resources in different languages and how turn projection is organized so as to achieve social coordination in interactions conducted in those languages.

2. Projection and grammatical structure

This section provides a brief overview of the previous research on projection (mainly in the framework of Conversation Analysis) that has shown that projection is a central resource for coordinated action in interaction (Section 2.1). It then reviews the recent body of work that has argued for ‘delayed projectability’ in Japanese (Section 2.2), on which the exploration of the action-projecting use of the demonstrative *are* in the following sections is based.

2.1. Projection as a key resource for coordinated social action

The ethologist J.M. Cullen once argued that “all social life in animals depends on the coordination of interactions between them” (Cullen, 1972: 101; cited in M. Goodwin, 1990: 1). Arguably, the same can be said about human social life. Indeed, one of the main concerns of the research on human social interaction in the last four decades has been to uncover mechanisms by which interactants accomplish intricate coordination of their conduct in interactions with one another (cf. Kendon, 1967, 1970; Jefferson, 1973, 1983;
Sacks et al., 1974; Goodwin, 1979, 1980, 1981, 1996, 2000a,b, 2003; Goodwin and Goodwin, 1987; Lerner, 1987, 1991, 1996, 2002; Schegloff, 2000; Streeck, 1995; Erickson and Schutz, 1982; Heath, 1986, 1992; among many others). One central resource for the achievement of such coordination is projectability – the feature of human conduct that prefigures possible trajectories of how an action (or a sequence of actions) might develop in the next moment, and which thereby allows interactants to negotiate and accomplish coordinated action in the subsequent course of interaction. Thus, as C. Goodwin (2000a: 149) argues:

The accomplishment of social action requires that not only the party producing an action, but also others present, such as its addressee, be able to systematically recognize the shape and character of what is occurring. Without this it would be impossible for separate parties to recognize in common not only what is happening at the moment, but more crucially, what range of events are being projected as relevant nexts, such that an addressee can build not just another independent action, but instead a relevant coordinated next move to what someone else has just done.

Work in Conversation Analysis and related fields has shown that there are various levels of organization that are involved in projection. On a higher level, there are, for instance, ‘story prefaces’ (Sacks, 1974), such as The funniest thing happened to me today, which projects the speaker’s subsequent action (i.e., the story proper) that is often performed in a large unit of talk (i.e., in a multi-unit turn; cf. also Schegloff, 1982). In addition to the speaker’s own subsequent action, a story preface also projects relevant next actions for the recipients to perform at specific points in the course of storytelling. Thus, a preface like the one above, for example, projects (i) either a go-ahead or blocking response (e.g., What happened? or I know. John told me.) as a relevant next move by the recipients upon completion of the preface, and (ii) once the story proper is launched, it projects laughter at the climax of the story as an appropriate subsequent action by the recipients. There are also other types of prefatory activities in conversation (e.g., pre-request, pre-invitation, pre-announcement, pre-closing, pre-pre, etc.) that project specific subsequent action by the speaker and recipients across turns (cf. Schegloff and Sacks, 1973; Terasaki, 1976; Schegloff, 1980; Levinson, 1983).

Projection does not just operate across turns. Within the boundary of a turn, earlier parts of a turn-in-progress may project subsequent parts and thereby allow the recipients to “systematically recognize the shape and character of what is occurring,” i.e., the shape that the turn will eventually take and the action that it will eventually perform. This type of projection, which is sometimes referred to as ‘turn projection,’ plays a central role in Sacks et al.’s (1974) seminal work on the organization of turn-taking in conversation. According to their account, the structural unfolding of a speaker’s ongoing turn (or turn-constructional unit, TCU) displays a progression toward a possible completion point, where an opportunity is provided for others to start talking and become the next speaker. What is crucial in their account of turn-taking is that the

---

2 For a useful review of the work in a wider range of fields on interactional coordination, including behavioral synchrony and speech rhythm, see Pelose (1987).
interactants orient to the projection of possible turn completion in advance of the actual completion point and organize their conduct vis-à-vis that prospect. Thus, for instance, a current non-speaker intending to be the next speaker often starts to talk slightly before the actual completion of the current turn in anticipation of an imminent transition-relevance place, and this results in common occurrences of what Jefferson (1983) calls ‘terminal overlap.’ The current speaker may also act in anticipation of the imminent transition-relevance place to interdict such potential incomings by current non-speakers. The ‘rush-through’ (Schegloff, 1982) is one such practice—the speaker speeds up the pace of the talk towards the end of the current TCU and ‘rushes through’ the possible completion point so as not to give others an opportunity to start up.

Turn-taking is not the only concern for which the participants draw on turn projection. For instance, in Goodwin and Goodwin’s (1987) analysis of the processes of two interactants’ simultaneous engagement in assessment activities, it is demonstrated that a recipient of an assessment utterance (i.e., an utterance devoted to making an assessment of some object, person, event, etc.) orients in fine detail to the projection of the unfolding structure of the utterance as a resource to join in the ongoing assessment activity before the speaker’s actual production of an assessment segment (e.g., an adjective that carries the core of the assessment). It has also been shown that turn projection provides a major resource for the achievement of ‘anticipatory completion,’ i.e., a recipient’s completing the speaker’s utterance-in-progress (cf. Lerner, 1991, 1996; Tanaka, 1999; Hayashi, 2003a). Thus, turn projection plays a critical role in the organization of concurrent operation on the ongoing talk by multiple interactants within the boundaries of a TCU.

A large part of coordinated action in interaction, therefore, rests on the possibility of projection at various levels. What, then, furnishes human action with such a versatile resource? A number of studies have been conducted to explore this question, and have shown that a wide array of organizations of human conduct are involved in projectability, ranging from segment-level phonetics (e.g., Jefferson, 1974; Local and Kelly, 1986; Clark and Fox Tree, 2002), prosody (e.g., Local, 1992; Local et al., 1986; Auer, 1996; Couper-Kuhlen, 1996; Selting, 1996, 2000; Fox, 2001; Local and Walker, 2002), turn-design (e.g., Schegloff, 1968, 1980; Sacks, 1974; Terasaki, 1976; Levinson, 1983; Goodwin and Goodwin, 1987; Lerner, 1991; Ford, 2001) to the overall organization of a whole conversation (e.g., Schegloff and Sacks, 1973), as well as non-vocal conduct including gesture, gaze and posture (e.g., Goodwin, 1981; Schegloff, 1984; Heath, 1986; Streeck and Hartge, 1992; Streeck, 1995). While this body of research has provided many valuable insights into how projection is accomplished, one weakness is that most of this research is based on the examination of English language data and, to a lesser extent, German language data.4 If we are to understand projection as a fundamental resource for human

---


4 Exceptions are Streeck and Hartge (1992) [Ilocano], Streeck (1995) [Thai], Steensig (2001) [Danish, Turkish], Levinson (2003) [Guuuy Yimithirr, Tzeltal], and Enfield (2003) [Lao]. Note, however, that these studies, except for Steensig (2001), focus on non-vocal conduct and its relation to projection, and pay little (if any) attention to the linguistic structure of the language in which the interaction they examine is conducted.
interaction, we ought to examine how projection is achieved in interactions conducted in other languages as well. This is so because, among the various semiotic modalities that make up human social action, language plays a central (if not the central) role in human interaction, and given the key contribution that grammar makes in organizing the ordering of elements in temporally unfolding linguistic action, it is conceivable that differences of grammatical structure in disparate languages might affect the way projection is achieved in interaction. In other words, as Auer (2002) puts it, “the grammars of human languages provide interlocutors with sedimentated and shared ways of organizing” projections; therefore, it is worth investigating how the ways in which projections are organized might vary with divergent grammatical structures in typologically different languages. It is along this line of thinking that the present study explores the relationship between grammar and projection in Japanese talk-in-interaction; it examines some aspects of how the grammar of Japanese provides its speakers with ‘sedimentated and shared ways’ of organizing projections.

