Properties of particle “omission” revisited

Mitsuaki Shimojo
University at Buffalo, SUNY

This study examines the so-called particle omission (or ellipsis) in Japanese and critically evaluates the previously suggested view that the zero marking represents the mere omission of certain overt particles. By examining speakers’ use of the zero form in Japanese conversation, rather than its acceptability in constructed examples, I point out an underlying discourse property of the zero form—defocusing of referents in subsequent discourse—and suggest that the observed discourse property is ascribed to absolute specification of referents—the inherent grammatical property associated with the zero marking. Furthermore, I claim that the high frequency of the zero form observed in conversations is due to the general function denoted by the form, which unspecifics a contrastive relationship. Overall, given its unique roles, the zero form is best considered as a zero particle—a separate and important paradigmatic choice in post-nominal marking.

1. Introduction

The use of a zero form is a salient property of Japanese, as is most clearly exemplified by the extensive use of zero anaphora, as shown in (1a). However, the array of zero forms includes the so-called particle omission or particle ellipsis, as shown in (1b)—a modest form of zero but it is observed extensively to encode subjects and objects, particularly in spoken Japanese.1

1 The following abbreviations are used in this paper: ACC=accusative, CONJ=conjecture COP=copula, F=filler, HON=honorific, IT=interactional element, LK=linker, NMZ=nominalizer, NOM=nominative, QT=quotative marker, TOP=topic.
Particle omission is commonly so-labeled because Japanese uses post-nominal markers for dependent-marking, whether a case marker such as the nominative marking as shown in (1c) or an adverbial marker such as the topic marker as shown in (1d). The four sentences given in (1) represent four different ways to represent the same propositional content, i.e. “the teacher came”, supposing that the referent in zero anaphor in (a) is so-identified.

With respect to particle omission, there are some cardinal questions which remain to be elucidated, as given in (2).

(2) Questions concerning particle omission
(i) Should it be considered as mere omission of overt particles or as a zero particle—a separate paradigmatic choice which is associated with unique properties?
(ii) What makes this encoding type so prevalent in use relative to overt marking, as reported in recent text analyses?

With respect to the first question above, a number of previous studies assumed or implied that the phenomenon in question results from the omission of an overt particle and their analyses are typically based on conditions in which given sentences remain acceptable, or result in unacceptability, without the overt post-nominal markings. There is an inherent problem in this type of approach because such analyses have little to offer to explain why the speaker chooses the zero form when the zero form as well as the overt form is acceptable in a given sentence. In order to probe the properties of the zero form, we need to broaden the scope and examine how the speaker’s choice is made. Furthermore, the performance aspect of the phenomenon presents a significant asymmetry in frequency of use in terms of speakers’ choice of post-nominal markers, as the zero form clearly outnumbers overt forms in text counts (Shimojo 2005). Does the prevalence in use relate to the functional properties of the zero form?

The present study examines conversational Japanese and attempts to elucidate functional properties of the zero marking with respect to the questions presented above. In Section 2, I will present a brief summary of previous claims, pertaining to the zero form as particle ellipsis and as a zero particle, and also previous observations in text-based studies. In Section 3, I will outline the grammatical properties of the zero form and lay out the basis to consider the zero form as a zero particle. I will also present quantitative data to support the view of a zero particle by pinpointing the unique function
associated with the form. In Section 4, I will discuss further the zero particle as a paradigmatic choice and argue that it represents an unmarked category of the paradigm of post-nominal marking, which is related to the high frequency in use.

2. Previous studies


The previous studies that consider the zero form as ellipsis claim a variety of factors that trigger the ellipsis. Among them, there are sociolinguistic factors, such as formality of speech and familiarity of referents. For example, Tsutsui (1984a) claims that the lower the formality level is, the more acceptable the particle ellipsis is. For instance, in (3a), with explicit case marking for each argument, the sentence is more formally sounding than (3b), in which there is no overt case marking (despite the same honorific form of the predicate).

(3)  a. asita sensei ga oosaka ni irassyaimasu
tomorrow teacher NOM Osaka to go.HON
‘The teacher will go to Osaka tomorrow.’

b. asita sensei oosaka irassyaimasu

It has also been pointed out that the acceptability of particle ellipsis correlates with assignment of pragmatic focus in a given sentence. For example, when NPs represent a contrastive relationship, as exemplified by (4), the particle may not be ellipsed (Tsutsui 1984a, Yatabe 1999).

