

*The pragmatic nature of the so-called subject marker *ga* in Japanese: evidence from conversation*



TSUYOSHI ONO
UNIVERSITY OF ARIZONA

SANDRA A. THOMPSON
UNIVERSITY OF CALIFORNIA, SANTA BARBARA

RYOKO SUZUKI
KEIO UNIVERSITY

Discourse Studies
Copyright © 2000
SAGE Publications.
(London,
Thousand Oaks, CA
and New Delhi)
Vol 2(1): 55–84.
[1461-4456
(200002) 2:1;
55–84; 009744]

ABSTRACT Since the inception of modern approaches to grammar, Japanese *ga* has been treated as a marker indicating the grammatical relation 'subject'. If this is an accurate characterization of *ga*, then we would expect *ga* to occur to mark a grammatical category consisting of 'A' (transitive subject) and 'S' (intransitive subject) (Comrie, 1978; Dixon, 1979). Our examination of the contexts in which *ga* is actually used in everyday Japanese conversations shows that this expectation is not borne out. Our findings suggest that it is not appropriate to describe *ga* in terms of a grammatical relation such as 'subject', and that *ga* may be much less a grammatical marker than a discourse-pragmatic one.

KEYWORDS: *case, case-marking, conversation, grammar, grammatical relations, Japanese, pragmatics, subject*

1. Introduction

Since the inception of modern approaches to grammar, the issue of grammatical relations in Japanese and their interaction with the postpositions *ga* and *o* has been extensively debated. An influential contributor to this discussion is Shibatani (1977, 1990: ch. 11.3), who explicitly argues that Japanese can be described in terms of 'subjects' and 'objects', with certain tests being appealed to to determine whether a noun phrase (NP) is in fact the subject of its clause. Two such tests identify subject as: (i) the NP that is referred to by an honorific suffix on

the predicate; and (ii) the NP that is co-referential with the reflexive *jibun* (Shibatani, 1977, 1990: 282–6).

On the relationship between the postposition *ga* and the grammatical relation ‘subject’, Shibatani (1977, 1990: 305) suggests that there is a ‘normal correspondence between the nominative *ga* and the syntactic subject’, but that it is ‘disrupted’ in certain well-defined syntactic environments.¹ Thus, while recognizing that *ga*-marked NPs do not have a one-to-one correspondence with NPs that he identifies as subjects by his tests, Shibatani still calls *ga* ‘nominative’ (1977, 1990). Further, he suggests that *ga*-marking itself is one ‘subject property’ in his prototype analysis of subject (Shibatani, 1990: 306).

From this research, a general consensus has emerged among Japanese linguists to the effect that, while there are circumstances in which a *ga*-marked NP is not the subject, the norm is that *ga*’s primary function is to mark subject. Thus we find *ga* almost universally glossed as NOM (‘nominative’) or SUB (‘subject’) in most discussions of Japanese grammar. Examples such as the following, taken verbatim from Shibatani (1990: 258), illustrate this convention:²

(1)(a)

Taroo *ga* *kita*
 NOM³ came

Taroo came.

(b)

Taroo *ga* *hon* *o* *katta*
 NOM book ACC bought

Taroo bought a book.

However, apart from the ‘disruptive’ contexts mentioned by Shibatani, when we turn to everyday conversation, we find several more troubling mismatches between the use of *ga* and the notion of subject: our examination of the contexts in which *ga* is actually used in Japanese conversations shows that *ga* is rare and that it does not seem to mark a grammaticized category that includes A and S in Japanese,⁴ as we would expect it to do if it were indeed a marker of subject; further, *ga*-marked NPs occur mostly in a very limited number of constructions with intransitive predicates with low semantic content. We interpret these findings as calling into question the widespread practice of describing *ga* in terms of a grammatical relation such as subject, and we suggest that *ga* may be much less a grammatical marker than a discourse-pragmatic one.

2. Previous discussions of *ga*

Many discussions of *ga* in Japanese are embedded into comparisons of *ga* with *wa* (often termed a marker of ‘topic’) in constructed examples. One of the best-known such discussions is that of Kuno (1972: 272, 1973a: 38–9), who, following Kuroda’s earlier work (1965), proposed a distinction among three types of *ga*, labeling them ‘descriptive *ga*’, ‘exhaustive listing *ga*’, and ‘objective *ga*’.

'Descriptive *ga*' is the term used for the *ga* which occurs in 'neutral descriptions of actions or temporary states', as in the following example (Kuno 1973a: 38):

- (2)
 Ame ga hutte imasu
 rain GA falling is
 It is raining.

'Exhaustive listing *ga*' is the term used for the *ga* which occurs in sentences expressing 'X (and only X) ...', or 'It is X that ...', such as (Kuno, 1973a: 38):

- (3)
 John ga gakusei desu
 GA student is

which Kuno says means something like '(Of all the people under discussion) John (and only John) is a student' or 'It is John who is a student'. We see later that this distinction helps us to understand some of the functions of *ga* in actual conversations. We return to 'objective *ga*' later.

Kuno is also among the earliest to relate the use of *ga* to 'new' information, that is, information which is unpredictable from the context, thus suggesting that the function of *ga* is not restricted to marking subject. Here is an example (the glosses and the translation are ours):

- (4)
 A: Taroo to Hanako to Natsuko no uchi de dare ga ichiban segataikai ka
 and and of inside in who GA most tall Q
 Who is the tallest among Taroo, Hanako, and Natsuko?
 B: Taroo ga ichiban segataikai
 GA most tall
 Taroo is the tallest. (Kuno 1973b: 209)

Kuno claims that when *ga* is attached to the 'subject NP' (in a 'matrix' clause), it always indicates that that NP is new, i.e. unpredictable (Kuno, 1972: 273). We return to this point later.⁵

Shibatani (1990, 1991) also compares *ga* with *wa* in constructed sentences such as the following (1991: 97–8):

- (5)(a)
 Hi wa noburu
 sun WA rise
 The sun rises.
 (b)
 Hi ga noburu
 sun GA rise
 The sun rises.

Shibatani suggests that in a sentence such as (5)(a), the *wa*-marked NP 'iso-

lates an entity from [the] state of affairs to be known' (1991: 97), while in (5)(b), with the *ga*-marked NP (what Kuno would term a 'descriptive *ga*'), 'a clause is more tightly organized . . . and the subject shows a tighter bondage with the predicate' (1991: 98–9). We appreciate Shibatani's insight about the clause with *ga* being 'more tightly organized', as shown below in our proposal of the function of *ga*. For further discussion, see Shibatani (1990: 262–306).

Aoyama (1983) explicitly rejects the identification of *ga* with 'subject', with which we are in agreement, but analyzes *ga* in terms of an opposition to *wa*, suggesting that *ga* signals 'low focus' while *wa* signals 'high focus'. The assumption of such a structural opposition as a basis for understanding the use of a given form is not motivated by Aoyama, and does not seem justified to us; we do not make such an assumption in our analysis of *ga*.

Tsutsui (1983, 1984) is among the few to have discussed *ga* from the point of view of its omissibility. Tsutsui proposes a number of phonological, syntactic, semantic, and pragmatic conditions for the 'ellipsis' of *ga* based on constructed data, TV scenarios, and grammatical judgment questionnaires. In particular, Tsutsui (1983) suggests that the 'ellipsis' of *ga* is highly unnatural if the NP is the most 'emphatic' element in the clause and that it is more natural when the utterance carries expected or shared information. Some of these claims are clearly compatible with our conversation-based findings. Similarly, noting that others had observed 'case drop' to occur in colloquial speech, Masunaga (1988) suggests that *ga* can be omitted under certain discourse conditions, which she characterizes as 'a certain lack of focus on the pertinent NP' (p. 148). Supporting Kuno's notion of *ga* as having something to do with 'new information', she suggests that when an NP represents 'shared information', *ga* seems to be easy to omit. To illustrate her point, Masunaga gives the following intriguing pair of constructed examples (149–50):⁶

(6)

Ah! basu \emptyset kita
oh bus came
Oh! A bus is coming.

(7)

*Ah! kyukyusha \emptyset kita
oh ambulance came
Oh! An ambulance is coming.

According to Masunaga, (6) can be said without *ga* in a situation in which people are waiting for a bus and one of them sees a bus coming toward them. But (7) is not acceptable without *ga* in a similar situation where one of the people waiting for a bus sees an ambulance. As we later show, our analysis of the use of *ga* based on empirical data confirms Masunaga's intuition.

Several linguists (Shibatani, 1977, 1991; Sugamoto, 1982; Watanabe, 1984; Yanagida, 1985) have discussed what Kuno termed the 'objective *ga*', as seen in sentences such as (8), taken from Kuno (1973a: 85):⁷

(8)

Anata wa nihongo ga wakarimasu ka?
 you WA Japanese GA understand Q
 Do you understand Japanese?