This study builds on, and contributes to, an emerging body of research into the relationship between grammar and projection in Japanese (cf. Fox et al., 1996; Tanaka, 1999, 2000, 2001a,b; Hayashi, 1999, 2003a, 2004). The point of departure for this study is the observation made by Fox et al. (1996) and Tanaka (1999) that, due to several typological features of Japanese grammar, turns in Japanese talk-in-interaction (clausal and sentential ones, in particular) are typically structured in such a way that the type of turn being produced and the action being performed by that turn are made available relatively late in the progress of an emerging turn (in comparison to what typically happens with turns in English talk-in-interaction). The next subsection discusses the notion of ‘delayed projectability’ in Japanese, which will provide a basis for my analysis of the ‘action-projecting’ use of the demonstrative pronoun are as a turn-constructional practice designed to facilitate early projection of action-in-progress.

2.2. ‘Delayed projectability’ in Japanese

Two recent cross-linguistic studies of interactional phenomena in Japanese and English talk-in-interaction—Fox et al.’s (1996) study on same-turn self-repair and Tanaka’s (1999) study on turn-taking—present converging observations regarding the difference in how turn projection is typically achieved in the two languages. They suggest that, while turns in English tend to be structured in such a way that the actions being performed by the turns-in-progress are projected relatively early in their progress, turns in Japanese tend not to facilitate such an early projection of turn shape and action type from the beginning of turns. Both studies argue that this difference in turn projection stems largely from differences in the grammatical practices typically employed for turn construction in the two languages. For instance, English syntax requires the presence of core constituents in most clauses and their order is fairly rigidly SV(O). The grammatical subject in conversational English is typically short in terms of its phonological size (often a pronoun), and is regularly followed by a verb complex (e.g., auxiliary + verb, etc.), which often carries core information about the type of action being implemented by the unfolding TCU. This then leads to the situation in which, at least as far as clausal or sentential TCUs are concerned, elements that present a fair amount of information about
the kinds of actions being implemented by the ongoing turns are placed towards the beginning of turns rather consistently—which suggests that turn beginnings (or TCU beginnings) play an important role in turn projection in English. Indeed, Schegloff (1987) argues for this point in the following statement:

One important feature of turn construction . . . and the units that turn construction employs (e.g. lexical, phrasal, clausal, sentential constructions) is that they project, from their beginnings, aspects of their planned shape and type. . . . For example, question projection: obviously enough, starting a turn with a ‘wh-word’, though it doesn’t necessarily entail that ‘a question’ is going to be constructed, powerfully projects that possibility for the turn’s development, with potential consequent constraints on next turn. Or: beginnings can project ‘quotation formats’; starting a turn with ‘He says’ projects the strong possibility of quotation as the type of turn to be developed . . . Or: a beginning like ‘I don’t think’ can project, in certain sequential environments, ‘disagreement’ as a turn type for its turn. . . . Again: turn beginnings are important because they are an important place for turn projection, and, given the importance of turn projection to turn-taking, they are important structural places in conversation. (pp. 71–72; emphasis added)

Now, let us consider grammatical practices typically employed for turn construction in conversational Japanese. First, Japanese is a so-called predicate-final language in which a verb or other predicate tends to be placed at the end of a clausal or sentential TCU. While this predicate-final orientation is very rigid both structurally (cf. Kuno, 1973) and interactionally (cf. Tanaka, 1999), the ordering of other constituents that precede the predicate within a TCU exhibits a fair degree of flexibility. Thus, Martin (1975) states that:

. . . so long as you put the predicate . . . at the end, where it belongs in a well-planned sentence, you are free to present each of the build-up phrases early or late as you see fit. (p. 35)

Another notable feature of conversational Japanese is the prevalence of so-called ‘ellipsis,’ i.e., unexpressed syntactic constituents. Unlike English, which, in most cases, requires the presence of overtly expressed core arguments, Japanese allows its speakers to leave syntactic elements unexpressed when those elements are identifiable from contextual information, or their mention is unimportant, to be avoided, etc.5

Now, if word order before the TCU-final predicate is quite flexible, and if many syntactic elements (including core arguments) can be, and often are, left unexpressed, then it is conceivable that the presence of some constituent at the beginning of a turn provides relatively little information as to how the turn will develop structurally in the temporal unfolding of interaction, except (perhaps) that some kind of predicate will eventually be produced towards the end of the TCU-in-progress. In other words, unlike the fairly regular occurrences of the subject–verb combination towards the beginning of clausal and sentential TCUs in English, TCU beginnings in Japanese tend not to present consistent elements that serve to project what is to follow structurally. Furthermore, since verbs,

5 For more general discussions of characteristics of conversational Japanese, see Maynard (1989, Chapter 3), Fox et al. (1996), Tanaka (1999), and Hayashi (2003a, Chapter 2).
auxiliaries, and other predicates, which often carry information about the type of action being implemented by the ongoing TCUs, tend to be placed towards the end of TCUs in Japanese, the indication of action being implemented appears to be more end-oriented in Japanese than beginning-oriented (as Schegloff, 1987, has suggested for English). Tanaka (1999) articulates this point as follows:

English syntax facilitates an early projection (relative to Japanese) of the type of turn being produced, since the social action performed by a turn is typically made available early in the progress of a turn. In other words, the substance of what is being talked about is commonly produced after the turn-shape has already been projected. Roughly the reverse can be said to hold in Japanese. Partially as a result of the predicate-final orientation and postpositional grammar, turns in Japanese are massively structured so that the substance of what is being talked about is articulated before the social action bearing up that substance is made known. Put another way, turns in Japanese do not necessarily project from their beginnings what their ultimate shape and type will be. (p. 141; emphasis in original)

For instance, in Japanese, a simple yes-no question is typically formulated by placing the question particle \textit{ka} (or a variant thereof) at the end of a sentence, as seen in the following example. That is, unlike the subject-auxiliary inversion at the beginning of a yes-no question in English, the grammatical element that encodes the social action performed by a turn like the following is placed at the end of the turn,\(^6\) and therefore, at least from a syntactic point of view, the action import of the turn (i.e., asking a question) is made known only after the ‘substance’ of the turn has been articulated.\(^7\)

\begin{equation}
(2) \text{[TYC 41]}
\end{equation}

\begin{verbatim}
Shoko: kore: (.) yamamoto san ga katta n desu ka:?
this Yamamoto TL SP bought N CP Q
“Did ([(you)]) buy this ([(yourself)], Mr. Yamamoto?”
\end{verbatim}

It should be noted, however, that, by suggesting the possibility of ‘delayed projectability’ in Japanese, neither the previous studies nor I claim that recipients in Japanese conversation are completely unable to make predictions about what the speaker might say next or what action s/he is implementing as the utterance goes on. Obviously, any native speaker of a language has countless experiences with how utterances tend to go in their language. Moreover, prosody, bodily behavior, and the structures of the sequence and activity in which an utterance is embedded all contribute to the projectability of turn-shape and action-type. What Fox et al. (1996) and Tanaka (1999) have suggested, then, is that the grammatical practices of English appear to have built-in design features that would make

\(^6\) If it were not for the sentence-final question particle \textit{ka} in the sentential TCU in (2), the rest of the utterance would be heard as a declarative sentence (at least from a syntactic point of view) which means, "((You)) bought this ((yourself)), Mr. Yamamoto."

\(^7\) There are other ways to formulate yes-no questions, such as mobilizing rising intonation at the end of syntactically declarative sentences, among others. However, there is no systematic way to grammatically encode a yes-no question at the beginning of a sentence in Japanese.
early projections of turn-shape and action-type a more straightforward process than do the grammatical practices of Japanese. In other words, the built-in design features of the grammatical practices of Japanese do not seem to facilitate early projections as a structurally straightforward process.

Against this general background that Japanese grammatical practices tend to lead to relatively delayed projectability of turn-shape and action-type, the present study shows that Japanese speakers are nonetheless not without a means to achieve early projection when such early projection is sequentially and interactionally relevant. One such means is the turn-constructional practice involving the ‘action-projecting’ use of the distal demonstrative *are* discussed briefly in Section 1. In the next section, I present detailed analyses of the sequential and interactional environments in which the turn-format in question is used. Through these analyses, I demonstrate how the turn-format fulfills the sequentially and interactionally motivated need to project forthcoming action-type early in the progress of a turn.