(4)  taroo wa/*Ø iku kedo jon wa/*Ø ikanai
Taro TOP go but John TOP go.NEG
‘Taro will go but John won’t.’

Likewise, when the NP carries the so-called narrow focus of the sentence or an exhaustive-listing interpretation, particle ellipsis is not allowed, as shown in (5B).

(5)  A: dare ga kita no?
who NOM came IT
‘Who came?’

B: taro ga/*Ø kita
NOM came
‘Taro came.’

The description from the viewpoint of constraints on the use of the zero form more or less characterizes the form as an optional means of encoding of an NP and that under certain conditions—whatever they are—the optional form is unavailable. This type of approach is not useful for the purpose of the present study since descriptions of the
acceptable range of the zero form do not pinpoint speakers’ use of the form and fail to delineate any inherent properties associated with the form.


Contrary to the approach outlined above, the zero form may be considered as a zero particle, which itself is a full-fledged paradigmatic choice and denotes properties separate from those of the other overt forms. The primary motivation for this view is that the zero marking is sometimes obligatory; in such cases, none of the overt particles is a better choice than the zero. In fact, some previous analyses suggest that there are two types of zero marking—the optional one and the obligatory one. Consider the examples in (6).

(6) a. sensei Ø/ga/wa inakunaramasita
   teacher NOM/TOP disappeared
   ‘The teacher has disappeared.’

b. watasi Ø/*ga/?wa bikkurimasita
   I NOM/TOP was surprised
   ‘I was surprised.’

In (6a), any of the three markers—the zero particle, the case marker, and the topic marker—may be used to mark the subject NP without causing anomaly for the sentence given in isolation. In (6b), on the other hand, the zero particle is the best choice if the sentence is given without a particular context. Ga is unacceptable unless the subject represents narrow pragmatic focus, as in an answer to a question “Who was surprised?” Wa, on the other hand, brings out a sense of objective judgment about the self, as well as the sense of contrast, and therefore subdues the emotional tone of the utterance, contrary to the expected speaker’s emotional expression. Thus, the use of wa results in anomaly unless a marked contrastive reading or an objective description of the self is intended. It has been pointed out (Shibatani 1990, Lee 2002) that the obligatory zero particle is typically associated with expressing the speaker’s own feeling or emotion. Therefore, the contrast between the two examples in (6) would suggest a varying degree of preference for the zero particle. As far as the acceptability of a particle choice is concerned, while the zero form in (6a) is on a par with the other particles, it is almost obligatory in (6b).

In the present study, however, I do not employ a categorical distinction of the zero form on the basis of obligatoriness in particle choice, as described above. There are some inherent problems with such a distinction, at least for the purpose of this study. First, the distinction between particle ellipsis and a zero particle is hardly clear-cut, relying on a notion of obligatoriness that is empirically difficult to define. Obviously, contexts play a significant role in speakers’ choice of a post-nominal marker and the acceptability of the zero form varies accordingly. For instance, even in (6b), the zero form is not entirely obligatory as the use of wa fits particular contexts, such as those that entail a contrastive reading. Furthermore, the separation of the two types of the zero form, even if the obligatory zero could be defined somehow, faces a serious limitation of scope of investigation when one attempts to probe the underlying properties of the zero form.
Also, as mentioned earlier, descriptions of the zero form in terms of its acceptability (or unacceptability) alone would blur the properties associated with speakers’ actual use of the form.

Given the preceding arguments, I should discuss some recent studies which have examined the zero form as it is used by speakers in spoken Japanese.

2.3 Recent text-based analyses of the zero form

2.3.1 Syntactic and semantic factors

Fry (2003) examined the CallHome Japanese corpus consisting of telephone conversations between friends and family members and reported the statistical significance for the following syntactic and semantic factors.

