THE DISTRIBUTION OF FILLERS IN LECTURES IN THE JAPANESE LANGUAGE

Michiko Watanabe and Carlos Toshinori Ishi

University of Tokyo, Tokyo, Japan

watanabe@gavo.t.u-tokyo.ac.jp  c_t_ishi@gavo.t.u-tokyo.ac.jp

ABSTRACT

The present study investigated the distribution of five types of fillers in Japanese-language lectures. It was found that the distribution pattern differed depending on the kind of fillers; e, eto and ma tended to appear more often at major syntactic boundaries such as sentence- and clause-boundaries than ano and sono. Sono never occurred at sentence boundaries. These findings suggest that the use of different kinds of fillers may reflect different processes in speech delivery.

1. INTRODUCTION

1.1. General

Fillers are such utterances as uh, um in English and ano, eto in Japanese. They have neither clear grammatical functions nor semantic meanings, but are commonly observed in spontaneous speech. Among Japanese fillers referred to in previous studies are a, ano, de, dessane, e, eto, kono, ma, n, nito, nanka, sono, sodesane, nante iimasuka (what should I say) and word final vowel lengthening ([5], [6], [8], [15]). However what is considered as fillers varies among researchers.

Studies on fillers fall into three general categories; fillers as indicators of mental states and processes, fillers as discourse management device, and fillers as obstacles in speech recognition. In the first area, research has been conducted on relationship between fillers and syntactic boundaries, speaker’s cognitive loads and affective-states ([1], [9]). In the second area, function of fillers in discourse and conversation has been investigated. It has been claimed that fillers contribute to smooth communication by conveying speaker attitude toward the utterance and the interlocutor, as well as by giving information about discourse structure ([11], [13], [15]). The most well-known function pointed out in this field will be that fillers help getting and holding the floor. In the third area acoustic characteristics of fillers have been investigated so that they can automatically be detected for effective speech recognition ([3]). Recent research in this field has also been trying to make use of fillers for natural and cooperative dialogue systems, having taken insights of the second area into consideration ([3], [4]).

1.2. Specific

Across these areas, some researchers have been interested in whether there are distributional or functional differences between different types of fillers. Schreiber reported that um is more typical for sentence initial position than uh in English and noted that the former may be used relatively more often during planning of larger units, and uh may be more likely to reflect local lexical decision-making ([12]). Swerts et al. pointed out that um occurred more frequently at major discourse boundaries than uh in Dutch ([14]). Fox Tree found out that um seems to help listeners recognizing the following word faster, while uh does not ([2]).

As for Japanese fillers, Takubo and Kinsui ([15]) claims that ano and eto have different functions, though they both signal that speaker is in search of memory or in the process of calculation; ano is uttered when the speaker is searching for suitable form of utterance, not the content, while eto is used shortly before or while he is searching for knowledge or doing calculation using his knowledge. Their claim is based on observation that one says eto, not ano, when one is doing calculation or when one is on one’s own, while one says ano, but not eto, when one speaks to somebody to ask to do something; in the former case the speaker is engaged in calculation, and in the latter he is paying attention to the form. Though their observation is correct, to elaborate their model, it seems necessary to clarify how to distinguish the process of searching for knowledge from that of searching for form.

Itoh et al. ([4]) on the other hand, questions Takubo and Kinsui’s argument and claims that there is hardly any functional difference between different kinds of fillers, based on the results of their experiments where they substituted one filler with another and analyzed listeners’ impression. The fillers substituted were rea, e, uh, o, n, ano, eto, ma, iya and de. None of their listeners reported that the speech whose fillers had been substituted with another unnatural. They concluded that functional difference between different types of fillers is subtle enough to be ignored, if any.

It is true that fillers are interchangeable in many cases. However their conclusion seems too hasty. First of all it may be pointed out that one cannot conclude that there is no functional difference because the listeners did not report any unnaturalness about substituted fillers; they may have received a different message from a replaced filler that also fits in the context. Secondly the method of experiments does not seem without problem. The listeners were instructed to write down their comments while or after listening to 30 seconds’ speech. If they wrote comments while listening, they may have missed the following part. If they wrote after listening, they may have forgotten what they felt by the end of the speech.

The present research aims to find out whether there are functional differences between different kinds of fillers in Japanese monologues. As the first step for this aim, the distribution of fillers in lectures was investigated. If distribution patterns differ,
different types of fillers may have different functions.