3. The ‘action-projecting’ use of the demonstrative *are* in Japanese

This section begins with a brief overview of demonstrative pronouns in Japanese so as for the readers to appreciate the fact that the particular use of the distal demonstrative pronoun *are*, which is investigated in this study, in fact represents a rather peculiar usage from the perspective of conventional grammatical descriptions of demonstratives in Japanese. I then discuss the design features of the turn-format in question that involves the action-projecting use of the demonstrative *are*. This is followed by a detailed examination of some instances in which the turn-format in question is used by Japanese speakers to project an upcoming action early in the progress of a turn and thereby cope with different interactional contingencies.

3.1. Demonstrative pronouns in Japanese

Unlike the demonstrative paradigm in English, which has a two-way contrast (e.g., *this* vs. *that*), the paradigm of Japanese demonstratives has a three-way contrast, as seen in the following partial list of demonstrative pronouns:

<table>
<thead>
<tr>
<th>proximal (ko-series)</th>
<th>medial (so-series)</th>
<th>distal (a-series)</th>
</tr>
</thead>
<tbody>
<tr>
<td>object/event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kore (‘this one’)</td>
<td>sore (‘that one’)</td>
<td><em>are</em> (‘that one over there’)</td>
</tr>
<tr>
<td>place</td>
<td></td>
<td></td>
</tr>
<tr>
<td>koko (‘this place’)</td>
<td>soko (‘that place’)</td>
<td><em>asoko</em> (‘that place over there’)</td>
</tr>
<tr>
<td>direction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kochira (‘this way’)</td>
<td>sochira (‘that way’)</td>
<td><em>achira</em> (‘that way over there’)</td>
</tr>
<tr>
<td>person</td>
<td></td>
<td></td>
</tr>
<tr>
<td>koitsu (‘this guy’)</td>
<td>soitsu (‘that guy’)</td>
<td><em>aitsu</em> (‘that guy over there’)</td>
</tr>
</tbody>
</table>

While the terms ‘proximal,’ ‘medial,’ and ‘distal’ are useful glosses for the three categories, there are further aspects to the contrast that are not adequately captured by such glosses. According to conventional grammars of Japanese, generally the three categories differ in the following way. In the spatial-deictic use of demonstratives, i.e., reference to an
entity physically present in the speech situation, the ‘proximal’ ko-series is used to refer to something near the speaker, while the ‘medial’ so-series refers to something closer to the hearer. The ‘distal’ a-series indicates something at a distance from both the speaker and the hearer. In the anaphoric use of demonstratives, i.e., reference to something mentioned earlier in the discourse, the ‘medial’ so-series is used as an unmarked form (Martin, 1975: 1067). The use of the ‘proximal’ ko-series for anaphoric reference adds the sense of ‘immediacy’ in the sense that reference is made “as if the object being talked about were visible and were at the speaker’s side” (Kuno, 1973: 290). The ‘distal’ a-series is used for anaphoric reference (i) when shared knowledge about the referent is recognized between the speaker and the hearer, or (ii) when the speaker makes reference in a monologic mode with a strong emotional attachment to the referent (Iwasaki, 2002).8

While demonstratives in Japanese are most commonly used spatial-deictically or anaphorically, the distal demonstrative pronoun are in its ‘action projecting’ usage, as briefly discussed in Section 1, departs from both of these more common usages of demonstratives in that its referent is not locatable in the physical setting of the speech situation or in the preceding discourse. Rather, its deployment points forward to the subsequent discourse as the place to find the referent, and projects a prospective specification of the referent as the relevant next action in the interaction. For the readers’ convenience, I will present a schematic representation of the example of the action-projecting use of are discussed in Section 1:

(3)

Akira: sono: saikin are na n desu yo “Uh:mm, recently, ((it))’s been that thing”

“ano::” (0.7) GASu kan aru ja nai desu ka: “Uh::m” (0.7) You know there are gas pipes, right?”

(.) >are zenbu ima< purasuchikku “(.) They’ve all been changing to

ni naritsutsu aru n desu yo.= plastic pipes now =

=DONdon.=TEtsu kara. =one after another=from metal.”

8 See also Hinds (1973), Kitagawa (1979), Coulmas (1982), and Yoshimoto (1986). For a comprehensive review of the past literature on Japanese demonstratives, see Kinsui and Takubo (1992).
The practice of making forward-looking reference to an object, person, place, etc., that will be specified in the subsequent discourse is called ‘cataphora’ in the literature (cf. Bühler, 1934; Halliday and Hasan, 1976). Curiously enough, however, conventional accounts of cataphora in Japanese in the past literature have maintained that it is the ‘proximal’ *ko*-series that is used for cataphoric reference (cf. Shoho, 1981; Kinsui and Takubo, 1992), rather than the ‘distal’ *a*-series. In fact, except in a recent body of research using naturally-occurring language as its data (cf. Fox et al., 1996; Uemura, 1996; Kitano, 1999; Hamaguchi, 2001; Hosoda, 2002; Hayashi, 2003a,b), the kind of cataphoric reference with the distal demonstrative pronoun *are* explored in the present study has received virtually no attention in the long history of research on demonstratives in Japanese linguistics. It is a goal of this study, then, to provide an empirically grounded account of the hitherto little explored practice of making cataphoric reference with the distal demonstrative pronoun *are*. In the remainder of this section, I focus on one type of the cataphoric use of *are*, i.e., its action-projecting usage, and explore functional motivations for its deployment in specific interactional contexts. The next subsection discusses the particular turn-constructional format in which *are* in its action-projecting use is presented.

3.2. Turn-constructional format for the action-projecting use of *are*

A key feature of the action-projecting use of the distal demonstrative pronoun *are* is that the demonstrative appears in a particular type of turn-constructional format, which I call a ‘semantically light sentence.’ A sentential turn-constructional format is ‘semantically light’ in that it typically consists of elements that carry low referential/lexico-semantic content. Some examples of ‘semantically light sentences’ observed in my data are:

(4) a. *Are* ja nakatta?  “Wasn’t ((it)) *that thing*?”
   b. *Are* deshoo?  “((It))’s *that thing*, right?”
   c. *Are* da/desu yo ne.  “((It))’s *that thing*, isn’t ((it))?”
   d. *Are* chau n.  “Isn’t ((it)) *that thing*?”
   e. *Are* na n da yo, or  “((It))’s *that thing*.”
   *Are* na n desu yo.  indicates that the speaker is providing an account

These sentential turn-constructional units all consist of the distal demonstrative *are* (whose referent has not yet been specified), followed by some type of copula (e.g., *ja, deshoo, da, desu, na*), auxiliaries (e.g., *nakatta*), and/or final particles (e.g., *yo, ne, n*). Note that none of these elements that compose the ‘semantically light sentences’

---

9 This may be due to the fact that, while cataphoric reference with distal demonstratives is interactionally motivated, conventional accounts of demonstratives in Japanese linguistics literature have been based mostly on the examination of intuited sentences or dialogs.
contains much referential/lexico-semantic content. Rather, these TCU-final elements make up common expressions that are used to indicate specific kinds of actions that the speaker is up to. The TCU-final expressions in (4a)–(4d) are used to indicate that the speaker is seeking confirmation/agreement, while the ones in (4e) are used when the speaker engages in some sort of account-giving. Thus, with the cataphoric are projecting a subsequent specification of its referent later in the turn, these sentential TCUs seem to serve as a device to prefigure the type of action that the speaker is up to before providing the details of the content of the turn with which the prefigured action is to be performed. In other words, they allow speakers to give advance notice of an upcoming action early in the developing course of a turn, while postponing a specification/elaboration of the details of the projected action until later in the turn.10

The significance of the practice being described becomes clearer when we consider it against the backdrop of earlier discussions of how turn projection is typically achieved in Japanese. Recall that, according to Tanaka (1999: 141), “turns in Japanese are massively structured so that the substance of what is being talked about is articulated before the social action bearing up that substance is made known.” This can be represented schematically as follows:

A sentential or clausal TCU (which may comprise a whole turn)

[substance of what is being talked about] + [action-indicating elements]

e.g.) Example (2):

kore: (.) yamamoto san ga katta n desu

substance of what is being talked about

ka:?

action-indicating element

“Did ((you)) buy this ((yourself)), Mr. Yamamoto?”