As noted by Sugamoto (1982), the occurrence of *ga* with what appear to be 'direct objects' in transitive clauses is restricted to clauses of very low transitivity (Hopper and Thompson, 1980), involving predicates such as *wakaru* 'understand' (as in (8)) and *dekiru* 'be able to do'. 'Objective *ga*' is problematic for the hypothesis that *ga* marks subjects since such tests as honorification and reflexivization do not identify 'objective *ga*' NPs as subjects; it is generally treated as exceptional. However, we in our approach, where *ga* does not mark subjects, it is not an anomaly.

While these studies have led to a number of important insights regarding the use of *ga* in Japanese, comparing the examples upon which many of the previous claims about *ga* are based to the usage we find in our conversational data leads fairly straightforwardly to the conclusion that the discrepancy between what speakers think they know about their language and what they actually do when they speak can be fairly wide. In constructing Japanese examples, linguists may be influenced by written Japanese (Linell, 1982; Miller, 1995; Miller and Weinert, 1998) (as well as by written and constructed English, by virtue of their training). We return to this issue in our conclusions. In particular, with few exceptions, hardly any linguistic treatments of *ga* have taken notice of its most conspicuous characteristics in spoken Japanese: (a) its rarity; (b) the general restriction of its use to S, rather than A, nominals; (c) the restriction of its use to NPs with a limited set of intransitive predicates; and (d) the extent of its involvement in grammaticized patterns. We return to all of these points later.

Several studies have considered *ga* in the context of actual discourse. Suzuki and Ono (1991) investigate the use of *ga* in a set of 10 elicited narrative retellings of an episode of the video version of the well-known cartoon *Sazaesan*, and suggest that *ga* tends to occur with one-argument predicates and serves to bring the 'information presented in the *ga*-NP to the primary focus of consciousness'. Highly relevant to the present study is their finding that the referent of the *ga*-marked NP is not 'expected in the situation described by the predicate'. Watanabe (1986) compares the information status of NPs with and without a particle such as *ga* in a planned spoken monologue, showing that NPs without a particle tend to be more 'predictable'. And Kondo (1997), based on a large database consisting of newspaper articles, reports that *ga*-marking of nominalized clauses is extremely skewed in terms of grammatical relations.

These discourse-oriented studies point towards an account of *ga* that links its use to pragmatic factors rather than factors involving grammatical relations, which is the hypothesis we argue for in this article.

3. Data and methodology

Our database consisted of 21 audiotaped conversations varying in length from 1.5 minutes to 18 minutes, totaling over 2 hours of talk and more than 180 pages of transcript. The number of participants in the conversations varied from two to five, all of whom were close friends, couples, or families, and all of whom were adults, speaking standard Japanese.⁸ We have both mixed-gender and same-gender conversations.

In order to determine the frequency of *ga* in the data, we developed the following procedure. We first went through our data and identified all the predicates. We then counted the number of *places* where a *ga*-marked NP either could or did occur with those predicates, and whether it appeared to be an A, an S, or an O.

This means that cases of so-called zero anaphora and of unmarked NPs or NPs marked with such particles as *wa* and *mo* were counted as 'places' as long as the addition of some *ga*-marked NP to the predicate, the addition of *ga* to an unmarked NP, or the replacement of particles such as *wa* or *mo* by *ga* yielded a grammatical utterance, however unnatural-sounding the outcome might be in the context. It also means that we counted as 'places' cases where the so-called 'objective *ga*' could or did occur. This means that some predicates were associated with two 'places' due to the possibility of occurring with two *ga*-marked NPs.

We adopted this methodology in order to get a rough idea of how many of those places are actually filled with NPs, how many of those NPs are unmarked or marked by any particle, and how many of them are marked by *ga*. This methodology, however, should not be interpreted as suggesting that we believe that the actual identification of referents is not problematic or that we believe that all predicates invoke grammatical slots that are recognizably either 'filled' or 'empty'. As we have shown elsewhere (Ono and Thompson, 1997), identifying referents often may not be possible since referents may not be *intended* in cases where they are not overtly expressed (so-called zero anaphora).

We estimate that there are about 4500 such places in our data where a *ga*-marked NP either did occur or could have occurred. From these we first took a random sampling of 146 places on which to perform the quantitative analysis presented later (in Section 4); 24 out of these 146 cases are 'objective *ga*'. They are included partly because we believe 'objective *ga*' is not a well-founded category and partly because including it did not change general patterns in our figures. As we show later, we added more cases to these 146 cases, again based on a random sampling, since the first sample yielded rather few cases of *ga*-marked NPs. We note, however, that the examples in this article come not just from the randomly sampled cases but from the entire database.

4. Findings

From the point of view of the predicate–argument relations in our conversational

data, our findings are striking. Here we first discuss the three major findings of our study: *ga* is rare in Japanese conversation; it occurs largely with S NPs and intransitive predicates; and it occurs mostly with a semantically highly constrained set of predicates. We then claim that these findings show that *ga* is accurately characterized as *marking that its NP is to be construed as a participant in the state-of-affairs named by the predicate in pragmatically highly marked situations*. What we mean by this is that *ga* is used in conversation only when it is pragmatically called for to indicate that its NP should be understood as going with the predicate. That is, we claim that *ga* is found in pragmatically highly marked situations where there is something ‘unpredictable’ about the relationship between the *ga*-marked NP and the predicate such that an explicit signaling of that relationship becomes interactionally or cognitively relevant. With bare NPs, that is those without *ga*, the relationship between the NP and the predicate, rather than being signaled explicitly, is a matter of inference. The low frequency of *ga* in Japanese conversation, discussed later, strongly suggests that most of the time making this inference is not problematic for conversationalists.

As we show, the *ga*-marked NP is typically an intransitive ‘S’; it is very rarely an ‘O’ and even more rarely an ‘A’. Furthermore, it tends to occur with predicates of certain semantic types. What we are interested in, then, is showing (a) that there is no evidence in our data that *ga* has a function of marking subject; and (b) that there is strong evidence that *ga*-marking is *pragmatically* motivated.

4.1. THE STATISTICAL FINDINGS AND THEIR IMPLICATIONS FOR *GA* AS A MARKER OF SUBJECT

4.1.1. *ga* is rare. One of the clearest findings relates to the frequency of *ga* in conversational Japanese. To support this point, we counted all the places in the data in which *ga*-marked NP could occur (calculated according to the methodology outlined earlier). As shown in Table 1, the majority of these places (66%) occur with no NP at all.

TABLE 1. *Places in which ga-marked NP could occur*

	<i>n</i>	%
∅	96	66
NP	50	34
Total	146	100

Obviously, then, one major reason for the rarity of *ga* in Japanese conversation is the rarity of NPs, due to the phenomenon known as ‘zero anaphora’.

However, as shown in Table 2, which considers the marking of the 50 NPs from Table 1, even among those NPs that could occur with *ga*, less than one-third of them do so.

TABLE 2. *NPs that could take ga*

	<i>n</i>	%
Bare NP	20	40
NP + another particle ⁹	16	32
NP + <i>ga</i>	14	28
Total	50	100

If we combine Tables 1 and 2, we get Table 3, which shows that a *ga*-marked NP occurs in only 10 percent of the places where a *ga*-marked NP could occur.

TABLE 3. *Places in which ga-marked NP could occur (Tables 1 and 2 combined)*

	<i>n</i>	%
∅	96	66
Bare NP	20	14
NP + another particle	16	11
NP + <i>ga</i>	14	10
Total	146	100

The following examples illustrate a *ga*-marked NP ((9)), and a bare NP ((10)):

(9)

(talking about the opening of a new store selling goods idolizing stars who belong to Janiizu Jimusho 'Janii's Office', a famous management company)

Harajuku ni **Janiizu Shoppu ga** dekita no
 Harajuku in Janii's Shop GA came:into:being FP
 Janii's Shop has opened in Harajuku.

In (9), the NP *Janiizu Shoppu* 'Janii's Shop' is being introduced into the conversation with *ga*. In (10), though *atashi* could be argued to be the subject of its predicate, its referent is pragmatically assumed in its context (discussed later), and *atashi* is not marked with *ga*:

(10)

atashi ∅ mada oboekirenai
 I yet remember:can:not
 I still can't remember (it).

The conclusion we draw from the numbers in Tables 1–3 is that the use of *ga* is very marked in Japanese conversation. Rather than assume that *ga* is a default nominative or subject marker, we are led by the frequency data to the conclusion that the *default* is *not* to use *ga*, and that it may thus not be realistic to view *ga* as indicating case or grammatical relations at all; rather, a better account of the grammatical resources available to speakers of Japanese might be found if we seek the *pragmatic* motivation for its use.

4.1.2. *ga* occurs largely with S. Since we found so few cases of *ga*-marked NPs, we decided to add more cases to our database, again based on random selection. Specifically we added 27 *ga*-marked NPs and 22 bare NPs. Unless specified otherwise, frequency information given from here on is based on these 83 NPs (41 *ga*-marked NPs and 42 bare NPs).

Table 4 shows the breakdown of the 41 *ga*-marked NPs in terms of their macro-roles of S, A, and O.

