The Usage of Fillers at Discourse Segment Boundaries in Japanese Lecture-style Monologues

Michiko Watanabe

The University of Tokyo, CREST/JST
watanabe@gavo.t.u-tokyo.ac.jp

Abstract

We examined whether fillers (filled pauses) in a Japanese lecture appeared more frequently after discourse segment boundaries (DSB) than after other sentence boundaries. Contrary to our hypothesis that fillers occur more often after DSB than after other sentence boundaries, the frequency of fillers in the first phrase after DSB did not differ statistically from that after other sentence boundaries. The location of fillers in the first phrase after DSB and after other boundaries did not show any clear difference, either. However, the types of fillers at the initial position of the first phrase after two kinds of boundaries were different; sentence initial 'eto' appeared exclusively at DSB. This result indicates that sentence initial 'eto' may help highlighting DSB, but not other types of fillers. Other kinds of fillers ('e', 'ma', 'ano', 'sono') seem to be mainly concerned with planning units of the utterance that are smaller than a sentence.

1. Introduction

1.1. General

Fillers are such utterances as 'uh' and 'um' in English and 'ano' and 'eto' in Japanese. They have neither clear grammatical function nor semantic meaning, but are commonly observed in spontaneous speech. Among Japanese fillers referred to in previous studies are 'a', 'ano', 'de', 'desune', 'e', 'eto', 'kono', 'ma', 'n', 'nto', 'nanka', 'sono', 'sodesune', 'nante iimasuka (what should I say)', and word final vowel lengthening ([1]-[5]). However, what are considered as fillers vary among researchers.

Research on fillers can be grouped into four categories; 1) one which regards fillers as defects in communication and tries to find out ways to decrease them ([6]); 2) one which tries to find out correspondence between occurrence of fillers and speaker's mental processes or emotional states ([7]); 3) one which admits functions as discourse markers in fillers ([8]); 4) one which tries to find out acoustic characteristics of fillers to allow automatic detection for effective speech recognition ([9]). While there is much research on English fillers, not many studies have been conducted on Japanese fillers except from the forth point of view above. The present research is mainly relevant to the second and the third groups.

[4] and [5] are among the few studies on Japanese fillers relevant to the second and third groups. In this literature the authors claim that interjections and responses are relevant to the speaker's mental operations such as input, output, search, registering and editing of information. Each item, according to them, corresponds to a different kind of processing. By using one of these devices, speaker can not only monitor and help his own processing, but also inform a listener of his mental states, and thus, keep smooth communication. They claim that fillers are concerned with output processes. They have certain features in common, but each of them has functionally different aspects. As for 'eto' and 'ano', according to them, they both hint that the speaker has trouble in output processes and needs more time to continue his speech. However, they differ in that 'eto' is uttered when the speaker is searching for knowledge or conceptualizing ideas using his knowledge, while 'ano' is uttered when he has trouble finding suitable forms for content. Their claim is based on such observations that one utters 'eto', but not 'ano', when one is engaged in calculation, and that one utters 'ano', but not 'eto', when one tries to remember names of things or persons that one knows, or when one asks for a favour politely. Although their view appears correct, their model needs to be tested more empirically and elaborated. It would also be necessary for functions of other fillers to be included in the model.

We have investigated university lectures and speech on academic conferences to find out frequent fillers and their distributions, supposing that different types of fillers have different functions if there is distributional difference ([10], [11]). What we have found so far is as follows;
1. The most frequent fillers in these monologues were 'ano', 'e', 'eto', 'ma' and 'sono'; they covered about 90 % of all the fillers.
2. Among these fillers 'e', 'eto' and 'ma' tended to appear at stronger syntactic boundaries such as sentence- and clause-boundaries more often than 'ano' and 'sono'. 'E 'was most frequent at clause boundaries, and 'ano' after a topic particle 'wa'. 'Sono' hardly ever occurred at sentence boundaries.

These findings indicate that 'e', 'eto' and 'ma' tend to be concerned with planning larger units of an utterance than 'ano' and 'sono'.

1.2. Fillers and Discourse Segment Boundaries

As we mentioned in the previous section, some researchers have investigated whether fillers convey information about discourse structure. Swerts *et al*. ([12]) found that, in Dutch, phrases right after major discourse boundaries contained more fillers than those after minor boundaries, and that fillers after stronger breaks tended to occur at phrase-initial position, while those after weaker breaks at phrase-internal position. They also pointed out that 'um' tended to occur phrase-

initially, whereas 'uh' phrase-internally. They concluded that fillers seemed to carry information about discourse segment boundaries, and that difference in types might reflect different planning processes.