Now, what the speaker accomplishes by deploying a ‘semantically light sentence’ with cataphoric are embedded in it is to move up the action-indicating elements earlier in a turn, while projecting at least another TCU (to be produced subsequently)

---

10 It should be clear by now that, although I speak of the action-projecting use of the distal demonstrative are, the practice that I am describing is not constituted by are alone, but by the whole sentential TCUs in which are is embedded.
in which the referent of the cataphoric *are* is elaborated. To represent this schematically:\(^{11}\)

![Diagram: projects subsequent elaboration]

Thus, against the general tendency of the late placement of action-projecting elements in clausal and sentential TCUs in Japanese, the practice discussed here can be seen as a means to devote the initial TCU of a turn to an early projection of the type of action that the speaker is going to engage in in the turn, while leaving the specification of the ‘substance of what is being talked about’ to later TCUs within the same turn.

Now, why do speakers make efforts to achieve early projection with this turn-constructional practice in the first place? Are there any motivations for them to do so? The following subsection answers these questions by examining some interactional environments in which the ‘semantically light sentences’ containing the demonstrative *are* are deployed by Japanese speakers, and investigating what tasks they are used to accomplish.

### 3.3. Interactional motivations for the action-projecting use of the demonstrative *are*

In this subsection, I explore some interactional motivations for the action-projecting use of the demonstrative *are* in specific interactional/sequential environments. Two such motivations are discussed: (i) making a bid for additional turn-space to deal with preliminaries to performing the projected action, and (ii) soliciting others’ alignment as recipients early in the progress of a turn in the context of (a) competing for speakership or (b) deploying gesture-accompanied action that requires recipients’ gaze. While these are the most common interactional environments in which the action-projecting use of *are* is observed in my data, they do not by any means exhaust all the interactional contingencies that can motivate such usage of *are*. Rather, the following data analyses are meant to be exemplary demonstrations of the complex workings of the deployment of the turn-constructional practice in question in specific interactional contexts.

---

\(^{11}\) As the schematic representation suggests, the referent that *are* in a semantically light sentence points to is often the whole proposition elaborated subsequently, rather than a co-referential nominal. In this sense, the type of cataphoric reference being investigated falls in the category of ‘discourse deixis’ (cf. Levinson, 1983; Himmelmann, 1996).
The sources of the data segments examined in this subsection are tape-recordings of naturally-occurring conversations among adult peers, as well as of talk shows aired on television, in which casual conversations between the hosts and guests take place in front of an audience. The participants are all native speakers of Japanese (with some regional variations).

3.3.1. Making a bid for turn-space to deal with preliminaries to performing the projected action

Let us first examine cases in which a semantically light sentence with cataphoric are in it is deployed when a more-than-minimal account or explanation of something is interactionally relevant. In such an interactional environment, the semantically light sentence serves to project upcoming account-giving at the outset of a turn, thereby inviting the other participants to align themselves as account-recipients, while making a bid for additional turn-space in which to work out the details that are preliminary to performing the projected account-giving. Because of the turn-initial projection of upcoming account-giving, the recipients are given the framework for interpretation with which to hear the talk following the semantically light sentence as related to and preliminary to the projected action.

Fragment (5) below presents a longer stretch of talk of which Fragment (1) is a part. It is taken from a conversation among three male speakers of Japanese, Akira, Fumio, and Yoshi, who are close friends. The turn with which we are concerned is Akira’s in lines 20 through 23. This turn is embedded in a larger Question–Answer sequence initiated by Yoshi in line 1, in which he asks a question about Akira’s job, i.e., Akira ima nani yatten no:. (“What are ((you)) doing now, Akira?”). Yoshi knows that Akira is a chemist who works as a researcher for a gas company named ‘Osaka Gas,’ and therefore, his question is heard as asking about what kind of research Akira does. Following this question, there is some discussion that lasts for 33 lines about other researchers’ research in Akira’s lab (that part of the conversation is omitted below since it is not directly relevant to our discussion). After that lengthy intervening discussion, Akira finally provides an answer (lines 2–3), in which he states that he is doing research on plastics. This answer prompts further sequences of talk (i.e., ‘post-expansions’ of a base adjacency pair – cf. Schegloff, 1990, in press) in the form of a confirmation-request by Yoshi (lines 4ff) and follow-up questions by Fumio (lines 10ff, 12ff), which seem to be prompted by some degree of incongruity between what is expected as part of the business of a gas company and Akira’s answer that his research is on plastics. Consider how these post-expansions lead up to Akira’s deployment of a semantically light sentence in line 20.

---

12 Talk observed in TV talk shows may potentially be affected by various contingencies that do not exist for mundane conversations, such as the presence of a large crowd of people watching the speakers in the studio, etc. However, as shown in the data analysis below, at least as far as the target phenomenon for the present study is concerned, what is observed in the data from those particular TV talk shows used for this study exhibits little difference from what is observed in the mundane conversations examined for the study.
(5) [TG 10-11]

1 Yoshi: Akira ima nani yatten no:
   Akira now what doing FP
   “What are ((you)) doing now, Akira?”
   :
   :
   :
   :
   :
   :
   :
   :
   :
   : ((33 lines omitted where the participants discuss others’ research at Akira’s lab.))
   :
   :
   :

2 Akira: >boku- boku-< boku ga:: ano:: PURasuchikkun no
   I I I SP uhm plastic LK
3 kenkyuu yatterun desu.=[PURasuchikkun]
   research doing CP plastic
   “>I- I-< I’m doing research on uhm plastics.”
   [ ]

4 Yoshi: [ e Akira ] ima PURasuchikkun
   Akira now plastic

5 yatten no.=
   doing FP
   “e Are ((you)) working on plastics now?”

6 Akira: =ee. (0.9) nazeka= 
   yes somehow
   “Yes. (0.9) For some reason.”

7 Yoshi: =oo(h)o [OH!] HEH HEH heh heh [heh heh heh .hhh]=
   oh oh
   “O(h)hhh! HEH HEH heh heh heh heh heh .hhh”
   [ ]
   :

8 Akira: [ee. ] [PURasuchikkun no ]=
   yes plastic LK
   “Yes.” “Plastics”

9 Yoshi: =.[hhh] PURasuchikkun- [hh hh hh hh hh] .HH
   plastic
   “.hhh plastics- hh hh hh hh hh .HH”
   [ ]
   :

10 Pumio: =[ e! ] [doo:: sun no sore.]
   how do FP that
   “e!” “What do ((you)) do with that?”
11 Akira: *purasuchikku:* plastics
“Plastics...”

12 Fumio: [oo]*saka gasu kankee nai no.* Osaka Gas relation not.exist PP
“Is ((it)) unrelated to Osaka Gas?”

13 Akira: *iya! kankee naku mo nai desu kedo [ne:=-maa]= well relation not.exist PT not CP but PP well
“Well, ((it))’s not unrelated, but uhm...”

14 Yoshi: [o:::::]= “Oh:::::”

15 Akira: *=iroiro: purasuchikku tsukatteru toko arimasu kara ne::: various plastic use place exist because PP
“... There are various places where ((we)) use plastics, so...”

16 Yoshi: a honma:. oh really
“Oh really.”

17 Akira: u::: Kigu toka: *°maa sono[::::°] yeah appliance etc. well uhm
“Yeah. Like appliances, °and uh::::m°”

18 Yoshi: [kigu] toka de::.
 appliance etc. in
“Like appliances”

19

(0.4)

20 → Akira: sono: saikin *are* na n desu yo uhm recently that CP N CP PP
“Uh:m, recently, ((it))’s been are (=that thing).”

21 °ano:::° (0.7) *Gasu kan aru ja nai desu ka::: uhm gas pipe exist CP not CP Q
“°Uh::::m° (0.7) You know there are gas pipes, right?”