(7) a. Wh-word objects (e.g. nani ‘what’ and dare ‘who’) are more likely to be zero-marked than non wh-word objects (71% vs. 53%).

b. Subjects and objects in questions are more likely to be zero-marked than those in non-questions (45% vs. 31% for subjects, 65% vs. 52% for objects).

c. Subjects and objects in short sentences are more likely to be zero-marked than those in long sentences (41% vs. 27% for subjects, 59% vs. 48% for objects).

d. Multisyllabic subjects are more likely to be zero marked than monosyllabic subjects (34% vs. 21%)

e. Verb-adjacent objects are more likely to be zero-marked than other objects (59% vs. 42%).

f. Animate subjects are more likely to be zero-marked than inanimate subjects (36% vs. 31%).

g. “Strongly definite” subjects (i.e. proper nouns or personal pronouns) are more likely to be zero-marked than “other” subjects (38% vs. 31%). But “other” objects are more likely to be zero-marked than “strongly definite” objects (54% vs. 43%).

h. Grammatically defocused subjects and objects (which co-occur with other elements marked with an emphatic particle, such as mo ‘also’, dake ‘only’, and discourse particles yo and zo) are more likely to be zero-marked than other subjects and objects (37% vs. 32% for subjects, 63% vs. 50% for objects).

Fry offers explanations for some of these correlations. For (7c, length of sentence), for example, he argues that longer sentences exhibit more syntactic complexity, therefore more potential ambiguities. Thus, an overt marking is preferred for disambiguation. Likewise, regarding animacy and definiteness in (7f, g), he argues that inanimate and indefinite subjects are less common and likely to cause ambiguity, which invites the use of overt marking. Despite Fry’s painstaking corpus analyses, several important issues remain unanswered.

First, each of the individual factors in (7a–h) is not strong enough to offer a satisfactory generalization. For example, while the factor in (b) captures only 45% of all
subject zero particles and 65% of all object zero particles in questions, it is not strong enough to predict overt marking in questions (376 tokens; 48% of all question tokens) and the use of the zero particle in non-questions (2426 tokens; 39% of all non-question tokens). Furthermore, it remains unclear how the suggested factors relate to each other and whether there is any underlying property of the zero form that is collectively defined by the variety of factors.

2.3.2 Discourse-functional properties

Contrary to the structural and semantic descriptions as seen in Fry (2003), some recent text-based studies examined discourse-functional properties in the use of the zero form. Suzuki (1995) examined Japanese interviews and suggested that the use of the form is associated with a particular type of topic shift, as stated in (8).

(8)  

**Topic continuity and prominence of referents** (Suzuki 1995)

a. The zero particle signals minor discourse boundaries, in contrast with zero anaphor, which represents topic continuity, and *wa*, which signals major discourse boundaries.

b. The zero particle marks secondary referents so that referents represented by zero anaphor and *wa* are foregrounded.

Suzuki characterizes the zero particle as part of a scale with the two other major encoding forms—zero anaphor and *wa*. As stated in (8a), the zero form represents the mid point of the scale in terms of topic continuity, being situated between zero anaphor, which represents total continuity of a topic, and *wa*, which represents a shift to a new prominent topic. This property correlates with the attenuated level of prominence represented by zero-marked referents, as stated in (8b). Minor discourse topics, initiated at minor topic boundaries, are by definition secondary referents in given contexts.

Suzuki’s claims may be summarized as in (9).

(9)  

<table>
<thead>
<tr>
<th>Referent Type</th>
<th>Topic Continuity</th>
<th>Prominence of Referents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero anaphora</td>
<td>No topic shift</td>
<td>Yes</td>
</tr>
<tr>
<td>Zero-marked phrase</td>
<td>Minor topic shift</td>
<td>No</td>
</tr>
<tr>
<td><em>Wa</em>-marked phrase</td>
<td>Major topic shift</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The attenuated prominence associated with the zero particle has also been suggested by Ono et al. (2000), who examined the use of the form for subject NPs in informal conversation. They observed that, while the nominative case marker *ga* was typically associated with newsworthiness or pragmatic focus, the zero particle typically appeared without such prominence.
Properties of the particle “omission” revisited

(10) Newsworthiness or pragmatic focus with respect to the use of ga and the zero particle (Ono et al. 2000).