Their research motivated us to investigate into occurrence of Japanese fillers at discourse segment boundaries. Fillers must share some features in common across languages. If fillers are relevant to speech planning, it is reasonable to suppose that they appear more frequently at discourse segment boundaries than at minor boundaries, because the speaker is supposed to do discourse level planning as well as more local planning there.

As is mentioned in the previous section, 'e', 'eto' and 'ma' tended to appear at deeper syntactic boundaries than 'ano' and 'sono', and seem to reflect planning a larger unit of the utterance. As a discourse segment is usually larger than a sentence, it is likely that the former group of fillers tends to occur more frequently at discourse segment boundaries than the latter. As for 'ma', distribution was divergent among speeches on academic conferences ([11]). Therefore, we focused our attention mainly on 'e' and 'eto' in the former group in the present research.

The hypotheses tested in this research were as follows;
1) Fillers appear more frequently in the vicinity of discourse segment boundaries than at other sentence boundaries.
2) Frequent location of fillers in a phrase after discourse segment boundaries differs from that in a phrase after other sentence boundaries.
3) 'E' and 'eto' tend to appear more frequently at discourse segment boundaries than 'ano' and 'sono'.

2. Method

2.1. Material

We used an excerpt from a university lecture as material. The lecture was about international law. It is a part of larger corpus of lectures. We chose this lecture, because the speaker was a native speaker of Tokyo Japanese, and his speaking speed was about an average of the lectures in our corpus.

The lecture was recorded on a DAT (Sony TCD-D100) using a microphone (Sony ECM-717) in a large audience room in university. Forty-one minutes of speech from the beginning of the lecture was taken for analysis.

The excerpt was transcribed in Japanese orthography. The transcription was divided into smaller units bounded by perceptual pauses by the author. The number of units amounted to 1527. Rough average duration of a unit with a following pause is 1.6 seconds. Hereafter, this unit is called an 'inter-pausal-unit' (IPU).

The material contained 682 fillers in total. This means that the speaker uttered a filler every 3.6 seconds on average. Kinds and numbers of fillers which appeared in the sample are shown in Table 1.

Table 1: Kinds and numbers of fillers, and the ratio of each type of fillers in the sample

Fillers	eto	e	ano	sono	ma	Others	Total
Frequencies	13	177	261	65	90	76	682
%	2	26	38	10	13	11	100

2.2. Procedures

Two people participated in segmenting the text. Both were university lecturers. They were instructed to segment the script based on the speaker's purpose of utterance. They were told not to divide an IPU except when they were definitely sure that there was a boundary in it. They were also instructed to write down a purpose of each segment. An example of a segmented text (part of another lecture) was shown before they started segmenting.

2.3. Method of analysis

Labeler A divided the text into 93 discourse segments, and labeler B, 87 segments. The Kappa value of inter-raters' agreement on locations was .63, which we regarded good enough to base our further discussions on.

We call the locations where both labelers marked boundaries 'discourse segment boundaries' (DSB), and each agreed segment a 'discourse segment' (DS). We got 60 DSB. 53 out of 60 DSB accorded with sentence boundaries, and seven with clause boundaries. As most of them were sentence boundaries, we decided to compare distribution of fillers after sentence boundaries of DSB with that after sentence boundaries which neither of the labelers marked as DSB. We call the latter sentence boundaries 'non-discourse segment boundaries' (NDSB). We got 56 NDSB, out of 140 sentence boundaries in total.

First, we examined what percentage of the first IPU after DSB and NDSB included fillers. Second, we investigated what kinds of fillers appeared at initial and non-initial positions in the first IPU after DSB and NDSB. We also examined kinds and frequencies of parts of speech that occurred right after DSB and NDSB.

3. Result

Table 2 shows numbers and percentages of the first IPU after DSB and NDSB that included fillers. There was no statistical difference between these percentages. The first IPU after DSB did not include fillers more often than those after NDSB.

Table 2: Numbers and percentages of the first IPU after DSB and NDSB that included fillers

	IPU with fillers	IPU without fillers	Total
DSB	25 (47%)	28 (53%)	53
NDSB	26 (46%)	30 (54%)	56

Table 3 shows numbers and ratios of fillers at IPU initial and IPU non-initial positions after DSB and NDSB. There is no statistical difference between these ratios.