22 → (. ) >are zenbu ima< purasuchikku ni naritsutsu aru that all now plastic PT is.becoming exist
“(.) They’ve all been changing to plastic pipes now=’”
As discussed above, Akira’s answer in lines 2–3 to Yoshi’s initial question in line 1 prompts further sequences of talk. In lines 4–5, Yoshi asks for confirmation on Akira’s answer (e Akira ima PUrasuchikku yatten no. “Are ((you)) working on plastics now?”), which Akira confirms in the following turn. Then, in line 10, Fumio asks a follow-up question, e doo sun no sore (“What do ((you)) do with that?”). Notice that both Yoshi’s confirmation request in lines 4–5 and Fumio’s follow-up question in line 10 are prefaced by the particle e, which is used to indicate that the speaker treats the prior talk that s/he is responding to as ‘unexpected.’ In addition, Yoshi’s laughter in lines 7 and 9 suggests that there is something incongruous (and humorous) about the fact that a chemist working for a gas company is doing research on plastics. Thus, there are some observable indications of puzzlement on Yoshi’s and Fumio’s part about Akira’s answer in lines 2–3 to Yoshi’s initial question.

Perhaps this puzzlement prompts Fumio’s second question, or even challenge, in line 12, about whether Akira’s research is at all related to his company’s business. To this, Akira defends himself, but only weakly, by claiming in line 13 that his research is not totally unrelated to his company’s business, and also by stating in lines 15 and 17 that there are various gas-related products for which plastic is used, such as parts of appliances. This account by Akira is only met with minimal receipts produced by Yoshi in lines 14, 16, and 18. Fumio does not respond; he does not even nod during the course of Akira’s account. Thus, there is still a sense in which the puzzlement remains, at least on Fumio’s part, as to how Akira’s research is relevant to his company’s business.
Note also that, in line 17, Akira trails off his utterance, i.e., produces the last part of his utterance in a low volume, which is indicated by the two degree signs in the transcript. And when Yoshi receipts this statement of Akira’s by repeating kigu toka (‘like appliances’) in line 18, he looks away and moves on to have a sip of tea for the first time since the beginning of the sequence. These aspects of the participants’ conduct may suggest that the post-expansion sequence initiated by Fumio’s question in line 12 regarding the relationship between Akira’s research and his company’s business is possibly complete at this juncture.

Thus, the 0.4-s silence in line 19 can be characterized as a juncture at which the sequence so far might be closed down, leaving the two recipients, or at least Fumio, not entirely convinced of the relevance of Akira’s research to his company’s business. Akira could leave the situation as it is, and do nothing further to convince his interlocutors. Another option would be to pursue a further explanation, a more substantial one, so that the others can see how his research can be important for the business of the gas company that he works for. Akira takes this second option, and that is when he deploys the ‘action-projecting’ are embedded within a semantically light sentence, i.e., sono: saikin are na n desu yo (“Uhm, recently, it’s been are [=that thing]”) in line 20.

As discussed above, the distal demonstrative pronoun are in this TCU is used cataphorically, i.e., projecting a prospective specification of its referent in the subsequent talk. The string of morphemes at the end of line 20, i.e., n desu yo, is regularly used when a speaker presents an account of something. The combination of these, then, serves the work of adumbrating, early in the ongoing turn, the type of forthcoming action to be performed, i.e., account-giving, while projecting a subsequent TCU or TCUs to be produced, in which the details of the ‘substance’ of the projected account are specified. Through this projection, the recipients are instructed that (i) there will be more talk to be produced after the current TCU, and that (ii) they should interpret what they will hear in the subsequent TCU(s) as part of the execution of the projected action, i.e., account-giving. This is a particularly apt device for Akira to mobilize at the interactional juncture in which he finds himself at line 20, as the prior sequence (in which he offered initial accounts of the relationship between his research and his company’s business) appears to be closing down with a less than optimal result, and a more substantial account might be needed to remedy the situation. By deploying the semantically light sentence in line 20, then, Akira can immediately display to his recipients what he is up to, i.e., a further attempt at account-giving, while making a bid for additional turn space to work on the production of a more substantial account.

Now, note that the second TCU of Akira’s turn in line 21, i.e., ‘ano::: (0.7) GASu kan aru ja nai desu ka: (“‘Uhm:::m’ (0.7) You know there are gas pipes, right?””), is not the account projected by the initial TCU in line 20. Rather, this second TCU is an understanding-check performed as a preliminary to the execution of the projected account. That this understanding-check is done as a preliminary to the projected account becomes clear in the third TCU of Akira’s turn in lines 22–23. After receiving a non-vocal acknowledgment, i.e., a nod, from Fumio to the understanding-check during the micropause at the beginning of line 22, and thereby establishing a mutual recognition of a new referent, i.e., gasu kan (‘gas pipes’), introduced in the talk, Akira refers anaphorically to it and goes on to build the projected account on it, i.e., >are zenbu
ima< purasuchikku ni naritsutsu aru n desu yo (‘They’ve all been changing to plastic pipes now.’). Note that Akira uses the same sentence-final expression n desu yo in line 23 (followed by two increments) just as in line 20, and thereby indicates that this third TCU embodies the account adumbrated by the account-projecting initial TCU in line 20.

The way the initial TCU in line 20 works somewhat resembles what Schegloff (1980) calls ‘pre-pre’ in English. According to Schegloff (1980), a ‘pre-pre’ like Can I ask you a question? projects an upcoming action, i.e., asking a question, without providing the details of what the question is about. What typically follows a ‘pre-pre,’ however, is not the projected question itself, but rather a preliminary to it. In other words, by projecting a forthcoming action, a ‘pre-pre’ secures some turn-space for the speaker to deal with preliminaries that need to be cleared before the projected action itself is launched. A similar operation is observed with the account-projecting TCU in line 20 above. With the turn-initial projection of upcoming account-giving, the recipients are given the framework for interpretation, i.e., that they are to hear the talk that follows the initial TCU under the auspices of the projected account-giving. This is what allows the speaker to devote additional turn-space to doing something other than the projected action itself, without being heard as doing some random thing that is unrelated to the sequentially/interactionally relevant action.13 Thus, an early projection of upcoming action-type through a semantically light sentence in line 20 is crucial for the speaker’s execution, and the recipients’ sequentially relevant understanding, of the preliminary action in the second TCU in line 21. This suggests that the practice being discussed here is well suited to the kinds of interactional environments in which a more-than-minimal account is called for.

Another instance of the ‘pre-pre’-like operation of a semantically light sentence is observed in Fragment (6) below, taken from a talk show aired on television. In this segment, one of the hosts of the show talks with the members of a pop music band who just made their debut. The host’s comment in lines 1–3 is preceded by a confession made by the band’s leader that the band’s members spend an unusually high amount of money on their costumes (which make them look like animated characters in popular video games). The host’s rather blunt comment in lines 1–3 is received by the audience with boisterous laughter (line 4). In overlap with this laughter, the band’s leader initiates his response to the host’s comment, and deploys a semantically light sentence, i.e., are na n desu yo (‘(It)’s are [=that thing]’), as his first TCU (line 5), thereby projecting an upcoming account of something to be specified in the subsequent TCUs in the turn. Consider how he devotes the following two TCUs (lines 7–8 and lines 10–12) to providing some background information before producing the projected account in lines 14–15. (Note: the Oricon Chart in line 7 refers to a published ranking of music albums, singles, etc., in Japan that is comparable to the Billboard music charts in the USA.)

13 The workings of the semantically light sentence discussed here also resemble what Goodwin (1996) analyzes as ‘prospective indexicals,’ i.e., linguistic expressions whose exact sense is not yet available to the recipients at the time of their production but is instead something that has to be discovered subsequently as the interaction proceeds.
(6) [TV Talk Show 1]

1  Host: ganbara n to kore::: (.) kanzenni work hard not if this completely

2  akaji ni natteshima:::masu in the red go into

3  [(yo). FP

“If ((you)) don’t work hard, you’ll go into the red completely, I’m telling you.”

4  Audience: [((laughter [ laughter laughter laughter ))] [ ]

5  ⇒ Leader: [sore ga ne? [ano uhm are that SP FP uhm CP N CP FP

   na n desu yo]: ::

   “But y’know? Uhmm ((it))’s are (=that thing).”