TABLE 4. *ga*-marked NPs¹⁰

	<i>n</i>	%
S	32	78
A	2	5
O	7	17
Total	41	100

Among the 41 *ga*-marked NPs, 32 (78%) are in the S role, and only 2 (5%) are in the A role; 7 (17%) of the 41 *ga*-marked NPs in our sample are what could be designated as O (instances of the so-called ‘objective *ga*’).

Table 5 considers just the A and the S roles from Table 4.

TABLE 5. *S* and *A* roles

	<i>n</i>	%
S	32	94
A	2	6
Total	34	100

Table 5 shows that of the 34 *ga*-marked NPs which are traditionally designated as subject (i.e. A and S), again most of them (32, or 94%) are S. Examples (9) and (11) from our database are prototypical examples of *ga*-marked S NPs. In (9), the *ga*-marked NP *Janiizu Shoppu* ‘Janii’s Shop’ is the S of the predicate *dekita* ‘came into being’.

(11)

(talking about a particular photo)

nanka **nikkori to suggoi kawaiku waratten no ga**
 uh smile with extremely cute smiling NOML GA
 atta n da yone
 existed NOML COP FP

There was uh (this) cute smiling one (i.e. photo).

In (11) the NP *nikkori to suggoi kawaiku waratten no* ‘(this) cute smiling one’ is being introduced, with *ga*, as an S with the existential predicate *atta* ‘existed’.

The statistics in Tables 4 and 5 thus show very clearly that (a) *ga* strongly tends to occur as the S with intransitive predicates; and (b) there is very little macro-role evidence to support a claim of 'subject' as the function of *ga*, since speakers sometimes use it with O, and hardly ever use it with A.¹¹

It is of interest to note that there are sporadic mentions in the literature of similar correlations between *ga*-marked NPs and the S macro-role. Suzuki and Ono (1991) show that *ga*-marked NPs highly correlate with S in spoken narrative discourse.¹² Mayes and Ono's (1993) study of the acquisition of *ga* by one Japanese child reports that there was a similar correlation between *ga* and S in the caretakers' language.¹³ Noda (1996: 89) also notes a similar tendency regarding the use of *ga* in his large database, consisting mostly of a variety of written genres, including novels, expository discourse, essays, and newspaper and magazine articles. Finally, based on a large database consisting of newspaper articles, Kondo (1997) demonstrates that *ga*-marked nominalized clauses very rarely occur in the A role. Most notably, he reports that As make up only 558 out of 26,832 *ga*-marked nominalized clauses with the nominalizer *koto* (2%) and 78 out of 25,605 *ga*-marked nominalized clauses with the nominalizer *no* (0.3%).

The very low frequency of *ga* in the A role can be attributed to two related factors, both discussed extensively in the literature on 'preferred argument structure' (Ashby and Bentivoglio, 1993; Du Bois, 1985, 1987; Du Bois et al., in press; Duranti and Ochs, 1990; Durie, 1988, 1994; Iwasaki, 1985; Matsumoto, 1997; Nakayama and Ichihashi-Nakayama, 1994). First, both cross-linguistically and in Japanese, A is a slot for given information. Such information is typically unexpressed in Japanese (so-called 'zero anaphora'), resulting in very few A NPs in the data. Second, as shown later, the types of predicates that occur with *ga*-marked NPs tend to be intransitive predicates which serve to introduce or present a referent into the conversation; transitive predicates do not tend to play this role with respect to their A arguments.¹⁴

This fact also casts strong doubt on the claim that *ga* is a marker of a subject grammatical relation, since, as noted, we would expect the subject marker to treat the A of transitive predicates and the S of intransitive predicates in a similar manner.

4.1.3. *ga* occurs most frequently with a highly constrained set of predicates. We become even more suspicious of the subject claim when we see that the predicates with which *ga* tends to occur fall into a small number of semantic classes, as shown in Table 6:

TABLE 6. *Predicates with ga-marked NPs*

	<i>n</i>	%
Nominal/adjectival/adverbial predicates	9	22
Motion verbs	4	10
Existentials	9	22
Presentatives	5	12
Others	14	34
Total	41	100

From Table 6 we can see that out of 41 *ga*-marked NPs, 9 (22%) occur in constructions with nominal, adjectival, and adverbial predicates, and 18 (44%) occur in constructions with existentials (e.g. *iru* 'exist' and *aru* 'exist'¹⁵), motion verbs (e.g. *kuru* 'come'), and presentatives (e.g. *dekiru* 'come into being').

Examples (9) and (11) illustrate the use of *ga* with a presentative *dekiru* 'come into being' and an existential, respectively. In the following example, *ga* is used with the adjectival predicate *ii* 'good':

(12)

(discussing why photos come out well in California)

koosen ga ii

light GA good

The sunlight is good.

We thus see that *ga* is very often used with existentials, motion verbs, and presentatives, all of which serve to introduce a referent or concept into the discourse (see Ashby and Bentivoglio, 1997; Cumming, 1994; Hetzron, 1971; Sanchez-Ayala, 1998 on 'presentatives'). It is also often used with nominal, adjectival, and adverbial predicates, which assign attributes to certain entities. We later motivate the use of *ga* with this set of predicates.

Thus, the skewing in the types of predicates used with *ga* further suggests that *ga* may not be a marker of nominative case or subject because we would not expect the use of such a marker to be skewed to occurring with such a small set of intransitive predicates as illustrated in Table 6. It further suggests the primarily pragmatic nature of *ga*, which we later examine.

4.1.4. *Summary.* The results of our conversation-based study of *ga* show that describing *ga* as a marker of subject is not warranted: *ga* is rare, and we do not find it marking S and A, as a well-behaved subject marker should do. Instead, we find it occurring mostly with a very limited number of intransitive predicates with low semantic content. Our next question, then, is what *are* the motivations for those rare uses of *ga*?

4.2. WHEN DO SPEAKERS USE GA?

Recall that we proposed earlier that *ga* is well characterized as *marking that its NP is to be construed as a participant in the state-of-affairs named by the predicate in pragmatically highly marked situations*. That is, we claim that *ga* is found in pragmatically highly marked situations where there is something unpredictable about the relationship between the *ga*-marked NP and the predicate such that an explicit signaling of that relationship becomes interactionally or cognitively relevant.

In this section we show how the conversational data support our hypothesis that *ga* is used in those rare situations in which an explicit marking of the relationship between the NP and the predicate seems interactionally or cognitively relevant to the interlocutors in the conversation. We see that there are several contexts in which Japanese speakers find it appropriate to mark, or signal, that a 'participant' relationship holds between its NP and the predicate. While a

multivariate approach is thus inevitable, we nevertheless believe that a single (pragmatic) account of the use of *ga* in Japanese conversation can be given.

The most obvious kinds of interactional situations in which it is appropriate to mark an NP as a participant of the state-of-affairs named by its predicate, as we have noted, are those in which the concept referred to by that NP is unpredictable, as when the referent is being introduced into the conversation, as in (9), repeated here, or when there is some interactional problem with a referent, as seen in (13):

(9)

(talking about the opening of a new store selling goods idolizing stars who belong to Janiizu Jimusho 'Janii's Office', a famous management company)

Harajuku ni **Janiizu Shoppu ga** dekita no
 Harajuku in Janii's Shop GA came:into:being FP
 Janii's Shop has opened in Harajuku.

(13)

(discussing the fact that a friend canceled a trip to Australia. R clarifies that it's the previously paid traveling expenses which were 150 thousand yen, rather than the cancellation fee)

R: kyanseruryoo toshite
 cancellation:fee as
 As (the) cancellation fee

dakara juugoman kakatta wake [ne]
 so 150:thousand cost NOML FP

I mean (it) cost 150 thousand (yen).

H: [un]
 mhm

R: ano **ryohi ga**¹⁶
 uh traveling:expenses GA
 uh traveling expenses.

These examples illustrate our hypothesis: *ga* is used to mark an NP as a participant in the state-of-affairs named by the predicate in contexts where marking that relationship is relevant. In the rest of this section, we present five pieces of evidence in favor of this hypothesis.

4.2.1. Clauses without ga are the norm. As we have shown earlier, of the NPs that were judged to be able to occur with *ga* in our data, only 28 percent do so. Thus, strictly in terms of frequency, the norm is *not* to use *ga*, but to use a bare NP, suggesting that when *ga* is used, it is used for a recognizable purpose. We take this fact to indicate that in many contexts, there is no need to mark an NP as being related to the predicate; the relationship can be inferred with no difficulty.

A further piece of evidence in support of this point is the insight of Masunaga (1988), mentioned earlier, who claims that in constructed examples, when an NP represents 'shared information', *ga* seems to be 'easy to omit'.