Table 3: Numbers and ratios of fillers at IPU initial and non-initial positions after DSB and NDSB

	IPU initial	IPU non-initial	Total
DSB	11(44%)	14(56%)	25
NDSB	15(58%)	11(42%)	26

Table 4 shows kinds and numbers of fillers at sentence initial positions right after DSB and NDSB. 'Eto' was the most frequent filler after DSB. All the sentence initial 'eto' appeared at DSB, and none of them at NDSB. 'Eto' occurred more often than 'ano' and 'sono' at DSB, as we had expected. However, contrary to our hypothesis, 'e' occurred only once at DSB. It occurred more often at NDSB.

Table 4: Kinds and numbers of fillers at sentence initial positions after DSB and NDSB

	eto	e	ano	sono	ma	Total
DSB	6	1	3	0	1	11
NDSB	0	6	2	1	6	15

Table 5 shows kinds and numbers of fillers that appeared at non-sentence initial positions in the first IPU after DSB and NDSB. As we can see from the table, there is no clear difference between the distributions of fillers in the two locations.

Table 5: Kinds and numbers of fillers appearing at non-sentence initial positions in the first IPU after DSB and NDSB

	eto	e	ano	sono	ma	Total
DSB	0	0	8	3	3	14
NDSB	0	1	6	3	1	11

Table 6 shows frequencies of parts of speech including fillers that appeared at sentence initial position after DSB and NDSB. Figure 1 shows their ratios. χ^2 analysis revealed that there was statistical difference in the distribution of fillers, connectives and others (including nouns, demonstratives and adverbs) ($\chi^2 = 30.4$, df = 2, p < .00). There was statistical difference between the ratios of connectives and others. However, there was no statistical difference between the frequencies of fillers after DSB and NDSB.

From Figure 1, it is clear that connectives occurred significantly more often at DSB than at NDSB. 75% of the sentences after DSB started with a connective, whereas those starting with a connective after NDSB were 29%.

Table 6: Frequencies of parts of speech (including fillers) appearing at sentence initial position after DSB and NDSB

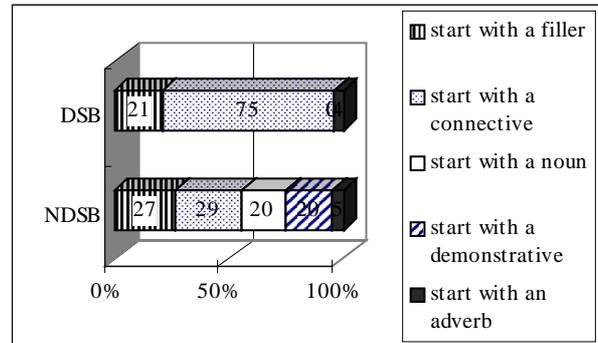|			Others			
	Filler	Connective	Noun	Demonstrative	Adverb	Total
DSB	11	40	0	0	2	53
NDSB	15	16	11	11	3	56

Figure 1: Ratios of parts of speech including fillers that appear at sentence initial position after DSB and NDSB

Table 7 shows the details of connectives that appeared at sentence initial position after DSB and NDSB. 'De' (abbreviated form of 'sorede') was by far the most frequent connective particularly after DSB. About a half of the sentences after DSB (27 out of 53) started with 'de', whereas sentences starting with 'de' after NDSB were 16% (9 out of 56).

Table 7: Connectives at sentence initial position after DSB and NDSB

	DSB	NDSB
de (and)	27	9
sorede (and so)	1	
sorekara (and then)	4	
dakara (so)	4	4
tokorode (by the way)	4	
soredemo (even so)		1
tokoroga (however)		1
aruiwa (or)		1
Total	40	16

4. Discussion

Our hypothesis, 1) Fillers appear more frequently in the vicinity of discourse segment boundaries than at other sentence boundaries, was not supported by the result. IPU

92

right after DSB did not contain fillers more often than those after NDSB. This means that, unlike Dutch, difference in quantity of fillers near boundaries cannot be a marker of DSB in Japanese.

As for the frequent location of fillers, while fillers tended to appear more often at phrase initial position after major boundaries, and phrase internally after minor boundaries in Dutch, we had no such difference in Japanese. Therefore, hypothesis 2) was not supported by the result, either. However, when we looked at types of fillers at phrase initial position after boundaries, 'eto' appeared exclusively after DSB. This means that sentence initial 'eto' can be a marker of DSB.

As for hypothesis 3), 'eto' appeared more frequently at DSB than 'ano' and 'sono', but not 'e'. 'E' hardly ever occurred at DSB. This indicates that 'e' is hardly ever concerned with planning a DS. Based on these results, our hypothesis 3) should be modified to "'eto' tends to appear more frequently at discourse segment boundaries than 'ano' and 'sono'".