6  Host: [u : : : : n.]

   “Uh huh”

7  Leader: bokura mo:::, orikon chaato toka ki ni naru we also Oricon Chart etc. worry about

8  ja [nai su ka de]byuu kyoku [ga::: CP not CP Q debut song SP

   “You know, we worry about, like, the Oricon Chart, how ((our)) debut single ((fores in it)).”

9  Host: [u:::n] [u::n] [u::n] u::n.

   “Uh huh” “Uh huh” “Uh huh, uh huh”
Just as in the previous instance, *are na n desu yo* in line 5 serves to adumbrate a forthcoming action-type, i.e., account-giving, at the outset of the turn, while projecting the production of subsequent TCUs in which the details of the ‘substance’ of the projected account are specified. Through this projection, the recipients not only recognize that a multi-unit turn is in progress, but they are also instructed as to how they should interpret what they will hear in the subsequent TCUs, i.e., not as any random remarks, but as preliminaries leading up to the projected action. Thus, in lines 7–8 and lines 10–12, the band’s leader produces two TCUs in which he provides some background information...
leading up to the projected account, and the host aligns with this move by producing a
series of acknowledgment tokens in lines 9 and 13 without attempting to initiate a full turn.
Then, when the band’s leader produces the fourth TCU of his turn in lines 14–15, he
recycles the action-indicating grammatical ending that he used in the semantically light
sentence in line 5, i.e., n desu yo, thereby indicating that this TCU provides the account
projected by the initial TCU in line 5. The host displays his understanding of the
completion of the account projected earlier by moving from a display of passive recipiency
through acknowledgment tokens to active speakership with an assessment in line 16 (cf.
Goodwin, 1986a).

Thus, in the face of a blunt remark by the host suggesting the possibility of a financial
failure in the future (which appears to be heard as a tease, at least by the audience, who
responds with boisterous laughter), the band’s leader launches into a more-than-minimal
account of the situation that they are in, which embodies a counter-argument to the host’s
suggestion, i.e., that the band is so successful that their debut single is ranked 10th in the
nationwide Oricon Chart. Here again, an early projection of upcoming action-type through
a semantically light sentence with are in it serves as an apt device for securing turn-space to
provide the details of a more-than-minimal account, while guiding the recipients’ inter-
pretation of those details in a sequentially relevant manner.

The next subsection discusses a different kind of interactional contingency that
motivates the use of a semantically light sentence with cataphoric are embedded in it.

3.3.2. Soliciting others’ alignment as recipients early in the progress of a turn

Another type of motivation for the use of semantically light sentences is found in the
interactional environments in which the speaker seeks others’ alignment as recipients early
in the progress of a turn for various reasons. For example, when two speakers are
competing for the attention of a third party to establish sequential implicativeness
(Schegloff and Sacks, 1973) of their talk, one means to win such a competition is to
display to the third party early in the turn that s/he (i.e., the third party) will need to respond
(e.g., agree/disagree, confirm/disconfirm, etc.) to what the speaker is about to say. In other
words, by showing at the outset of a turn that what will be said subsequently will make the
addressee’s response relevant at its completion, the speaker can motivate the addressee’s
attentive listening in prospect of becoming the next speaker and thereby win the
competition with another potential current speaker for the floor. Given that the grammatical
elements that display the current TCU’s status as a first pair-part (i.e., action that makes the
recipient’s response relevant) tend to come after the production of the substance of a turn’s
content in Japanese (see Section 2.2 above), moving up the deployment of such elements
towards the beginning of a turn with a semantically light sentence and postponing the
specification of the details of the substance of the turn can serve as an apt device to solicit
others’ alignment as recipients early in the progress of a turn.

Fragment (7) presents a case in point. It is taken from the same TV talk show as
Fragment (6), and the following segment occurs just after the two hosts introduced the
guests (two singers arranged to form a duo on a temporary basis) to the audience.
Immediately after one of the guests thanked the hosts for introducing them in line 1, one
of the hosts, Host A, initiates talk addressed to the guests (line 2). While he is still in the
middle of constructing his turn’s talk, the other host, Host B, initiates talk in overlap with
Host A’s turn (line 3). Given that, in this talk show (and in many other ones as well), the introduction of guests is followed by a slot in which the first topic talk is initiated (usually by the hosts), this overlapping talk can be seen as a competition between the two hosts for gaining the floor to introduce the first topic. Consider how Host A deploys a semantically light sentence containing cataphoric are that projects forthcoming confirmation-request (are desu yo ne ‘(It)’s that thing, isn’t ((it))?’), and succeeds in aligning the two guests as recipients of his talk, thereby winning the competition.

(7) [TV Talk Show 2]

1 Guest A: arigatoo gozaimasu:::=
   “Thank you very mu::ch.”

2 → Host A: =maa maa a[no::::::][are de] su yo ne? ((To Guests A & B))
   now now uhm that CP PP PP
   “Now, uh::::::m, ((it))’s are (=that thing), isn’t ((it))?”

3 Host B: [IYA:::::::::::::::::::] ((To Guests A & B))
   well
   “WE::LL WE::LL,”

4 Guest A: e[h:?] ((To Host A))
   “Huh?”

5 Guest B: [eh:]:? ((To Host A))
   “Huh?”

6 Host A: koozyu ii kata suru to mi mo huta mo nai like.this say way do if too.frank

7 no kamo: shirenai desu [kedomo,]=
   N maybe not.know CP but
   “I guess this may be too frank, but”

[ ]
As stated above, when Host B initiates talk in line 3 in overlap with Host A’s talk in line 2 addressed to the guests, the two hosts enter the state of competition for gaining the floor to launch the first topic. During this competitive talk, Host A initiates a semantically light sentence, i.e., are desu yo ne, which consists of cataphoric are and the sentence-final expression (desu yo ne) that is recurrently used when the speaker seeks confirmation about something from the recipients. This turn-format is designed to give advance notice of upcoming confirmation-request, while postponing the specification of the details of what the speaker is seeking confirmation about. The strategic import of deploying this turn-format and addressing it to the third-party addressees in the context of competitive talk with another speaker is as follows. By announcing early in the turn that the addressees’ confirmation will be relevant on completion of the remainder of the turn, the speaker can motivate the addressees to align themselves as recipients of his talk, rather than the other speaker’s, in
prospect of becoming the next speakers. Note then that, on completion of the semantically light sentence in line 2, both guests display their alignment as recipients of Host A’s talk by uttering eh? in lines 4 and 5 and thereby prompting Host A to elaborate on the referent of the demonstrative are. Having established this speaker–recipient relationship with the two guests, Host A legitimately emerges as the sole speaker (lines 6–7) and goes on to provide the specification of the referent of the cataphoric are (lines 12).14

The next fragment shows an additional interactional contingency that motivates a semantically light sentence to be deployed by an emerging speaker to solicit others’ alignment as recipients early in the progress of a turn. Besides making a bid for the floor during simultaneous talk, what seems relevant in this instance is the organization of eye gaze with regard to gesture. That is, given that a gesture must be seen by the addressee(s) in order to be effective as part of communicative conduct, there is a systematic motivation for a speaker who is about to deploy a gesture to secure the addressee’s gaze (cf. Goodwin, 1986b; Streeck, 1988). The following fragment shows how a speaker mobilizes a semantically light sentence not only to gain the floor, but also to secure the addressees’ gaze in time for the deployment of a gesture that accompanies the projected action.

Fragment (8) is taken from an ordinary conversation among three female speakers of Japanese, Asami, Kyoko, and Chika, who are close friends. It is a part of a longer stretch of talk in which Kyoko tells the other two participants about a mutual friend’s wedding, which only Kyoko attended among the three. In lines 1 and 3, Kyoko states that the wedding reception was such a large one that, even though she was seated in the front section of the reception hall, she was still very far away from the bride and groom. Then, consider how Chika deploys a semantically light sentence in line 6 as she emerges as the next speaker responding to (and making a joking comment on) what Kyoko has just stated.