Ga tends to resist deletion when the subject NP requires pragmatic focus, namely:
(i) when a subject NP is other than a pronoun or a demonstrative,
(ii) when a subject NP is a complex NP,
(iii) when a subject NP represents ‘exhaustive listing’ and new (unpredictable) information,
(vi) when the identity of a referent represented by the NP-ga is under negotiation.

Overall, the findings reported by Suzuki and Ono et al. point to some sort of subdued prominence exhibited by zero-marked referents in contrast with referents represented in other forms. The assumptions employed by these studies are appropriate for the purpose of the present study in terms of the following respects: (i) analyses should be based on speakers’ use of the zero form, not its acceptability in a given sentence, and (ii) the functional properties of the zero form should be described as a paradigmatic choice vis-à-vis the other available forms. In what follows, I will take up these assumptions and attempt to advance the previous claims and elucidate why the observed properties are associated with the zero particle.

3 Properties of the zero particle revisited

3.1 Grammatical properties—absolute specification of referents

As a starting point, I need to lay out the grammatical properties of the zero marking. Whether the zero particle is optional or obligatory in a given sentence, the zero marking exhibits a particular property in contrast with other possible overt markers. In this regard, I follow the notion of absolute specification of referents claimed by Lee (2002), which is described in (11).

(11) “The grammatical property of the zero particle is ‘absolute specification’, by which the speaker specifies an object or event represented by the NP, without referring to other objects/events” (ibid, 662).

By using the zero form, the sentence subdues the tone achieved by an overt particle which would be used otherwise, such as the contrastive tone achieved with wa and the focus reading associated with ga. Particularly noteworthy is the opposing relationship of the zero particle and wa with respect to contrastiveness, as shown in (12).

(12) Zero particle and wa with respect to ways of referring to other objects/events (Lee 2002).

<table>
<thead>
<tr>
<th>zero particle (absolute specification)</th>
<th>wa (topic)</th>
<th>wa (contrast)</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>implicit</td>
<td>explicit</td>
</tr>
</tbody>
</table>
Unlike *wa*, which denotes a contrastive relationship either explicitly or implicitly, the zero particle denotes complete attenuation of contrastiveness by electing the absence of the overt marking. Consider the example from Lee (2002) given in (13).

(13)  

A: *kono kuruma *Ø *enzin* Ø *ii* desu ne  
this car engine good COP IT  
‘This car, its engine is good.’

B: *syatai wa*/*Ø doo?  
car.body TOP how  
‘What about the body?’

The utterance in (13B) represents an explicit contrastive relationship involving the two referents—‘engine’ and ‘body’. However, the use of the zero particle does not fit such a context as the zero marking attenuates the contrastiveness imposed by the context.

Given the property of the zero particle outlined above, it is expected that the zero particle does not fit a context where *ga* represents an exhaustive listing interpretation, which is exemplified by (14).

(14)  

*matsu* ga/*Ø* itiban hoomuran o *utu  
Matsui NOM most home.run ACC hit  
‘Matsui hits most home runs.’

This sentence implies a context in which Matsui is contrasted with other baseball players, by way of the implied exhaustive listing reading. Again, the use of the zero marking cancels the contrastive link with the other referents excluded, as schematically shown in (15).

(15)  

<table>
<thead>
<tr>
<th>Matsui</th>
<th>X, Y, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>hits most home runs</td>
<td>do not hit most home runs</td>
</tr>
</tbody>
</table>

On the other hand, there is no conflict in the use of the zero particle in the so-called ‘neutral description’ reading. In (16), for example:

(16)  

*matsu* ga/*Ø* nihonn *kuru* yo  
Matsui NOM Japan to come IT  
‘Matsui is coming to Japan.’

This sentence, if given in isolation, does not evoke any contrastive interpretation with other potentially related referents. In other words, the sentence represents absolute specification of the referent and thus fits the property of the zero particle.

To summarize, I spell out the grammatical function of the zero particle as in (17).
(17) *Absolute specification of referents achieved by the zero particle*

The zero particle denotes referents to be specified in an absolute manner by withholding reference to other referents which are potentially related to the proposition denoted by the sentence.

Given this inherent property of the zero particle, how does it relate to speakers’ use of the form? In what follows, I will present text data using quantitative measures and attempt to relate absolute specification of referents achieved by the zero particle to attenuated prominence observed in the use of the particle.