4.2.2. *Expressions referring to 1st or 2nd persons, 3rd person pronouns, and demonstratives.* An interesting point in favor of our claim is that expressions referring to 1st or 2nd persons, 3rd person pronouns,¹⁷ and demonstratives tend to be bare NPs. In the present discussion we will refer to this combination of forms as ‘123D’.¹⁸

Examples (10), repeated here, and (14) illustrate the use of a 123D expression as a bare NP:

(10)

atashi \emptyset mada oboekirenai
 I yet remember:can:not
 I still can't remember (it).

(14)

sore \emptyset zettai ii
 it absolutely good
 It is absolutely good.

Table 7 shows the relationship between 123D expressions and *ga*-marking.

The distribution of 123D and other expressions is clearly skewed in relation with *ga*-marking; the chi-square test shows that these results are statistically significant. Looking at the first row of Table 7 horizontally, we can see that of the 23 123D expressions in our data, the great majority, 19 (83%), are bare NPs, not marked with *ga*. Looking at the first column of Table 7 vertically, we can see that out of the 41 *ga*-marked NPs in our data, again the great majority, 37 (90%), are not 123D expressions.

This fact is predicted by our hypothesis: it is reasonable to assume that expressions referring to 1st or 2nd persons, 3rd person pronouns, and demonstratives are rarely used in pragmatically marked situations, because their referents are generally identifiable in their contexts.

4.2.3. *ga with complex NPs.* Our hypothesis about the function of *ga* in conversation provides a powerful account for the fact that *ga* tends to appear when the relevant NP is a complex NP. Examples of *ga* with complex NPs, that is, NPs which

TABLE 7. 123D expressions and *ga*-marking

	ga-marked		Bare		Total	
	n	%	n	%	n	%
'123D'	4 (10%)	17	19 (45%)	83	23	100
Other	37 (90%)	62	23 (55%)	38	60	100
Total	41	100	42	100	83	

$\chi^2 = 13.04$, d.f. = 1, $p < .01$

are either nominalizations or occur with a long modifier, include (11), repeated here, and (15):

(11)

(talking about a particular photo)

nanka **nikkori to sugoi kawaiku waratten no ga**
 uh smile with extremely cute smiling NOML GA
 atta n da yone
 existed NOML COP FP

There was uh (this) cute smiling one (i.e. photo).

(15)

natsu da kara sa **inai koto ga** ooi no yone
 summer COP because FP exist:not NOML GA much NOML FP
 (We) are often not (here) in the summer.

[more literally] Since (it) is summer, (us) not being (here) is much.

Table 8 gives the percentage of complex NP arguments that take *ga*.

The distribution of complex vs. other NPs is again clearly skewed in relation with *ga*-marking; the chi-square test shows that these results are statistically significant. In particular, we see that 17 out of the 26 complex NPs (65%) occur with *ga* and that 33 out of 42 bare NPs (79%) are non-complex NPs.¹⁹

The frequent use of *ga* with complex NP arguments provides particularly strong support for our hypothesis that *ga* is used to mark a relationship between its NP and the predicate in pragmatically highly marked situations. In real time, we suggest, the role of a complex NP is often not easily recognized with respect to its predicate. There is, therefore, a communicative advantage in *marking* the complex NP, thereby providing a signal of the *segmentation* of the utterance into participant and predicate.

Interestingly, a similar correspondence holds for *ga* with one-mora NPs and foreign words. Tsutsui (1984), using constructed examples, shows that not using *ga* for one-mora NPs is less natural, as seen in the following examples (p. 99):

TABLE 8. *Complex NPs and ga-marking*

	<i>ga</i> -marked		<i>Bare</i>		<i>Total</i>	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Complex NPs	17 (41%)	65	9 (21%)	35	26	100
Other NPs	24 (59%)	42	33 (79%)	58	57	100
Total	41	100	42	100	83	

$\chi^2 = 3.87, \text{d.f.} = 1, p < .05$

(16)

Boku wa e ga/?Ø hoshii
 I WA picture GA want
 I want a picture.

(17)

Boku wa kuruma ga/Ø hoshii
 I WA car GA want
 I want a car.

Tsutsui's examples suggest that the use of *ga* for one-mora NPs is essentially grammaticized. We claim that this can also be explained in terms of the fact that very short NPs, like long and complex NPs, are often not easily recognizable as participants, and are therefore deemed by speakers as appropriate candidates for *ga*, marking that NP as a participant and as a signal of how the utterance is to be segmented. This seems to have led to grammaticization.

Fascinating evidence for this suggestion is that in the Kansai dialects, spoken in the western part of Japan, where the non-use of so-called case particles such as *ga* and *o* is frequent (Martin, 1975), one-mora NPs without the particle do occur, but, remarkably, they are lengthened into two morae (Tsutsui, 1984). Comparing Tsutsui's example (16) with his (16'), we can see that the non-use of *ga* would be acceptable for *e* 'picture' for the Kansai dialects as long as the NP is lengthened into two morae (p. 100):

(16')

Boku wa ee Ø hoshii
 I WA picture want
 I want a picture.

It is as if the shortness of the one-mora NPs is compensated for by lengthening, so that it is 'long enough' to be recognizable as a participant in the state-of-affairs denoted by the predicate.²⁰

Yamaji (1998), based on some of the same data as our study, suggests that foreign words show a similar pattern. Our database is too small to confirm this claim, but we have found that among six foreign words in the database, four were marked by *ga*, which lends support to Yamaji's claim. Example (9) about Janii's Shop supports this claim, as does the following example from Yamaji:

(18)

(talking about when the addressee met his girlfriend for the first time when she was still a high school student)
sono imeeji ga mada nokotten deshoo
 that image GA still remaining:NOML COP:SUP
 (You) still have the image (of her being a high school girl), don't you?
 [more literally] That image is still remaining (with you)?

We suggest that foreign words are similar to complex NPs and one-mora NPs in being phonologically and/or pragmatically unusual, and thus potentially difficult to recognize as belonging to the rest of the utterance. We therefore find

that their relationship to the predicate as a participant is frequently marked with *ga*.

We have shown that complex NPs strongly tend to be marked with *ga*, and we have suggested that Tsutsui's and Yamaji's findings for one-mora NPs and foreign NPs may be related to this fact. Earlier hypotheses, based on constructed examples, have neither recognized nor provided a persuasive account of such apparently phonological constraints favoring the use of *ga*. In particular, accounts of *ga* as a marker of subject or even nominative are not able to explain the strong association of *ga* with complex NPs and the tendencies suggested by Tsutsui and Yamaji.

In contrast, our account, according to which *ga* is used in those pragmatically marked situations in which the marking of the relationship between the NP and the predicate is called for, provides a functional motivation for what looks like a purely phonological constraint on the use of *ga* in a way that earlier accounts have not been able to do.

4.2.4. *'Exhaustive listing' meaning and new (unpredictable) information.* So far we have examined three pieces of evidence in favor of our hypothesis about the pragmatic nature of *ga*. We have seen that the rarity of *ga*, the pattern of its use with 123D expressions, and its use with complex NPs all support our hypothesis.

Our fourth piece of evidence for our hypothesis is that in the data, the presence of *ga* typically carries a recognizable pragmatic meaning: it strongly correlates with what Kuno termed an 'exhaustive listing' ('singling out') meaning, as shown in (19) and (20):

(19)

uchi tte hora **onna ga** minna dekai ja nai?
 house TOP see woman GA all big COP not
 My family, see, women are all big, aren't they?

Example (20) is an instance of reported speech, where the speaker reports both a question and its answer:

(20)

'nani ni hain da'
 what in enter:NOML COP
 'What (club are you going) to join?'
 '**handbooru ga** ii kana toka omotte' toka tte sa
 handball GA good FP QT think QT say FP
 '(It's) handball (I want to join), (I) think', (I) said.
 [more literally] '(I) think handball may (be) good', (I) said.

We believe that when a speaker uses *ga* to single out a certain NP as related to a predicate, the meaning of 'X (and only X) ...' or 'It is X that ...' is naturally brought out. That is, the so-called exhaustive-listing interpretation often emerges as a pragmatic implicature on those rare occasions in which a certain NP is

explicitly marked as being a participant of the state-of-affairs named by its predicate.

A corollary to this point can be seen if we try to remove *ga* from the clauses where it is used in the data or put it in the clauses where it is not used. If we take out *ga* from those few clauses in our database that have it, in most cases they still appear to be ‘grammatical’, very little of the sense of the utterance is lost, and there would be no serious disruption to the conversation. Thus (19) and (20) without *ga* also sound right to most speakers of Japanese. More strikingly, if we try to mark our bare NPs with *ga*, they tend to acquire an ‘exhaustive listing’ meaning. So if we look at (10) again, we see the typical bare NP:

(10)
atashi \emptyset mada oboekirenai
 I yet remember:can:not
 I still can’t remember (it).

Here the speaker is simply saying she cannot remember the name of the model of her brother’s fancy sports car. She is not saying that out of a set of people it is she who cannot remember it. Now, if we try to add *ga* to *atashi* in (10), we get (10’), which has a strong implication of exhaustive listing, as suggested by our English translation:

(10’)
atashi ga mada oboekirenai
 I GA yet remember:can:not
 It’s me who can’t remember (it).