(8) [KG 19]

1 Kyoko: ... moo KAIsha no hito no TSugi gurai ya noNI::,
EMP company LK people LK next about CP though
“...Even though ((I)) was sitting next to ((her)) colleagues at ((her)) company,”

2 (0.7)

3 Kyoko: too:: tte kanji ya (mon).
far QT impression CP FP
((Kyoko brings gaze to Asami; Figure 1))
“((I)) was so:: far:: away ((from her)).”

4 (0.7) ((Kyoko’s gaze remains with Asami; Figure 1))

5 Kyoko: [ (moo) ] ((To Asami))
EMP

14 Unlike in Fragments (5) and (6), the specification of the referent of the cataphoric are in this instance is not marked by the recycling of the same sentence-final expression as that used in the semantically light sentence (desu yo ne in line 2 vs. desu wa na in line 12). Note, however, that the latter expression, i.e., desu wa na, is also used in the context of seeking confirmation, and in that sense, recycling is achieved at the level of action, if not at the level of grammar.
Let us examine the segment above in terms of the organization of gaze among the participants. For the most part of her turn in lines 1–3, Kyoko looks up in mid-air, looking at neither Asami nor Chika. In the middle of line 3, however, she brings her gaze to Asami as seen in Fig. 1, and maintains this gaze orientation throughout the 0.7-s silence in line 4. Thus, at the moment at which Chika initiates her response in line 6 to what Kyoko has just described, her addressed recipient, i.e., Kyoko, is looking away at another participant. It is in this configuration of gaze orientation that Chika produces a semantically light sentence, i.e., aren’t (it) are (=that thing)? Like bird watching, (they) are doing this, right? hh hehh hehh"

“Isn’t (it) are (=that thing)? Like bird watching, (they) are doing this, right? hh hehh hehh"

Let us examine the segment above in terms of the organization of gaze among the participants. For the most part of her turn in lines 1–3, Kyoko looks up in mid-air, looking at neither Asami nor Chika. In the middle of line 3, however, she brings her gaze to Asami as seen in Fig. 1, and maintains this gaze orientation throughout the 0.7-s silence in line 4. Thus, at the moment at which Chika initiates her response in line 6 to what Kyoko has just described, her addressed recipient, i.e., Kyoko, is looking away at another participant. It is in this configuration of gaze orientation that Chika produces a semantically light sentence, i.e., aren’t (it) are (=that thing)?”, which projects a forthcoming confirmation-request to be produced in the subsequent TCU(s). By announcing at the outset of the turn that Kyoko’s confirmation will be relevant on completion of the remainder of the turn, Chika can motivate Kyoko’s alignment and attention as the recipient of Chika’s talk in prospect of the need to respond to Chika’s turn on its completion. Indeed, as Chika produces the semantically light sentence, Kyoko withdraws
from her talk addressed to Asami (line 5) and begins to shift her gaze toward Chika in the process of aligning herself as the recipient of Chika’s talk. This then allows Chika to produce the substance of the confirmation-request in the second TCU of her turn in mutual gaze with the addressed recipient. This is significant because the substance of the confirmation-request, which turns out to be a joke about the size of the wedding reception, is accompanied by a hand gesture depicting a pair of binoculars used by a bird-watcher, as seen in Fig. 2.

Thus, in this instance, the advance notice of the relevance of the addressee’s uptake of a forthcoming confirmation-request given by a semantically light sentence works not only as a floor claimer, but also as a gaze solicitor for achieving mutual gaze with the addressee when the speaker elaborates on the substance of the projected action with a gesture.

In this section, I have explored two types of interactional motivations for mobilizing semantically light sentences to achieve early projections of forthcoming action in specific interactional/sequential environments. The analysis of the fragments examined above showed that, given the general tendency in turn-constructional practices in Japanese that the grammatical elements that indicate the social action performed by a turn are placed relatively late in the turn, the practice of announcing upcoming action early with a semantically light sentence while postponing the specification of its details until subsequent TCUs is well suited for dealing with such interactional contingencies as (i) a need to make a bid for multi-unit turn-space to present a more than minimal account, and (ii) a need to solicit others’ alignment and attention as recipients early in the turn to gain the floor in the environment of competitive talk or to secure the addressee’s gaze in the context of deploying a gesture.

4. Implications and conclusion

Through a close examination of the ‘action-projecting’ use of the distal demonstrative *are* in Japanese, this study has investigated the dynamic interplay of grammar, projection, and the interactional work that gets done under specific sequential/interactional
contingencies. By way of conclusion, I will discuss several implications of the findings of the study and possible directions for future research.

1. The point of departure for the present study was an observation made by Fox et al. (1996) and Tanaka (1999) about the possible relationship between grammar and projection in Japanese and English; that is, that grammatical practices employed for turn construction in Japanese tend to lead to rather ‘delayed’ projection of unfolding turn-shape and action-type, as compared to early projection facilitated by grammatical practices used for turn construction in English. Against this background, the analysis presented in this study demonstrated that Japanese speakers nonetheless have available turn-constructional means to achieve early projection of forthcoming action and thereby counteract the general tendency of delayed projectability. This finding suggests that, while there may be certain important relationships between grammatical structure (especially its typological features) and the way projection is achieved, we should not hasten to regard the relationship as a simple one-to-one correspondence (such as regarding a language with certain grammatical structure as always furnishing early or late projection). Rather, the relationship between grammar and projection is much more complex and needs to be investigated further.

For instance, we can investigate what repertoire of turn-constructional practices speakers of a given language have available to build utterances that facilitate projection. Exploring this can be relevant and important for the students of grammar because, besides general typological features, there is the possibility that these turn-constructional practices designed for projection, or what may be called ‘projection techniques,’ might be grammaticized, to a varying degree, into ‘grammatical constructions’ (in the sense of Goldberg, 1995) in that language, and become regular parts of the grammar of the language. There are, for instance, several grammaticized

---

15 Besides the features of conversational Japanese and English discussed in Section 2.2, there are important differences in recurrent phrase structures between the two languages (and among typologically different languages in general) that are relevant to projection. For example, English has many phrase structures in which an element that implies the presence of another precedes that implied element, such as:

- Determiner (e.g., articles a, the) → Noun
- Preposition → Noun Phrase
- Verb (transitive) → Noun Phrase (direct object)

The presence of a preposition, for instance, implies the presence of a noun phrase ‘governed’ by the preposition, and thereby projects the production of a noun phrase subsequently. Compare this to the corresponding phrase structures in Japanese, which lacks articles, employs postpositions rather than prepositions, and places transitive verbs after direct objects. In structural terms, one could say, for example, that a postposition implies the presence of a noun phrase, but in terms of on-line production, it does not translate into a projection relation as in the case of English prepositional phrases, since a postposition occurs after the noun phrase that it ‘governs.’ (Note also that not all noun phrases are followed by postpositions.)

Given that the structural dependency relations in syntax such as those exemplified above from English provide ‘strong projective force’ (cf. Auer, 2002), it cannot be denied that their presence or absence in the grammars of different languages plays a significant part with respect to how projection is achieved in those languages. Therefore, while I suggest that the relationship between grammar and projection is not a simple one, I believe, along with Fox et al. (1996) and Tanaka (1999), that there could be general tendencies in how projection is achieved according to typological features of different grammatical structures.
projection techniques in English, such as the following discussed by Levinson (2003) (the list is not meant to be exhaustive; the arrow starting at the ‘projecting element’ and ending at the ‘projected element’):

(9) i. left-movement

*Was it* that Mullah they said they were going to arrest __? 

ii. variables

*What I meant to say was that* these issues are not easy.

iii. fixed expressions

*We thought tabs might be being kept on us.*

iv. hierarchy of constituents

*Unless he comes, we are lost.*

In these constructions, the projecting element foreshadows a specific element to be produced later *within the same (sentential) TCU*. Levinson (2003), citing others’ work, argues that the same sorts of projection techniques are found also in *sequential* structure (as opposed to *sentential* structure), in which projection works *across turns* or TCUs. Examples of the first three projection techniques realized across turns follow:

(10) i. left-movement (cf. Geluykens, 1992: 35 on left-dislocation in English)

A: *now (em) the last paragraph-*

B: *yes *

A: *[e:m] I seem to remember it. being different from what’s printed*

ii. variables (cf. Terasaki, 1976: 53 on pre-announcements)

A: You wanna know who I got stoned with a few w(hh)eeks ago?

