

Factors influencing ratios of filled pauses at clause boundaries in Japanese

Michiko Watanabe¹, Keikichi Hirose², Yasuharu Den³, Shusaku Miwa² and Nobuaki Minematsu¹

¹Graduate School of Frontier Sciences, University of Tokyo, Japan

²Graduate School of Information Science and Technology, University of Tokyo, Japan

³Faculty of Letters, Chiba University, Japan

<https://doi.org/10.36505/ExLing-2006/01/0057/000057>

Abstract

Speech disfluencies have been studied as clues to human speech production mechanisms. Major constituents are assumed to be principal units of planning and disfluencies are claimed to occur when speakers have some trouble in planning such units. We tested two hypotheses about the probability of disfluencies by examining the ratios of filled pauses (fillers) at sentence and clause boundaries: 1) the deeper the boundary, the higher the ratio of filled pauses (the boundary hypothesis); 2) the more complex the upcoming constituent, the higher the ratio of filled pauses (the complexity hypothesis). The both hypotheses were supported by filler ratios at clause boundaries, but not by those at sentence boundaries. The results are discussed in light of speech production models.

Introduction

Disfluencies such as filled pauses (fillers) and repetitions are ubiquitous in spontaneous speech, but rare in speech read from written texts. Therefore, they are believed to be relevant to on-line speech production: When speakers have some trouble in speech planning, they tend to be disfluent.

It has been claimed that disfluencies are more frequent at deeper syntactic and discourse boundaries in speech. In early studies it was argued that disfluencies are frequent at the points at which transition probabilities of linguistic events are low and as a consequence the information value is high. Major syntactic boundaries are assumed to be a type of such locations (Maclay & Osgood, 1959). More recent studies have shown that disfluencies tend to cluster at the point at which what the speaker talks about widely shifts (Chafe, 1980), or near the point at which many listeners recognise a boundary (Swerts, 1998). We call the claim that the deeper the boundary, the higher the disfluency ratio *the boundary hypothesis*.

There have been arguments whether speech planning is incremental or hierarchical. Holmes (1995) examined filler ratios at the beginning of basic and surface clauses in English and French, and found no significant differ-

ence in ratios between the two types of locations. Holmes argued that speakers plan one basic clause at one time and that there is no added difficulty even if the upcoming clause contains other clauses. Clark and Wasow (1989), on the other hand, examined repetition rates of articles and pronouns and found that the more complex the following constituents, the higher the ratios. They claimed that the complexity of the following constituents does affect speakers' planning load and consequently the ratio of disfluencies. We call Clark and Wasow's view *the complexity hypothesis*, following their naming. We tested the two hypotheses by examining filler ratios at sentence and clause boundaries in a Japanese speech corpus.

Japanese adverbial clauses are marked by connective particles or certain conjugations of verbs, adjectives or copula markers at the end of the clauses. They always precede main clauses. The clause order of Japanese complex sentences is shown below.

((adverbial clause <connective particle>) (main clause))

Adverbial clauses are classified into three groups according to the degree of dependency on the main clauses (Minami, 1974). *Type A* clauses are the most dependent on the main clauses. Grammatically they can have neither their own topics nor subjects. *Type B* clauses can contain their own subjects, but not their own topics. *Type C* clauses can have both their own topics and subjects. Therefore, *Type C* clauses are the most independent of the main clauses. Consequently, it is assumed that boundaries between *Type C* and the main clauses are deeper than the boundaries between *Type A* or *Type B* and the main clauses and that boundaries between *Type B* and the main clauses are deeper than the boundaries between *Type A* and the main clauses. Therefore, it is predicted from the boundary hypothesis that filler ratios are highest at *Type C* boundaries and lowest at *Type A* boundaries. We considered sentence boundaries as well. As sentence boundaries are assumed to be deeper than clause boundaries, we predict that filler ratios at sentence boundaries are even higher than those at *Type C* clause boundaries.

We predict from the complexity hypothesis that the more complex the upcoming clause, the higher the filler ratio at the boundary. We employed the number of words in the clause as an index of complexity. Details of the experiment are described in the remaining sections.

Method

We analysed 174 presentations (69 academic and 105 casual presentations) in *the Corpus of Spontaneous Japanese* (CSJ) (Maekawa, 2004). The classification of three clause types employed in the present study is described in Table 1. Minami (1974)'s classification was partly modified based on the relevant studies such as Takanashi et al. (2004).