Again, we take these facts to provide strong support for our hypothesis that *ga* is a *pragmatic* marker, indicating that its NP is to be understood as a participant of the state-of-affairs named by the predicate. At the same time, we take these facts as strong support *against* the claim that *ga* is a *grammatical* marker of subject, since the variable use of subject markers would not be expected to correlate with such noticeable shifts in pragmatic interpretation.²¹

Finally, as noted, *ga*-marked NPs are often associated with ‘new’ (‘unpredictable’) information. Examples (9), (11), and (20), repeated here, illustrate this phenomenon.

(9)
 (talking about the opening of a new store selling goods idolizing stars who belong to Janiizu Jimusho ‘Janii’s Office’, a famous management company)
 Harajuku ni **Janiizu Shoppu ga** dekita no
 Harajuku in Janii’s Shop GA came:into:being FP
 Janii’s Shop has opened in Harajuku.

(11)
 (talking about a particular photo)
 nanka **nikkori to sugoi kawaiku waratten no ga**
 uh smile with extremely cute smiling NOML GA
 atta n da yone
 existed NOML COP FP

There was uh (this) cute smiling one (i.e. photo).

(20)

'nani ni hain da'
what in enter:NOML COP

'What (club are you going) to join?'

'**handobooru ga** ii kana toka omotte' toka tte sa
handball GA good FP QT think QT say FP

'(It's) handball (I want to join), (I) think', (I) said.

[more literally] '(I) think handball may (be) good', (I) said.

Again, examples such as these, demonstrating the ways in which *ga* typically carries a recognizable meaning of 'new' or 'unpredictable', show the pragmatic nature of *ga*. They further suggest that *ga* would be a rather unusual type of subject marker since subjects, in languages which have them, tend to be strongly cross-linguistically associated with *given* information (Chafe, 1994; Givón, 1979; Keenan, 1976: 319).

4.2.5. *ga* is appropriate in contexts where referent identity is under negotiation. Finally, as our fifth piece of evidence supporting our claim of *ga* as a pragmatic marker, we note that *ga* is often found in contexts where the identity of a referent as a participant in the predicate's state-of-affairs is under negotiation, as shown in (21) and (22):

(21)

(A is making a generic statement to her friends, but B, one of her children, asks who A specifically has in mind)

A: 'meshi! okane!'

meal money

sore dake kkya iwanaku naru mono ne, [dandan]

it only only say:not become NOML FP gradually

'Meal! money!' (kids) gradually get so that is all they say.

B:

[dono ko]?

which child

Which kid?

A: a, ue, **oniichan ga**

uh older older:brother GA

Uh, older, your older brother.

(22)

(K has been discussing some of her friends who have had successful relationships with men and who are themselves highly successful)

K: Hokkaidoo no tomodachi wa

Hokkaidoo of friend WA

(My) friend in Hokaidoo

hora ano shihooshiken ganbattete

see uh bar:exam trying:hard

is trying hard (for), see, uh, the bar exam and

hontoo wa

truth WA

actually

M: sono **sono onnanoko jitai ga?**

that that girl self GA
the girl herself?

K: un

mhm

In examples (21) and (22), we see two ways in which the referent of the *ga*-marked NP is under negotiation. First, in both instances an interlocutor expresses a need to have a referent clarified with a next-turn repair initiator (Schegloff et al., 1977). In (21), the repairing utterance is a *ga*-marked NP, and in (22), the repair initiator is a *ga*-marked NP. Second, in both examples, from a clause-structure point of view, an argument is provided after the predicate, in contrast to the normal predicate-final order of Japanese (Ono and Suzuki, 1992).²²

Again, we believe that interactional situations like those shown in the preceding examples strongly support our hypothesis that *ga* is primarily used in contexts where there is a pragmatic advantage to marking its NP as a participant in the state-of-affairs named by the predicate. Contexts in which a referent is being negotiated and in which the NP is expressed in an ‘unusual’ position are precisely two such situations.

4.2.6. *Summary.* We have endeavored to show how a range of types of evidence supports the central claim in this article, that *ga* is not best understood as a marker of either nominative case or subject, but is better characterized as a pragmatic marker of the relationship between its NP and the state-of-affairs named by the predicate.

As for subject, we have argued that, even if we grant the mismatch between subjecthood and *ga*-marking, as other researchers are ready to do, subject is a grammatical relation that minimally includes A and S, which does not characterize *ga* well. Furthermore, *ga* is highly infrequent, and when it does occur, its use appears to have to do with pragmatic factors rather than with grammatical relations.

As far as our pragmatic hypothesis is concerned, we have presented five pieces of evidence to support our argument that *ga* is used when it is ‘called for’ to mark an NP as a participant in the state-of-affairs named by the predicate in the context in which it is used: (1) *ga* is rare, which suggests that it is only ‘called-for’ in specific pragmatic situation; (2) 123D expressions tend not to be used with *ga*, since they are generally identifiable in their contexts; (3) *ga* tends to be used with complex NPs (and other phonologically ‘unusual’ NPs), whose relationship with the predicate can be difficult to recognize in real time; (4) *ga* generally correlates with a specific pragmatic meaning of ‘singling out’ or ‘unpredictable’; and (5) *ga* is often found when the identity of the referent of the single NP in a clause is under negotiation. The weight of this evidence in favor of our hypothesis is enhanced by the extent to which these pieces of evidence are somewhat disparate, including interactional, cognitive, and even phonological factors.

Before we conclude, we would like to comment on the status of so-called ‘objective *ga*’ and to point out another characteristic associated with *ga*: lexicalization/grammaticization.

5. ‘Objective *ga*’

Here we briefly return to the issue of ‘objective *ga*’, mentioned earlier. We consider two examples from our data. Example (23) is a kind of example that has been analyzed as a case of a transitive verb with an ‘objective *ga*’ NP, perhaps partly because of the English translation of ‘having’ a boyfriend:

(23)

kareshi ga ita n desu yo
 boyfriend GA existed NOML COP FP
 (She) had a boyfriend.

Example (24) contains a clause which seems to be composed of a predicate *iya* ‘dislike’ and its ‘object’, a *ga*-marked NP. K is basically saying that she enjoys talking for a long time if people call her first, but does not like initiating calls. Providing a collaborative completion to what T has just said, she produces an ‘objective *ga*’ NP *jibun kara nanka koo kakeru no ga* ‘calling (from myself)’ for the predicate *iya* ‘dislike’:

(24)

K: kakete kita nagadenwa toka wa
 call came long:call like WA
 Like a long call (which they) made
 sugoi sukina n da ked[o]
 awfully like NOML COP but
 (I) really like (it) but

T:

[u]n **jibun kara nanka koo**
 mhm self from uh uh
 mhm, from (your)self uh

K: **kakeru no ga** ne
 call NOML GA FP
 calling (from myself)

T: a [a]
 ah ah

K: [iyana no]
 dislike FP
 (I) don’t like (it).
 (It’s calling from myself (i.e. originating calls) that I) don’t like.

The issue can be stated as follows: if *ga* is a nominative or subject marker, then its occurrence to mark ‘objects’ with predicates such as *iru* ‘exist’ and *iya* ‘dislike’ is anomalous: how can a single ‘case’ marker be found marking both ‘subjects’ and ‘objects’? In our approach to *ga*, this fact is not an anomaly for the simple reason that we do not consider *ga* to be marking either nominative or subject. If the function of *ga* is to mark an NP as a participant of the state-of-affairs

named by the predicate, then whether that NP is ‘subject’ or ‘object’ is not a problem.

At this point one might argue that instances of ‘objective *ga*’ could be found in clauses which *also* contained another (‘subject’) argument. In fact such instances seem to be extremely rare: in our data, none of the seven *ga*-marked NPs that we coded as instances of ‘objective *ga*’ has another (‘subject’) *ga*-marked NP appearing with it. This finding indicates that, as far as *ga*-marking is concerned, predicates with ‘objective *ga*’ NPs are essentially one-argument predicates. If such predicates are seen as *taking* just one argument, then a *ga*-marked NP occurring with them will be an S, and there is no motivation for suggesting a category such as ‘objective *ga*’. With this new understanding, the percentage of *ga*-marked NPs that are S would rise from 78 to 95 percent (39 out of 41) (see Table 4), making our general pattern even more general.

As a further comment on ‘objective *ga*’, there is no interactional or functional motivation for positing a separate category ‘objective *ga*’. The pragmatic function of the so-called ‘objective *ga*’ does not seem to be any different from that of the *non*-‘objective *ga*’, which we described earlier, as illustrated by our examples (23) and (24): (23) is a case of an introduction of a new participant in the discourse, and the *ga*-marked NP in (24) has the meaning of exhaustive listing, as our translation indicates.

6. *ga* is often lexicalized or grammaticized

Finally, we would like to point out another characteristic of *ga*, which, as far as we know, has not been discussed in the literature, and which we think may further illuminate the function of *ga*.