B: Who?

A: *Mary Carter ‘n her boy(hh)frie(hhh)nd. hh*
iii. fixed expression (generating the predictability of ‘okay’ + ‘bye bye’)

A: O:Kay.=

B: =Okay | Buh buy

A: | Bye bye

Structural parallels can thus be observed between the operation of projection within sentence structure and that across turns or TCUs in sequence structure. There is, of course, a difference between the two types of projection in that the former is based on knowledge about syntactic structure, while the latter is based on knowledge about the sequencing of activities. However, the dividing line between the two is not always easy to draw (‘left-dislocation’ in (10i) is a case in point; cf. Geluykens, 1992). In fact, Auer (2002), basing himself on similar observations of structural parallels between ‘syntactic projection’ and ‘interactional projection,’ goes so far as to claim that “syntax can be seen (among other things) as the historical result of a sedimentation and (partly normative) regularisation of certain interactional projection techniques”. While this claim remains to be substantiated by further empirical studies, it is certainly an intriguing hypothesis to explore and test about the relationship between grammar and interaction. It also provides the present study with a guiding light for possible directions in future research. The present study has documented but one projection technique employed by Japanese speakers that works across TCUs. It may be worthwhile, then, to investigate what range of other projection techniques are mobilized by Japanese speakers in what sorts of interactional environments and to what extent they have (if ever) been sedimentated and grammaticized as part of the grammar of the language. We may be able to have a glimpse of ‘emergent grammar’ (Hopper, 1987, 1988) motivated by the mechanism of projection.

2. Of the four projection techniques cited from Levinson (2003) above, the turn-constructional format examined in the present study resembles (10ii), i.e., the practice of building utterances with ‘variables’ to project a forthcoming element across turns or TCUs. English has a number of formulaic expressions that fall into this category, such as Guess what, I want to tell you something, Can I ask you something?, Listen to this, etc., which have been analyzed in Conversation Analysis under the rubrics of pre’s and pre-pre’s (cf. Terasaki, 1976; Schegloff, 1980; Levinson, 1983) and prospective indexicals (cf. Goodwin, 1996). There is a striking functional similarity between these expressions in English and those with are in Japanese examined in this study. In both English and Japanese, these expressions with ‘variables’ are used to adumbrate upcoming actions while postponing the specification of the details of the projected actions, thereby inviting recipients to prepare to hear what follows within specific frameworks of interpretation. This observed similarity across the two fairly distinct languages prompts us to speculate as to whether this sort of practice is observed in other languages as well, and if so, whether there are some ‘universal’ motivations in human interaction that have led to the availability of such practices in different languages. This, of course, calls for empirical studies based on interactional data in diverse languages.
From a linguistic point of view, one could also explore what kinds of elements are mobilized as ‘variables’ (or ‘dummy terms’; cf. Schegloff, 1980: 107) in different languages for the sort of action projection being discussed, and whether there are any semantic/pragmatic commonalities among such elements. For instance, English employs wh-words (e.g., what, who), general, non-specific terms like something, and demonstratives (mostly this, but sometimes that; cf. Schegloff, 1996: 62). In Japanese, besides the use of the distal demonstrative pronoun are described in this study, the nominalizer koto (which literally means ‘thing’) may be used, as in o-kiki shitai koto ga aru no desu ga (‘There is something I’d like to ask you.’). Prototypical cataphoric devices like proximal demonstratives (e.g., koo shimashoo ‘Let’s do this’) may also be used. If we have enough data from many other languages, perhaps we could specify the range of meanings in which these items fall, and come up with a typology of ‘variables’ used for action projection (e.g., demonstratives, general/indefinite terms, etc.). We might then be able to begin to examine the hypothesis discussed above on the grammatization of projection techniques by exploring the extent to which the expressions containing such ‘variables’ in different languages have been grammatized/formulaicized for the function in question, i.e., action projection.

3. Related to this last point just discussed, the fact that the distal demonstrative pronoun are is used as a ‘variable’ in action projection in Japanese can be pursued further from a slightly different perspective. In Japanese, there is another interactional environment in which the distal demonstrative are is used cataphorically, and that is the environment of ‘word searches.’ In this latter usage, when a speaker encounters difficulty in remembering/articulating the next word due, are is deployed as a placeholder for the searched-for item, while postponing the specification of its referent until later (cf. Hayashi, 2003b; Hayashi and Yoon, 2004). This placeholder use of are in word searches and the action-projecting use described in this report are clearly

---


16 For example, in the following excerpt, are (line 2) is deployed during a word search as a placeholder for the then-unavailable item, which is subsequently specified as ongaku bangumi (‘music programs’) in line 5.
related in that, in both uses, *are* projects a prospective specification of its referent in the subsequent course of talk.

Now, recent studies have reported that demonstratives (especially distal ones) are often used as placeholders during word searches in several typologically unrelated languages, including Korean (Kim and Suh, 2002; Yoon, 2003), Mandarin (Zhao, 2002); Ilocano (Rubino, 1996), Indonesian (Wouk, 2003), Russian (Nichols, 1993), Romani (Matras, 1998), and Maliseet-Passamaquoddy (Ng, 2003). A question then arises as to whether, in these other languages, demonstratives are ever used as ‘variables’ in action projection in ways similar to how *are* is used in Japanese. If not, are there other forms used as ‘variables,’ and if so, what is the relationship between those forms and demonstratives? If demonstratives are indeed used as ‘variables’ in action projection, do they work similarly as in word searches? Are they embedded in ‘semantically light sentences’ like *are* in Japanese? Answering these questions will contribute to our understanding of hitherto little understood aspects of interactional uses of demonstratives from cross-linguistic perspectives (see Hayashi and Yoon, 2004, for a beginning).

Acknowledgements

Earlier versions of this paper were presented at the Conversational Analysis Advanced Studies Institute held at UCLA (29 July–8 August, 2002), and at the Workshop on Projection held at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands (5–7 March, 2003). I am grateful to the participants at both meetings for their constructive criticism and feedback. I am also indebted to Barbara Fox, Andrea Golato, Charles Goodwin, Kyung-eun Yoon, and two anonymous reviewers for their helpful comments at various stages of the development of the paper.

Appendix A.

A.1. Transcript symbols

| [ | The point where overlapping talk starts |
| ] | The point where overlapping talk ends |
| (0.0) | Length of silence in tenths of a second |
| (.) | Micro-pause |
| underlining | Relatively high pitch |
| CAPS | Relatively high volume |
| :: | Lengthened syllable |
| - | Glottal stop self-editing marker |
| = | “Latched” utterances |
| ?/!/ , | Rising/falling/continuing intonation, respectively |
| ! | Animated tone, not necessarily an exclamation |
| ( ) | Unintelligible stretch |
| (word) | Transcriber’s unsure hearings |
(( )) Transcriber’s descriptions of events – e.g., ((sniff))

hh Audible outbreath
.hh Audible inbreath
(hh) Laughter within a word
> < Increase in tempo, as in a rush-through
○ ○ A passage of talk quieter than the surrounding talk

A.2. Abbreviations used in the interlinear gloss

CP Various forms of copula verb be
FP Final particle
MIM Mimetics
O Object particle
Q Question particle
SP Subject particle
TL Title marker
EMP Emphasis marker
LK Nominal linking particle
N Nominalizer
PT Particle
QT Quotative particle
TAG Tag question
TP Topic particle

References

Ford, Cecilia E., Fox, Barbara A., Hellerman, John, 2002. The Phonetic Composition of “No” Turns. Paper given at the International Conference on Conversation Analysis, held at the University of Copenhagen, Denmark.


Local, John, Kelly, John, 1986. Projection and ‘silences’: notes on phonetic and conversational structure. Human Studies 9, 185–204.


Makoto Hayashi is Assistant Professor of Japanese Language and Linguistics at the University of Illinois at Urbana-Champaign. His research interests revolve around the study of how participants in conversation achieve socially coordinated action and how grammar and other semiotic resources are involved in that achievement. He is the author of Joint Utterance Construction in Japanese Conversation (John Benjamins, 2003).