### 3.2 Discourse-pragmatic properties—cataphoric defocusing of referents

#### 3.2.1 Conversation data

The text data discussed in the present study consists of two-party casual conversations between friends (8 pairs of native speakers in total, 10 females and 6 males—mostly from the Kanto area of Japan) which I recorded at the University at Buffalo in 2002. The entire data contains 4 hours of conversations and about 7,900 clausal units (see Shimojo 2005 for detailed discussions of the data).

#### 3.2.2 The five encoding forms and frequency of use

The quantitative analysis in the present section is based on the use of the zero particle for subjects and direct objects in the conversation data, as the zero particle may be used for these two grammatical roles most freely without structural constraints. The results from the quantitative measurement are presented for five commonly observed forms including the zero particle (see Table 1) so that the property of the zero particle may be captured in relation to the other forms.

In terms of overall frequency, it should be noted that the two “zero” forms—zero anaphor and the zero particle—are the two most frequent forms among the five, which reflects the prominence of the zero forms in Japanese conversation. It is particularly noteworthy that the zero particle is noticeably more frequent than the case markers and it is thus the unmarked post-nominal marking in terms of not only the form (i.e. NPs are zero marked) but also the token frequency. The nominative marker *ga* is slightly more frequent than the “topic” marker *wa*. The accusative marker is least frequent; however, this is partly due to the nominative tokens consisting of both transitive and intransitive subjects.

<table>
<thead>
<tr>
<th>Table 1: Frequency of the five encoding types in Japanese conversation (subject and direct object only).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero anaphor</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>3269</td>
</tr>
</tbody>
</table>

---

2 Oblique elements typically require overt post-nominal marking.
3.2.3 Anaphoric saliency

The first discourse property to be discussed is anaphoric saliency because there is a previous claim which associates the zero particle with given or shared information. Table 2 shows the results from the referential distance measurement for each token, counting the number of clauses back to the most recent co-referential expression.\(^3\)

Table 2: The five forms in terms of anaphoric saliency (subjects and direct objects).

<table>
<thead>
<tr>
<th></th>
<th>Zero anaphor</th>
<th>Wa</th>
<th>Zero particle</th>
<th>O</th>
<th>Ga</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salient (RD1-10)</td>
<td>2985</td>
<td>91</td>
<td>199</td>
<td>59</td>
<td>449</td>
</tr>
<tr>
<td>Non-salient (RD11+)</td>
<td>284</td>
<td>9</td>
<td>137</td>
<td>41</td>
<td>480</td>
</tr>
<tr>
<td>Total</td>
<td>3269</td>
<td>100</td>
<td>336</td>
<td>100</td>
<td>929</td>
</tr>
</tbody>
</table>

I arbitrarily define the tokens as salient if their RD value ranges from 1 through 10 and non-salient if their RD is 11 clauses and over. As shown in the table, the index of anaphoric saliency is useful only for zero anaphor and the overt case markers – the former is typically associated with anaphorically salient information and the latter with non-salient information. In terms of this index, *wa* and the zero particle are ambivalent as they may mark either salient or non-salient information, which suggests that the overall usage of the zero particle as well as *wa* is not associated with a particular profile of anaphoric saliency.

3.2.4 Cataphoric saliency

Saliency can be examined in terms of cataphoric context as well—namely, does a given referent continue to be referred to in the subsequent context? Saliency in the sense of cataphoric persistence may be defined in at least two different ways—(i) frequency of mention and (ii) uninterrupted reference. In the former, persistence is defined in terms of how many times a given referent occurs in a certain range of cataphoric discourse (within 10 cataphoric clauses in this study), and in the latter, how far into the cataphoric context there is continuous reference for the referent in question. The index of persistence in either sense suggests how important the referent in question is, since referents that are important for the purpose of the given discourse would occur often and be referred to persistently.

Table 3 summarizes the token counts for cataphoric frequency of reference in the range of 10 cataphoric clauses.