First, we marked *A*, *B* or *C* at each adverbial clause boundary and *D* at sentence boundaries. Then, the number of words in each clause between the boundaries was counted. The clauses were grouped into three according to the number of words in the clause: short (1- 8 words), medium (9-16 words), and long (more than 16 words). Filler rate for each length group of the following clauses at each boundary type was computed for each presentation, and the mean values of the conditions were compared.

Table 1: Classification of adverbial clauses

Type	connective	meaning, usage
A	~nagara, ~tutu	expresses accompanying actions
	~mama	expresses continuous accompanying actions
	~tari, dari	Lists actions or situations
B	~to, ba, tara, nara	if ~
	~te, te kara, te mo	~ and, after ~, even if ~, respectively
	~yoo ni	so that ~
	adverb forms	~and
C	~kara, node	as ~ (reason)
	~noni, ke(re)do	though ~,
	~ga	although ~, ~ but
	~si	~ and (lists similar actions or features)
	~de	~ and
	~masite, ~desite	~ polite auxiliary verb + and

Results and discussions

Type A clauses were excluded from analysis because of low frequency in each presentation and treated as adverbial phrases. Table 2 illustrates mean filler ratios in nine conditions. Because of the space limitation, we describe only the main results. Repeated measures ANOVA showed main effects of the boundary type and the length factor. An interaction between the two factors was significant, $F(4, 680) = 4.02$, $p < .005$. We first compared the ratios by boundary type. For *Type B* boundaries, the filler ratio was higher, the longer the following clause, long vs. short: $t(170) = 5.18$, $p < .001$; long vs. medium: $t(170) = 2.95$, $p < .05$; medium vs. short: $t(170) = 3.08$, $p < .007$. For *Type C* boundaries, only the filler ratio before long clauses was significantly higher than that before short clauses, $t(170) = 2.75$, $p < .05$. For *Type D* boundaries, there was no significant difference among length factors, $F(2, 169) = .25$, $p = .78$. When we compared the ratio by length group, there were significant differences among boundary types in all the length groups. Paired comparisons showed that in all the length groups, the ratios of fillers at *Type C* and *Type D* boundaries were significantly higher than those at

Type B boundaries, but that there were no significant differences between the ratios at *Type C* and *Type D* boundaries.

The complexity hypothesis was supported by the results for *Type B* boundaries, and also for *Type C* boundaries with less degree, but not supported by the results for *Type D* boundaries. The boundary hypothesis was supported by the difference between *Type B* and *Type C* boundaries and the difference between *Type B* and *Type D* boundaries, but there was no significant difference between *Type C* and *Type D* boundaries.

We speculate that these results derive from difference in the most influential factors at different types of boundaries. At deeper boundaries such as *Type C* and *Type D* boundaries, most of speakers' attention and time tend to be devoted to message conceptualisation. As a consequence, speakers cannot plan linguistic units far ahead at those points and the complexity effects are relatively small. In contrast, at shallower boundaries at which cognitive loads for conceptualisation are not so heavy, the complexity effects seem to play a significant role.

Table 2: Rate of clause boundaries with fillers (%) before short (1-8 words), medium (9-16 words) and long (over 16 words) clauses.

Boundary type	1-8 words	9-16 words	17- words
B	26	29	33
C	39	40	43
D	40	39	40

References

- Chafe, W. 1980. The deployment of consciousness in the production of a narrative, in Chafe, W. (ed.) *The Pear Stories, Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, 9-50. New Jersey, ALEX Publishing Corporation.
- Clark, H. H. & Wasow, T. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37, 201-242.
- Holmes, V. M. 1995. A crosslinguistic comparison of the production of utterances in discourse. *Cognition*, 54, 169-207.
- Levelt, W. J. M. 1989. *Speaking*. The MIT Press, Cambridge, Massachusetts.
- Maclay, H., & Osgood, C. E. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15, 19-44.
- Maekawa, K. 2004. Outline of the Corpus of Spontaneous Japanese. In Yoneyama, K. and Maekawa, K. (eds.) *Spontaneous Speech: Data and Analysis*. National Institute for Japanese Language.
- Minami, F. 1974. *Gendai nihongo no kouzou (The structure of modern Japanese)*. Taisyukan syoten, Tokyo.
- Swerts, M. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30, 485-496.
- Takanashi, K., Uchimoto, K., & Maruyama, T. 2004. Identification of clause units in CSJ. In vol1, the Corpus of Spontaneous Japanese.