Throughout our investigation, we have noticed that there are some examples with *ga* which may be seen as grammaticized or lexicalized. (We use the term ‘grammaticized’ instead of the more cumbersome ‘grammaticized or lexicalized’ in the remainder of this discussion, since the differences between the two are not crucial for our purposes.) The following examples illustrate two grammaticized expressions with *ga*, namely *hazu ga* and *hoo ga*:

(25)

daitai sa soo yuu tokoro no hanashi wakaru **hazu ga**
 generally FP so say place of talk understand reasoning GA
 nai n da yone
 not:exist NOML COP FP

Obviously there is no way that (he) (would) understand talk about that area.

[more literally] Obviously there is no reasoning that (he) understands talk about that area.

(26)

doobutsuen ni ireteru **hoo ga** tanoshii ja nai ka yo
 zoo in putting direction GA fun COP not Q FP
 Wouldn't it be more fun to keep (him) in a zoo?

[more literally] Isn't the direction (choice) of keeping (him) in the zoo more fun?

As the translations indicate, *hazu* 'reasoning' in (25) and *hoo* 'direction' in (26) (called *keishikimeishi* 'formal nouns' in Japanese grammar) are relatively weak in lexical content. In fact, it may be somewhat difficult for average Japanese speakers to come up with these lexical meanings though they would certainly know what the entire utterances mean. Both of these utterances thus show their grammaticized nature, but the status of *ga* appears to be different between the two: it seems to be more optional in (25) and more necessary in (26).

Besides these more clearly grammaticized cases, some examples which we have already discussed seem to be associated with certain degrees of conventionalization. Not using *ga* in those examples may not yield completely ungrammatical utterances, but they can sound somewhat unnatural for some speakers. (12) and (23), repeated here, are two such examples:

(12)

(discussing why photos come out well in California)

koosen ga ii

light GA good

The sunlight is good.

(23)

kareshi ga ita n desu yo

boyfriend GA existed NOML COP FP

(She) had a boyfriend.

As it happens, such (partially) grammaticized expressions as illustrated in the preceding four examples are associated with constructions with two broad types of predicates: (a) presentatives, existentials, and motion verbs; and (b) nominal, adjectival, and adverbial predicates. We have pointed out that the former type introduces a participant into the discourse, and the latter type assigns attributes to certain entities. Both of these are environments in which it would be most natural to mark that NP as a participant in the state-of-affairs described by the predicate.

It is interesting to note further that these are the most frequent verb types used with *ga*, as seen in Table 6. For this reason, we hypothesize that the frequency of constructions with particular verb types is playing a role in this grammaticization (Ashby, 1981; Bybee, forthcoming a, b; Bybee and Thompson, in press; Bybee et al., 1994; Du Bois, 1987; Hopper, 1987).

We also hypothesize that the semantics of these predicates play a role. That is, existentials, presentatives, and motion verbs are typically semantically rather empty. With respect to the nominal, adjectival, and adverbial predicates shown in Table 6, with which *ga* also frequently occurs, we suggest that a comparison with verbal predicates can reveal a motivation for the prevalence of *ga*-marking with these non-verbal predicates: verbal predicates typically express events and states, which naturally invoke participants, while nominal, adjectival, and adverbial predicates typically do not. Instead they only name a class or a property without

necessarily invoking any participants. To signal that a given NP is intended to go with such a predicate, *ga* plays a more important role than with verbal predicates, whose semantics typically specify something about the nature of the participants involved.

In both instances (that is, (1) existentials, motion verbs, and presentatives; and (2) nominal, adjectival, and adverbial predicates), the relationship between the single arguments of these predicates and the predicates themselves is underdetermined by the semantics of these predicate types. This may explain why it is common for NPs to be explicitly marked with *ga* when they appear with these types of predicates.²³

Interestingly, we also found that both these predicate types which most frequently occur with *ga* are often associated with different degrees of conventionalization, with examples like (25) and (26) among the most conventionalized cases. We believe this non-discreteness comes from the fact that we are essentially dealing with ongoing change, and our use of such terms as 'grammaticization' and 'lexicalization' reflects this belief. Not being able to judge the grammaticality of examples categorically is a natural consequence of this.

Finally, it should be noted that we do not exclude these conventionalized cases from our investigation for several reasons. First, most cases seemed to allow the choice of using or not using *ga*; and second, it was impossible to make a decision as to which cases should be excluded since they are associated with different degrees of conventionalization. But most importantly, following Bybee et al. (1994), we felt that in order to seek a deeper understanding of the function and the patterning of *ga*, we need to examine the whole spectrum along which different uses of *ga* may be aligned on the continuum of conventionalization. We suggest that, due to the frequency of their association with the pragmatic functions which we have outlined here, and to the relative lack of specificity of these predicate types, these cases are excellent candidates for grammaticization.

We have seen that some utterances containing *ga* are associated with certain degrees of conventionalization. Those utterances take either (i) presentative/existential/motion verb; or (ii) nominal/adjectival/adverbial predicates, which are the most frequent types of predicates in our database. Presentative/existential/motion verb predicates are used to introduce participants into the discourse and nominal/adjectival/adverbial predicates name a class or a property without necessarily invoking any participants. In terms of semantics, they do not contribute much. That is, for these predicates, the explicit marking of the relation between the participant and the state-of-affairs which they express is an integral part of the discourse function they serve. We suggest this may have led to the grammaticization of these expressions.

7. Concluding remarks

We have taken an empirical approach to the Japanese nominal particle *ga*, which has been nearly universally viewed as a marker of either nominative case or subject. We

have brought evidence to support our contention that *ga* has very little to do with either case or grammatical relations, but rather has to do with pragmatic issues involving the interactional advantage in marking an NP as a participant in the state-of-affairs named by the predicate.

We have suggested that this advantage arises most readily with one-argument predicates, where the semantics of the verb may reveal rather little about the relationship between that argument and that predicate. Most of the time, as we have shown by the rarity of *ga* in the data, the nature of this relationship is not in general an issue for Japanese conversationalists, but on occasion, the speaker finds it appropriate to signal that the NP and the predicate 'go together' in the state-of-affairs named by the predicate, especially in contexts in which the NP:

- names a concept is being singled out with respect to other potential concepts ('exhaustive listing' contexts);
- names a concept which is being introduced (presentative/existential/motion verb contexts), or is 'new' or 'unpredictable';
- is phonologically 'unusual' in being relatively long, short, or 'foreign';
- names a concept which is under negotiation.

What we have proposed as a new way to understand the use of the nominal particle *ga* in Japanese has parallels throughout Japanese grammar. The nominal particle *wa* has long been claimed to be pragmatically motivated (see especially Hinds et al., 1987; Kuno, 1973a, 1973b; Maynard, 1980, 1981; Noda, 1996; Shibatani, 1990). Matsuda (1996) and Fujii and Ono (1998) show that the likelihood of using the nominal particle *o*, often thought of as an 'object' or 'accusative' case marker in Japanese, is similarly low, and, just as we have claimed for *ga*, can be shown to be similarly motivated by pragmatic, rather than grammatical, factors. Putting these facts together, we suggest that a realistic account of Japanese grammar would approach both *ga* and *o*, as well as *wa*, in terms of their pragmatic functions rather than in terms of grammatical relations or case.

Finally, we return to an issue mentioned in our discussions of previous literature: the issue of 'register'. Upon hearing our findings, several linguists commented that our conclusion that *ga* is not a marker of case or grammatical relations in conversational Japanese does not match their assumptions about their language. We suggest that this may have to do with 'register', especially differences between written and spoken language (Cumming and Englebretson, in press; Linell, 1982; Miller, 1995; Miller and Weinert, 1998). As we have noted, it is quite possible that uses of *ga* in written Japanese might influence linguists' intuitions about the grammar of spoken Japanese.

It remains to be seen to what extent *ga* can be analyzed as a marker of nominative or subject in any genre of written Japanese. However, Fujii's (1991) diachronic study of the increase in the use of *ga* in successive translations of *Genji Monogatari* [The Tale of Genji] provides some evidence for *ga* being more grammaticized in written Japanese. Fujii's analysis strongly suggests that even this grammaticization is due to influence from western languages (which have been

more successfully analyzed in terms of a grammatical relation of ‘subject’) starting from the mid-19th century. At any rate, if, as seems to be the case, specialists in Japanese grammar are inclined to view *ga* as a marker of subject, even in the face of the contradictory findings from our conversational data, we believe that the practice of constructing examples, influenced from written language, provides the basis for a ready explanation.

ACKNOWLEDGEMENTS

The seeds for this article were planted in Suzuki and Ono (1991), which gave us the impetus to examine *ga* in conversational Japanese. We are also grateful to Susanna Cumming, Aryeh Faltz, Makoto Hayashi, Shin Ja Hwang, Robert Kirsner, Kyoko Masuda, Tomoyo Takagi, and especially Joan Bybee and Shoichi Iwasaki for valuable input in the preparation of this article. Any remaining errors are our responsibility.