---

\(^3\) See Shimojo (2005) for detailed discussions of referential distance and persistence measurement of the conversation data. See also Givón (1983) for discussions of the basic tenets of the quantitative analyses.
Table 3: Number of clauses containing co-referential expression within 10 cataphoric clauses (cataphoric frequency of reference)

<table>
<thead>
<tr>
<th>RP-f</th>
<th>Zero anaphor</th>
<th>O</th>
<th>Wa</th>
<th>Ga</th>
<th>Zero particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>550</td>
<td>24</td>
<td>81</td>
<td>117</td>
<td>319</td>
</tr>
<tr>
<td>1</td>
<td>492</td>
<td>29</td>
<td>48</td>
<td>83</td>
<td>188</td>
</tr>
<tr>
<td>2</td>
<td>464</td>
<td>11</td>
<td>59</td>
<td>77</td>
<td>137</td>
</tr>
<tr>
<td>3</td>
<td>435</td>
<td>11</td>
<td>37</td>
<td>47</td>
<td>96</td>
</tr>
<tr>
<td>4</td>
<td>328</td>
<td>11</td>
<td>21</td>
<td>37</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>316</td>
<td>6</td>
<td>39</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>236</td>
<td>12</td>
<td>18</td>
<td>25</td>
<td>29</td>
</tr>
<tr>
<td>7</td>
<td>206</td>
<td>3</td>
<td>20</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>132</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>76</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>34</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>3269</td>
<td>115</td>
<td>336</td>
<td>438</td>
<td>929</td>
</tr>
</tbody>
</table>

Mean RP-f 3.2  2.8  2.7  2.3 1.9

This table summarizes the number of tokens for the frequency from zero to 10; for example, for the zero particle (i.e. the right-most column), there were 319 cases in which the referent was never represented again in the cataphoric range, and in 188 cases it was mentioned again only once, etc. On average, referents marked with the zero particle were mentioned 1.9 times, which is the lowest among the five encoding types. Zero anaphor is ranked highest in this measure, with the mean frequency of 3.2 times. The observation is summarized in (18).

(18) Referents marked with the zero particle were least likely to be given again in the cataphoric context among the five encoding types—zero anaphor, o, wa, ga, and the zero particle.

Table 4 presents the token counts for uninterrupted referential persistence, which is indicated by the number of clauses for a given referent to continue to be present within the range of ten cataphoric clauses.
Table 4: *Uninterrupted referential persistence within 10 cataphoric clauses.*

<table>
<thead>
<tr>
<th>RP</th>
<th>O</th>
<th>Wa</th>
<th>Zero anaphor</th>
<th>Ga</th>
<th>Zero particle</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>36</td>
<td>88</td>
<td>629</td>
<td>143</td>
<td>392</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>45</td>
<td>261</td>
<td>92</td>
<td>193</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>29</td>
<td>189</td>
<td>47</td>
<td>79</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>31</td>
<td>144</td>
<td>33</td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>23</td>
<td>94</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>12</td>
<td>64</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>11</td>
<td>44</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>10</td>
<td>29</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>6</td>
<td>32</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3</td>
<td>23</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>14</td>
<td>62</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>104</td>
<td>272</td>
<td>1571</td>
<td>390</td>
<td>770</td>
</tr>
<tr>
<td>Mean RP</td>
<td>3.0</td>
<td>2.6</td>
<td>2.1</td>
<td>1.9</td>
<td>1.1</td>
</tr>
</tbody>
</table>

In 392 cases of the zero particle, the encoded referents did not persist at all (i.e. they were never referred to again), and in 193 cases the referents were mentioned only in the immediately following clause. In terms of mean referential persistence, the zero particle is ranked lowest, as in the case of cataphoric frequency of reference (Table 3). The result here is summarized in (19).

(19) The referents marked with the zero particle were least likely to continue to be present in the cataphoric contexts among five encoding types.

To summarize, while speakers’ use of the zero particle is not satisfactorily described in terms of anaphoric saliency, the index of cataphoric persistence, whether it is defined by frequency of reference or duration of uninterrupted reference, points to attenuation of reference achieved by this encoding type. This property of the zero particle may be rephrased using one’s cognitive focus of attention, as stated in (20).

(20) The zero particle is associated with the effect of defocusing of referents because the attentional focus of the speaker and the hearer is shifted away from the referents as termination of reference occurs in the cataphoric context.