2. METHOD

2.1. Materials

Excerpts from eight lectures were taken for analysis. All the lectures were freshman-level general courses given at University of Tokyo in April and May, 1999. All the lecturers were native speakers of Japanese. Morphologically all the speakers used standard Japanese which is based on Tokyo Japanese. The age of the speakers ranges from 32 to 55, and two of them were female. The samples were recorded by a DAT (Sony TCD-D100) using a microphone (Sony TCD-100). All the lectures were video-recorded simultaneously for complementary use. An initial two to three minute utterance and a medial two to three minute utterance of each lecture were taken, amounting to 36 minutes of speech.

2.2. Measurements

The samples were digitized and analyzed on a computer. They were transcribed in Japanese orthography and aligned using software packages for sound analysis, xwaves and HTK. Silent periods longer than 100ms, excluding closure of stops, were defined as pauses. In case a voiceless stop followed immediately after short silence, the silence was recognized as a pause when it was longer than 100ms + duration of closure of the same stop in the vicinity. Fillers were defined as segments of speech which did not effect the meaning of the utterance including them but contributed to slowing the delivery of lexicalized information. Though de (a contracted form of sorede, meaning therefore, so, and) was categorized as a type of filler in some studies, it was not regarded as a filler here, because the author viewed that it had meanings and functions as connective. Also, though word-final vowel lengthening was included in fillers in some research, it was not counted as a filler in this study.

Based on the definition mentioned above, fillers were first extracted and their frequency of occurrence was calculated. Then the location of each type of fillers was examined in relation to silent pauses. A unit called an ‘inter-pausal unit’ (IPU), a stretch of speech bounded by pauses was introduced for this purpose. The frequencies of each type of fillers in an IPU-initial position, an IPU-medial position, and an IPU-final position were counted as well as fillers as an independent IPU.

Next, the location of each kind of filler was investigated in syntactic contexts. The number and the ratio of occurrences of each type of filler at sentence boundaries, at clause boundaries, after connectives and after sentence topics ending with wa, a topic particle, were calculated. Definition of clause was based on Minami ([7])’s definition of ‘subordinate phrase’, whose right boundary is marked with one of the following connectives (ga, kara, keredo, shi, node, tara, temo, to, nara, noni, ba, nagara, tsusu) or with te-form (a conjugated form ending with te or de of verbs, adjectives and copulas) or with renyo-form (another conjugated form of verbs and adjectives). These positions are relatively deep syntactic boundaries, where pauses are reported to occur in high probabilities.

3. RESULTS

3.1. General

The material contained 613 fillers in total. Table 1 shows the frequencies of fillers used by each speaker. Each speaker used at least four kinds of fillers. The most frequent fillers in the lectures were ano, e, eto, ma and sono. Hereafter the discussion focuses on these five fillers.

<table>
<thead>
<tr>
<th>Speakers' ID</th>
<th>ano</th>
<th>e</th>
<th>eto</th>
<th>ma</th>
<th>sono</th>
<th>others</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>St</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>13</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>13</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Ln</td>
<td>2</td>
<td>21</td>
<td>26</td>
<td>9</td>
<td>2</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Lw</td>
<td>27</td>
<td>16</td>
<td>10</td>
<td>14</td>
<td>21</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>In</td>
<td>9</td>
<td>49</td>
<td>2</td>
<td>5</td>
<td>46</td>
<td>119</td>
<td></td>
</tr>
<tr>
<td>Ph</td>
<td>13</td>
<td>14</td>
<td>8</td>
<td>31</td>
<td>22</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>Bi</td>
<td>19</td>
<td>0</td>
<td>39</td>
<td>7</td>
<td>0</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Hs</td>
<td>8</td>
<td>17</td>
<td>3</td>
<td>4</td>
<td>13</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>sum</td>
<td>114</td>
<td>121</td>
<td>85</td>
<td>95</td>
<td>91</td>
<td>107</td>
<td>613</td>
</tr>
</tbody>
</table>

Table 1: The frequencies of fillers used by eight lecturers

3.2. The location of fillers in IPU

Next, the location of each kind of filler was investigated in syntactic contexts. The number and the ratio of occurrences of each type of filler at sentence boundaries, at clause boundaries, after connectives and after sentence topics ending with wa, a topic particle, were calculated. Definition of clause was based on Minami ([7])’s definition of ‘subordinate phrase’, whose right boundary is marked with one of the following connectives (ga, kara, keredo, shi, node, tara, temo, to, nara, noni, ba, nagara, tsusu) or with te-form (a conjugated form ending with te or de of verbs, adjectives and copulas) or with renyo-form (another conjugated form of verbs and adjectives). These positions are relatively deep syntactic boundaries, where pauses are reported to occur in high probabilities.