NOTES

1. These environments include ‘double nominative’ constructions, *ni*-marked ‘subject’ constructions, and *ga*-marked ‘object’ constructions.
2. Unless otherwise noted, as here, *ga*, *o*, and *wa* are glossed as GA, O, and WA, respectively, in this article.
3. The following abbreviations are used in this article: ACC, accusative; COP, copula; FP, final particle; NOM, nominative; NOML, nominalizer; Q, question; QT, quotative; SUP, supposition; TOP, topic.
4. By ‘A’ we mean the agent-like argument of a two-argument predicate, and by ‘S’ we mean the single argument of a one-argument predicate, following Blake (1994), Comrie (1978), Croft (1991), Dixon (1979), and Palmer (1994).
5. However, it has been pointed out in several studies that the connection between *ga* and new information is not always supported by actual discourse data (Maynard, 1980, 1981, 1987; Suzuki and Ono, 1991).
6. We have modified Masunaga’s examples slightly for the sake of clarity.
7. We have modified Kuno’s glossing slightly for the sake of clarity.
8. There seems to be dialectal variation in the use and non-use of *ga* (Martin, 1975). For this reason, we restricted our study only to standard Japanese.
9. The category ‘another particle’ includes *wa*, *mo*, and *tte*.
10. For this table, *ga*-marked NPs with predicates which could *also* be associated with either an NP marked with *o* (generally considered an ‘object’ marker) or an ‘objective *ga*’-NP are classified as A. The ‘objective *ga*’-NPs are classified as O. If the predicate could not be associated with either of these NP types, the *ga*-marked NPs are classified as S. The two NPs for so-called ‘double nominative’ constructions are classified both as S, following Shibatani (1990: 295).
11. Note that one cannot make the counterargument that *ga* should still be regarded as a subject marker because in general it ‘can’ occur only with A and S. Such a statement would not only be descriptively inadequate, considering what is observed in actual use, but would misrepresent the primarily pragmatic nature of *ga*-marking, which we are arguing for in this article.
12. Iwasaki (1985) finds some correlation with *ga* and A. However, even his figures indicated that *ga*-NPs are much more likely to be S than A.
13. Mayes and Ono (1993) also note that the order of emergence of *ga*-marking in the

child's language followed the frequency of different predicate types with *ga* in the caretakers' language.

14. Our claim that *ga*-marked NPs tend to be in the S role implies a clear distinction between A, S, and O macro-roles in Japanese. In fact, we have found many instances of NPs in our conversational data whose assignment to any of these macro-roles is quite problematic (for some discussion, see Ono and Thompson, 1997, and the discussion of 'objective *ga*' in Section 5). For the present purposes, our point is that (a) most examples of *ga* occur in slots that are not clearly A or O; and (b) clear examples of *ga*-marked A NPs are very rare.
15. The choice between *iru* and *aru* is generally described as being related to animacy, *iru* being used with animate referents and *aru* with inanimate. See Strauss (1993) for an alternative view.
16. In contrast to the 'canonical' predicate-final order, the *ga*-marked NP in (13) is expressed after the predicate for the purpose of clarifying the referent (Ono and Suzuki, 1992). We will return to this point in Section 4.2.5.
17. As has been reported in the literature (e.g. Clancy 1980), 3rd person pronouns are very rare in spoken Japanese; in our database there was only one example:

kare ne zettai hyooka dekite nai yo
 he FP absolutely evaluation being:able NEG FP
 Surely he is not able to evaluate (it).

18. The complexity in naming this group is partly due to the fact that expressions referring to 1st and 2nd person in Japanese are sometimes not pronouns, but various types of lexical NPs.
19. Watanabe (1986) finds a similar statistical skewing in regard to the use and non-use of nominal particles, including *ga*.
20. The Kansai dialects, incidentally, provide one further hint that *ga* is not best understood as a marker of grammatical relations or case, since these dialects are known for omitting more nominal particles (including *ga*) than dialects of other parts of Japan (Martin, 1975; Masunaga, 1988). If *ga*'s function were to mark subject, this would be an unexpected direction for dialect variation.
21. However, see Section 6, where we discuss how this principle may be affected by the process of grammaticization.
22. In our database, four out of five NPs which are expressed after the predicate are marked by *ga*.
23. Shoichi Iwasaki has pointed out to us that the account given in Section 4.2.3 regarding the tendency of complex NPs to be marked with *ga* also provides a motivation for the use of *ga* with such *keishikimeishi* 'formal nouns' as seen in (25) and (26), since these 'formal nouns' must occur with modifiers.


REFERENCES

- Aoyama, Takashi (1983) 'The Free-Floating Focus System in Japanese: Form-Content Analysis of WA and GA', *Gengo Kenkyu* 83: 41-60.
- Ashby, William J. (1981) 'The Loss of the Negative Particle *ne* in French: A Syntactic Change in Progress', *Language* 57: 674-87.
- Ashby, William J. and Bentivoglio, Paola (1993) 'Preferred Argument Structure in Spoken French and Spanish', *Language Variation and Change* 5: 61-76.
- Ashby, William J. and Bentivoglio, Paola (1997) 'Strategies for Introducing New Referents into Discourse: A Comparative Analysis of French and Spanish Presentational

- Structures', in Robert M. Hammond and Marguerita G. MacDonald (eds) *Linguistic Studies in Honor of Bohdan Saciuk*, pp. 9–26. West Lafayette, IN: Learning Systems.
- Blake, Barry J. (1994) *Case*. Cambridge: Cambridge University Press.
- Bybee, Joan (forthcoming a) 'Mechanisms of Change in Grammaticization: The Role of Frequency', in Rich Janda and Brian Joseph (eds) *A Handbook of Historical Linguistics*. Oxford: Blackwell.
- Bybee, Joan (forthcoming b) 'The Emergent Lexicon', *Chicago Linguistic Society* 34.
- Bybee, Joan, Pagliuca, William and Perkins, Revere (1994) *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago, IL: The University of Chicago Press.
- Bybee, Joan and Thompson, Sandra A. (in press) 'Three Frequency Effects in Syntax', *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society*. Berkeley.
- Chafe, Wallace (1987) 'Cognitive Constraints on Information Flow', in Russell Tomlin (ed.) *Coherence and Grounding in Discourse*, pp. 21–51. Amsterdam: Benjamins.
- Chafe, Wallace (1994) *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago, IL: University of Chicago Press.
- Clancy, Patricia (1980) 'Referential Choice in English and Japanese Narrative Discourse', in Wallace Chafe (ed.) *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, pp. 127–202. Norwood, NJ: Ablex.
- Comrie, Bernard (1978) 'Ergativity', in Winfred P. Lehmann (ed.) *Syntactic Typology*, pp. 329–94. Austin: University of Texas Press.
- Croft, William (1991) *Syntactic Categories and Grammatical Relations*. Chicago, IL: University of Chicago Press.
- Cumming, Susanna (1994) 'Functional Categories in the Lexicon: Referent Introduction in Indonesian Novels', *Text* 14(4): 465–94.
- Cumming, Susanna and Englebretson, Robert (in press) "'Tight" and "Loose" Syntax in Malay/Indonesian: Dimensions of Language Structure', in Carol Moder (ed.) *Discourse Across Languages and Cultures*. Amsterdam: Benjamins.
- Dixon, R.M.W. (1979) 'Ergativity', *Language* 55: 59–138.
- Du Bois, John W. (1980) 'Beyond Definiteness: The Trace of Identity in Discourse', in Wallace Chafe (ed.) *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.
- Du Bois, John W. (1985) 'Competing Motivations', in John Haiman (ed.) *Iconicity in Syntax*, pp. 343–65. Amsterdam: Benjamins.
- Du Bois, John W. (1987) 'The Discourse Basis of Ergativity', *Language* 63: 805–55.
- Du Bois, John W., Ashby, William J. and Kumpf, Lorraine, eds (in press) *Preferred Argument Structure: The Next Generation*. Amsterdam: Benjamins.
- Duranti, Alessandro and Ochs, Elinor (1990) 'Genitive Constructions and Agency in Samoan Discourse', *Studies in Language* 14(1): 1–23.
- Durie, Mark (1988) 'Preferred Argument Structure in an Active Language: Arguments Against the Category "Intransitive Subject"', *Lingua* 74: 1–25.
- Durie, Mark (1994) 'Pragmatic Linking in Acehnese', *Text* 14(4): 495–529.
- Fujii, Noriko (1991) *Historical Discourse Analysis: Grammatical Subject in Japanese*. Berlin: Mouton de Gruyter.
- Fujii, Noriko and Ono, Tsuyoshi (1998) 'The Occurrence and Non-Occurrence of the Japanese Direct Object Marker *o* in Conversation', unpublished manuscript, University of Oregon.
- Givón, Tammy (1979) *On Understanding Grammar*. New York: Academic Press.
- Hetzron, Robert (1971) 'Presentative Function and Presentative Movement', *Studies in African Linguistics Supplement* 2: 79–105.