In the conversation data of the present study, the zero particle is used frequently at the termination of referential continuity. An example is given in (21), in which two female speakers are talking about American-style cake.

(21) A382: Ø zibun dake no tameni kau

self only LK for buy

‘(You) buy just for yourself.’
In this part of the conversation, speaker B mentions cake in (383), and then, ‘cake’ continues to be the central topic in the conversation until (391), where ‘cake’ is marked with the zero particle. At this point, the topic shifts from ‘cake’ to ‘someone’s birthday’. In terms of encoding forms, ‘case’ is initially introduced in a full NP, then it switches between the zero anaphor and the overt NP, and eventually the zero particle at the topic shift.
4. Cataphoric defocusing and absolute specification of referents

Table 5 summarizes the discussions given so far, showing both grammatical and discourse properties of the zero particle. The properties of the zero particle are most clearly identified if it is paired with *wa* because the two forms are associated with the same functional scales but complementary functions.

**Table 5: Grammatical and discourse properties.**

<table>
<thead>
<tr>
<th>Manner of referent specification</th>
<th>Discourse function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero particle</td>
<td>Absolute</td>
</tr>
<tr>
<td><em>Wa</em></td>
<td>Contrastive</td>
</tr>
<tr>
<td></td>
<td>Defocusing</td>
</tr>
<tr>
<td></td>
<td>Focusing</td>
</tr>
</tbody>
</table>

Now the question to be raised is how do the grammatical and discourse properties relate to each other? More specifically, how does the absolute or contrastive specification of referents relate to defocusing or focusing of referents in discourse? In this respect, I suggest it is useful to explore the relationship between textual cohesion and contrastivity.

It has been observed that contrastive specification of referents facilitates coherence in text, as cited in (22).

(22) “The primary function of *wa* is to serve as a local cohesive device, linking textual elements of varying degrees of contrastivity” (Clancy & Downing 1987: 46).

In their analyses of story-telling discourse, Clancy & Downing have found that speakers’ use of *wa* is mostly associated with representation of a contrastive relationship of referents and propositions, rather than representation of a global-level theme. They claim that the primary effect brought out in the use of *wa* is textual cohesion achieved by denoting a contrastive relationship between referents and associated propositions, which would otherwise be represented disjointly. By contrasting a referent/proposition with another, the discourse units are connected with each other coherently, making the transition smooth.

The assumption above leads us to the association of the zero particle with disjoint representation of referents and associated propositions since the zero particle denotes absolute specification with total absence of contrastiveness. Furthermore, disjoint representation of referents would have an effect of suppressing cohesion. The functional contrast between *wa* and the zero particle is shown schematically in (23). The marking with *wa* signals that the new state of affairs is presented as further development of the previous one by way of the contrastive relationship denoted. On the other hand, if states of affairs are presented in a non-contrastive manner—i.e., if they are presented as discrete—little textual cohesion is achieved as if the newly given state of affairs is presented in isolation.
If cohesion is achieved by presenting further development of a given state of affairs, cohesion would correlate with cataphoric persistence of associated referents, contained in the current state of affairs. However, in absolute specification of referents achieved by the zero particle, the absence of contrastiveness suppresses cohesion and thus correlates with suppressed textual persistence of given referents.

In fact, in the present conversation data, the zero particle is frequently used to put the main topic on hold in order to interweave isolated side comments. An example is given in (24), in which one of the speakers is describing a pasta recipe.

(24) B198: Ø sosite soko ni eetto kuriimusoosu o irete
and there to F cream.sauce ACC put
‘Then (I) add cream sauce there…’

B199: [beekon itameru]toki ni
bacon panfry time in
‘when (I) panfry the bacon’

B200: Ø siokosyoo mo huttoite
salt.and.pepper also shake
‘(I) sprinkle salt and pepper also…’

B201: dee kuriimusoosu irete
and cream.sauce put
‘and (I) put in cream sauce.’

A202: Ø kekkoo irete siokosyoo-Ø?
quite put salt.and.pepper
‘Did (you) put in much salt and pepper?’

A203: beekon-Ø syoppai yatu?
bacon salty one
‘Is the bacon the salty kind?’