The discrepancy of the results of research on Dutch and Japanese fillers may be attributed to differences in the unit of analysis. Swerts et al. based their analysis on 'prosodic phrases', while we based ours on IPU. However, the difference in the first element after boundaries remains to be explained, because a difference in the unit of analysis did not play any role here.

Another possible explanation for the discrepancy of the results is that it derives from difference in speech samples. Although both samples were monologues, Swerts et al. used description of paintings, while we used lectures. Our speech material may well be less spontaneous, because lectures are usually planned in advance and often rehearsed. With more spontaneous speech, results may be different.

5. Conclusions

The present research tested the three hypotheses;
1) Fillers appear more frequently in the vicinity of discourse segment boundaries than at other sentence boundaries.
2) Frequent location of fillers in a phrase after discourse segment boundaries differs from that in a phrase after other sentence boundaries.
3) 'E' and 'eto' tend to appear more frequently at discourse segment boundaries than 'ano' and 'sono'.

Hypothesis 1) was not supported by the result. There was no quantitative difference in frequencies of fillers at the two boundaries. Hypothesis 2) was not supported, either. We did not find any clear difference in the two kinds of phrases. However, frequent types of fillers at initial positions were different. Sentence initial 'eto' exclusively appeared at DSB. Therefore, 'eto' is the only filler that can be a marker of DSB. Hypothesis 3) was true only with 'eto'. 'E' did not appear more frequently at discourse segment boundaries than 'ano' and 'sono'.

From these findings, it is likely that, among the four fillers investigated here, only 'eto' may convey information about DSB, but not the other fillers. The others seem mainly relevant to planning units of an utterance that are smaller than a sentence. Japanese fillers as a whole do not seem to have a lot to do with DSB marking, and discourse level planning in the present research.

The next step will be to increase the quantity of speech samples. Examining more spontaneous speech will help to find out whether Japanese fillers have little relevance to DSB marking and discourse level planning. As most fillers seem to be concerned with planning units of an utterance that are smaller than a sentence, it may be more reasonable to examine the usage of fillers at more local level to establish a comprehensive model of function of each type of fillers.

6. References

[1] Kawamori, K., and Shimazu, A. (1996) "On the form and function of Japanese discourse markers" *Technical Report of IEICE: NCL*96-5, pp. 27-32.
[2] Koide, K. (1983) "Hesitation" in *Lecture Series of Japanese Expression 3: Expressions in Speech* pp. 81-87. Mizutani, Y. (Eds.) Chikumashobou
[3] Nakagawa, S., Kobayasi, S. (1995) "Phenomena and acoustic variation on interjections, pauses, and repairs in spontaneous speech" *Journal of the Acoustical Society of Japan* 51 (3), pp. 202-212.
[4] Takubo, Y. and Kinsui, S. (1997) "The function of responses and interjections in discourse" *Speech and Grammar*, pp. 257-279. Spoken Language Working Group (Ed.) Kuroshioshuppan
[5] Sadanobu, T. and Takubo, y. (1995) "The monitoring devices of mental operations in discourse – a case of 'eeto' and 'ano(o)'-", *Gengo Kenkyu.,* 108: 74-93.
[6] Christenfeld, N. (1996) "Effects of a metronome on the filled pauses of fluent speakers" *Journal of Speech and Hearing Research* 39 (6) pp.1232-1238.
[7] Rochester, S. R. (1973) "The significance of pauses in spontaneous speech" *Journal of Psycholinguistic Research* 2, pp. 51-81.
[8] Schiffrin, D. (1987) *Discourse markers.* Cambridge: Cambridge University Press.
[9] Goto, M., Itou K., and Hayamizu, S. (1999) "A real-time system detecting filled pauses in spontaneous speech" *Information Processing Society of Japan SIG Notes* 99 (64) pp.9-16.
[10] Watanabe, M. and Ishi, C. T. (2000) "The Distribution of Fillers in Lectures in the Japanese Language" *Proceedings of the 6th International Conference on Spoken Language Processing.* Beijing：Volume3 167-170.
[11] Watanabe, M. (2001) "An analysis of usage of fillers in Japanese lecture-style speech" *Proceedings of the Spontaneous Speech Science and Technology Workshop.* Tokyo, pp. 69-76.
[12] Swerts, M., Wichmann, A. and Beun, R. (1996) Filled pauses as markers of discourse structure. *Proceedings of ICSLP* 96, pp. 1033-1036.