![Figure 1: The percentages of fillers appearing between two pauses (independent), at IPU initial, medial and final positions](image-url)
Figure 1 shows the percentage of fillers in each kind appearing at different locations in IPU. In all, fillers most frequently occurred at IPU initial position (53%). However, the distribution pattern differed depending on the kind; e and eto tended to appear more often at IPU initial position (68% and 66% respectively). They hardly occurred in the middle or at the end of IPU. No eto occurred at IPU final position. Including those consisting of IPU on its own, most e’s and eto’s were preceded by pauses (96% and 86% respectively). In contrast, about half of sono’s were not preceded by pauses. Compared with the average, much higher ratio of sono’s appeared at the end of IPU.

3.3. Fillers at syntactic boundaries

Figure 2 shows the percentage of fillers in each kind appearing at relatively strong syntactic boundaries; at sentence boundaries, at clause boundaries, after connectives and after sentence topics. The total percentages of fillers appearing at one of these positions are given in ‘sum’ column in the table. The sum column shows that about half of e, eto and ma appeared at one of these locations, but that only one forth of sono occurred there. Ano’s ratio of occurrence at one of these places was in between the former and the latter (39%). Sono never appeared at sentence boundaries. It is inferred that sono is more commonly used at minor boundaries such as phrase or word boundaries than at major ones. E occurred far more frequently at clause boundaries than in other positions. Ano occurred much more frequently after wa than in other positions.

4. DISCUSSION

The result is interesting from the point of searching for relationship between fillers and mental processes of speech production. It is widely hypothesized that pauses, including fillers, serve to mark the boundaries between syntactic units because the speaker is making linguistic decisions and need extra processing time.

E and eto appeared more often after pauses and at major syntactic boundaries than the other fillers. This indicates that e and eto are typically used when speaker is planning a larger unit of utterance. E occurred much more frequently at clause boundaries than other fillers. It may be because of te-form which often appears at the end of subordinate clause that filler e is often used at this location. The sound consisting of the same vowel as the preceding clause final vowel must be the easiest for a speaker to utter.

On the other hand, the fact that sono appeared less often after pauses or at strong syntactic boundaries suggests that sono is typically used when speaker has unexpected difficulty in delivering smaller units in the middle of utterance. Sono has a function as demonstrative adjective similar to English the. It seems possible to suppose that the content that speaker tries to express is pointed at with sono. As the content is not yet known to the listener when sono is being uttered, it is perceived as a filler. Sono was not investigated in Itoh et al.’s research. A sense of incongruity may be reported if sentence initial fillers are substituted with sono.

The ratio of ano occurring at major syntactic boundaries was lower than that of e, eto and ma, though not so low as sono. This indicates that ano may be concerned with delivery of rather smaller units of utterance. Like sono, ano has a function as demonstrative adjective similar to English that. Therefore, it seems possible to infer that speaker is pointing at what he tries to express with ano in mind. This explanation is in concordance with Takubo et al’s claim that ano is used when speaker is
5. CONCLUSION

This study has shed some preliminary light on the distribution of fillers in Japanese monologues. Though fillers share quite a few features in common, there were some differences in distribution. *E* and *eto* tended to appear at stronger boundaries more often than the other fillers. *Sono* seemed to occur at minor boundaries more often than the other fillers. *E* was most frequently used at clause boundaries, and *ano* after *wa*, a topic particle. It seems possible to suppose that *e* and *eto* tend to be concerned with planning larger units of utterance, and that *sono* corresponds to delivery of rather smaller units.

This information is useful for learners of Japanese as a second language. Rose ([10]) points out that appropriate use of fillers of the target language improves speaker’s apparent fluency. By using typical fillers at each position, speaker may be able to make his speech sound more native-like. The results can also be adapted to language modeling for speech recognition.

As the next step, we need to examine the context where each kind of filler appears more closely. It will be useful to see fillers in relation to other hesitation phenomena such as word lengthening, repairs and repetition. It will also be interesting to compare the use of fillers in monologues with that in dialogues. Finally it must help understanding the role of fillers more comprehensively to examine the effects of them on listeners’ perception and comprehension of speech.

6. REFERENCES