B204: [beekon wa kekkoo syoppakatta kara]
bacon TOP quite was.salty because
‘Because the bacon was quite salty’

B205: demo aan [men ga haitte
but F pasta NOM enter
‘but the pasta is added’

B206: soosu ga haitte]
sauce NOM enter
In the excerpt above, while speaker B describes the recipe, speaker A finds it necessary to interrupt briefly for clarification in (A202–203). A’s interruption puts B’s description of the recipe on hold briefly. The use of the zero particle fits the side comments given by speaker A since speaker A’s interruption does not represent a further development of the preceding state of affairs, as the interruption was made simply to invite clarification of what was given previously.

5. The zero particle in the post-nominal marking system

Given the properties of the zero particle I have proposed thus far, what makes this particular form so important in speakers’ paradigmatic choice of a post-nominal marker? Recall that in the conversation data of the present study, it is the most frequently used form next to zero anaphor. In this respect, I propose the following.

(25) The zero particle plays an important role in the paradigm of post-nominal marking because it is associated with the general—or unmarked—function in referent specification and thus plays the role of globally applicable alternative choice.

The system of post-nominal marking contains an array of encoding forms which are associated with specific functions. The use of case markers—ga and o—is associated with identifiability of referents, as they are used for (re)introduction of referents in order to properly identify previously non-salient referents to be described (see Table 2). On the other hand, the use of wa is connected with contrastiveness and achieves contrastive specification of referents, as discussed earlier. All encoding types mentioned above exhibit a specific or marked function, whether involving a shift of attention to newly identified referents or denoting referents in a contrastive manner. On the other hand, zero anaphor and the zero particle—the two most frequently used forms—exhibit more general and unspecified functions. The former denotes sameness continuing with the previously given referents and the latter specifies referents in isolation without denoting other referents to be contrasted with. Yet, whereas zero anaphor and the zero particle are both “zero” forms and associated with relatively general functions, there is a striking difference between the two. While zero anaphor exhibits cataphoric persistence—i.e. focusing of referents in the subsequent discourse, the zero particle is associated with cataphoric defocusing (Tables 3 and 4).

Overall, the zero particle, as well as zero anaphor, represents an indispensable paradigmatic choice due to its general and default function. Wa is the choice when there
is a need for contrastive specification. However, there is also a need for an alternative choice not to specify contrastiveness. Thus, the general category of the zero particle fills the gap between other functionally more specific choices of the paradigm and thus makes the system of alternatives complete. The situation is analogous to what has been observed in semantically motivated nominal classification systems with gender classes, numeral classifiers, etc., where there are general categories (e.g. the neuter class, a general classifier, etc.) which provide a means to categorize members in a semantically and/or pragmatically unspecified manner.\(^4\) The zero particle in Japanese plays an analogous function in the system of post-nominal marking, as it is used in the place of other alternatives which are associated with pragmatically more specialized functions.

6. Conclusion

The present study revisited the phenomenon commonly known as “particle ellipsis” in Japanese and attempted to probe some fundamental issues remaining to be elucidated.

First, the zero marking, which was previously considered as particle ellipsis by many, has unique properties and is thus best considered as a separate paradigmatic choice, which is more than the mere absence of a particular overt maker. Whether the zero marking is obligatory or optional in a given sentence, the use of the form achieves absolute specification of referents, and this grammatical function is associated with the particular discourse-pragmatic property of referent defocusing. I have also proposed that the two levels of properties—grammar and discourse—are not disjoint but there is a link between them by way of an interplay of contrastive representation of referents and cohesion in discourse.

Furthermore, it is worth noting again that, whether zero anaphor or the zero particle, these “zero” forms are used far more frequently than the overt forms in the representation of argument NPs. The zero forms are unmarked not only in form but also in their functions, and their high frequency in use suggests the cardinal role they play providing the means to represent an argument in a default manner. While the default function of zero anaphor lies in the representation of sameness in the absence of a referential shift, the zero particle denotes a referent without relating it to another and thus provides an alternative to a more specialized manner of referent specification achieved by \(wa\).

References


\(^4\) For example, see Zubin & Köpcke (1986) in this respect for discussions of the neuter class in German. See also Zubin & Shimojo (1993) for cross-linguistic discussions and typology of general categories in nominal classification systems.
MITSUAKI SHIMOJO