- Hinds, John, Maynard, Senko K. and Iwasaki, Shoichi, eds. (1987) *Perspectives on Topicalization: The Case of WA*. Amsterdam: Benjamins.
- Hopper, Paul J. (1987) 'Emergent Grammar', *Berkeley Linguistics Society* 13: 139–57.
- Hopper, Paul J. and Thompson, Sandra A. (1980) 'Transitivity in Grammar and Discourse', *Language* 56: 251–99.
- Hopper, Paul J. and Thompson, Sandra A. (1984) 'The Discourse Basis for Lexical Categories in Universal Grammar', *Language* 60: 703–52.
- Iwasaki, Shoichi (1985) 'The "Given A Constraint" and the Japanese Particle *ga*', in Scott DeLancey and Russell S. Tomlin (eds) *Proceedings of the First Annual Meeting of the Pacific Linguistic Conference*, pp. 152–67. Eugene: University of Oregon Press.
- Keenan, Edward L. (1976) 'Towards a Universal Definition of "Subject"', in Charles N. Li (ed.) *Subject and Topic*, pp. 303–34. New York: Academic Press.
- Kondo, Yasuhiro (1997) '*No koto ni yoru* Meishisetsu no Seishitsu [The nature of *no* and *koto* nominalized clauses]', *Kogugogaku* 190: 1–11.
- Kuno, Susumu (1972) 'Functional Sentence Perspective: A Case Study from Japanese and English', *Linguistic Inquiry* 3: 269–320.
- Kuno, Susumu (1973a) *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.
- Kuno, Susumu (1973b) *Nihon Bunpoo Kenkyuu* [A study of Japanese grammar]. Tokyo: Taishuukan.
- Kuroda, S.-Y. (1965) 'Generative Grammatical Studies in the Japanese Language', unpublished PhD thesis, Massachusetts Institute of Technology.
- Linell, Per (1982) 'The Written Language Bias in Linguistics', unpublished manuscript, Department of Communication Studies, University of Linköping.
- Martin, Samuel E. (1975) *A Reference Grammar of Japanese*. New Haven, CT: Yale University Press.
- Matsuda, Kenjiro (1996) 'Variable Zero-Marking in (o) in Tokyo Japanese', unpublished PhD thesis, University of Pennsylvania.
- Masanaga, Kiyoko (1988) 'Case Deletion and Discourse Context', in William J. Poser (ed.) *Papers from the Second International Workshop on Japanese Syntax*, pp. 145–56. Stanford, CA: CSLI.
- Matsumoto, Kazuko (1997) 'NPs in Japanese Conversation', *Pragmatics* 7(2): 163–81.
- Mays, Patricia and Ono, Tsuyoshi (1993) 'The Acquisition of the Japanese Subject Marker *ga* and its Theoretical Implications', in Patricia Clancy (ed.) *Japanese/Korean Linguistics* pp. 2, 239–47. Stanford, CA: CSLI.
- Maynard, Senko K. (1980) 'Discourse Functions of the Japanese Theme Marker -WA', Unpublished PhD thesis, Northwestern University.
- Maynard, Senko K. (1981) 'The Given/New Distinction and the Analysis of the Japanese Particles -*Wa* and -*Ga*', *Papers in Linguistics* 14: 109–30.
- Maynard, Senko K. (1987) 'Thematization as a Staging Device in the Japanese Narrative', in John Hinds, Senko K. Maynard and Shoichi Iwasaki (eds) *Perspectives on Topicalization: The Case of Japanese 'Wa'*, pp. 57–82. Amsterdam: Benjamins.
- Miller, Jim (1995) 'Does Spoken Language have Sentences?', in F. R. Palmer (ed.) *Grammar and Meaning*, pp. 116–35. Cambridge: Cambridge University Press.
- Miller, Jim and Weinert, Regina (1998) *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Oxford University Press.
- Nakayama, Toshihide and Ichihashi-Nakayama, Kumiko (1994) 'Discourse Perspective for Core-Oblique Distinction in Japanese', *Santa Barbara Papers in Linguistics* 5: 158–77.
- Noda, Hisashi (1996) '*Wa to ga* [*Wa* and *ga*]', in *Shin Nihon Bunpoo Sensho* 1 [Selected Works in Contemporary Japanese Grammar 1]. Tokyo: Kuroshio Shuppan.

- Ono, Tsuyoshi and Suzuki, Ryoko (1992) 'Word Order Variability in Japanese Conversation: Motivations and Grammaticization', *Text* 12: 429–45.
- Ono, Tsuyoshi and Thompson, Sandra A. (1997) 'Deconstructing "Zero Anaphora"', in *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society*, Berkeley.
- Palmer, F.R. (1994) *Grammatical Roles and Relations*. Cambridge: Cambridge University Press.
- Sanchez-Ayala, Ivo (1998) 'The Spanish Presentative Lexicon', unpublished manuscript, University of California at Santa Barbara.
- Schegloff, Emanuel, Jefferson, Gail and Sacks, Harvey (1977) 'The Preference for Self-Correction in the Organization of Repair in Conversation', *Language* 53: 361–82.
- Shibatani, Masayoshi (1977) 'Grammatical Relations and Surface Cases', *Language* 53: 789–809.
- Shibatani, Masayoshi (1990) *The Languages of Japan*. Cambridge: Cambridge University Press.
- Shibatani, Masayoshi (1991) 'Grammaticization of Topic into Subject', in Bernd Heine, and Elizabeth Traugott (eds) *Grammaticalization II*, pp. 93–133. Amsterdam: Benjamins.
- Strauss, Susan (1993) 'Do *aru* and *iru* Exist as Verbs of Existence in Japanese?', in Soonja Choi (ed.) *Japanese/Korean Linguistics*, Volume 3, pp. 271–87. Stanford, CA: CSLI.
- Sugamoto, Nobuko (1982) 'Transitivity and Objecthood in Japanese', in Paul Hopper and Sandra A. Thompson (eds) *Studies in Transitivity*, pp. 423–47. New York: Academic Press.
- Suzuki, Ryoko and Ono, Tsuyoshi (1991) 'Japanese *ga*, Spotlighting, and Intransitives in Spoken Narratives', *Santa Barbara Papers in Linguistics* 3: 144–52
- Tsutsui, Michio (1983) 'Ellipsis of *Ga*', *Papers in Japanese Linguistics* 9: 199–244.
- Tsutsui, Michio (1984) Particle Ellipses in Japanese', unpublished PhD thesis, University of Illinois.
- Watanabe, Yasuko (1984) 'Transitivity and Evidentiality in Japanese', *Studies in Language* 8: 235–51.
- Watanabe, Yasuko (1986) 'Two Kinds of Ellipsis in Japanese Discourse: A Quantitative Text Study', *Studies in Language* 10: 337–51.
- Yamaji, Harumi (1998) 'GA-Ellipses in Japanese', unpublished manuscript, University of Arizona.
- Yanagida, Yuko (1985) 'Informational Predictability and Case-Marking Variation in Japanese', *Text* 5: 123–45.



TSUYOSHI ONO is assistant professor in East Asian Studies at the University of Arizona. He specializes in conversation and grammar and language change, focusing particularly on the way in which frequent patterns in conversational interaction become habituated as grammatical structure. He has co-authored a number of articles on different morphosyntactic problems of both Japanese and English. He is currently building a corpus of conversational Japanese and preparing to write a book exploring the implications for the nature of Japanese grammar gained from the close analysis of conversational data. ADDRESS: Department of East Asian Studies, University of Arizona, Tucson, AZ 85721, USA. [email: ono@u.arizona.edu]

SANDRA A. THOMPSON is Professor of Linguistics at the University of California at Santa Barbara. She specializes in discourse and language universals, and is particularly interested in the role of patterns of conversational discourse in shaping morphosyntactic regularities. She is the co-author with Charles Li of *Mandarin Chinese: A Functional Reference Grammar*, and with Paul Hopper of *Transitivity in Grammar and Discourse*. She has co-edited *Studies in Transitivity* with Paul Hopper, *Clause Combining in Grammar and Discourse* with John Haiman, *Discourse Description* with William C. Mann, and *Interaction and Grammar* with Elinor Ochs and Emanuel Schegloff. ADDRESS: Department of Linguistics, University of California, Santa Barbara, CA 93106, USA. [email: sathomps@humanitas.ucsb.edu]

RYOKO SUZUKI is Assistant Professor of Department of Economics at Keio University in Japan. She specializes in discourse and grammar, language change and acquisition. She is currently interested in grammaticization in Japanese, in which various clause-final morphemes turn into pragmatic particles. She is interested in carrying out cross-linguistic research on grammaticization of quotative constructions. ADDRESS: Shin-kenkyushitsu, Keio University 4-1-1 Hyoshi, kohokuku, Yokohama 223-8521 Japan [email: ryoko@hc.cc.keio.ac.jp]