

Features of Pauses and Conjunctions at Syntactic and Discourse Boundaries in Japanese Monologues

Michiko Watanabe¹, Yasuharu Den², Keikichi Hirose³, Shusaku Miwa³, Nobuaki Minematsu¹

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

²Faculty of Letters, Chiba University, Japan

³Graduate School of Information Science and Technology, The University of Tokyo, Japan

{watanabe, hirose, shusaku, mine}@gavo.t.u-tokyo.ac.jp, den@cogsci.1.chiba-u.ac.jp

Abstract

Syntactic and discourse boundaries are signalled by prosodic cues as well as linguistic cues in speech. We investigated whether there is a correspondence between prosodic or linguistic cues and the boundary strengths. We measured the rates of filled pauses (FPs) and conjunctions, and the durations of silent and filled pauses and conjunctions at four types of boundaries in casual presentations in Japanese. The results showed that the rates of FPs and conjunctions and the durations of silent pauses correspond to the boundary strengths. However, no significant correspondence was found between the duration of FPs or conjunctions and the boundary strengths. The results suggest that how long the speaker pauses and whether he or she utters a FP or a conjunction is relevant to the boundary strengths. However, the durations of FPs and conjunctions are likely to be affected by the other factors such as planning difficulties of the following parts of speech.

Index Terms: syntactic boundary, discourse boundary, silent pause, filled pause, conjunction

1. Introduction

Prosodic cues are known to be relevant to syntactic and discourse structure across languages. Sentence boundaries and topic shifts are often marked by a long pause, a pitch range reset, change in intensity and speaking speed, and disfluencies [1], [2], [3]. It is reported that some prosodic features correspond to boundary strengths in speech. The deeper the discourse boundary, the longer the silent pause duration [4]. The rate of filled pauses increases with syntactic or discourse boundary strengths [4], [5], [6]. However, most previous studies analysed syntactic or discourse boundaries separately. The two types of boundaries were not well integrated in the analyses. We assume sentences as components of a discourse unit and clauses as components of a sentence. As a consequence, we regard discourse boundaries as stronger than sentence boundaries, and sentence boundaries as stronger than clause boundaries. Linguistic items such as connectives, adverbs and adverbial phrases also signal boundaries [7], [8]. However, it has not been elaborated on whether and how the existence of such linguistic items and prosodic features at the boundaries interact.

In this paper, we investigated the correspondence between boundary strengths and prosodic and linguistic features at the boundary. More specifically, we examined the rates and the durations of filled pauses and conjunctions as well as the duration of silent pauses at four types of boundaries in Japanese monologues. We compare the results with those of previous

studies, and discuss relevance of prosodic and linguistic features to discourse structure.

2. Method

2.1. Data

Twenty-four casual presentations excerpted from *The Corpus of Spontaneous Japanese* (CSJ) were analyzed [9]. The corpus comprises sound files, transcripts and morphological analyses of 660 hours of 3302 speeches with 7.7 million words [10]. The presentations examined were given by speakers of Tokyo Japanese, aged between the 20s and the 60s. They were paid volunteers. A half of them were male and another half female speakers. They talked about a general topic such as “the happiest (saddest) experience in my life” or “my town”. The topics were given to the speakers previous to the talk. They were instructed to prepare a note for the presentation, and not to read aloud a manuscript. Each presentation lasted between 10 and 15 minutes.

2.2. Procedures

First, we classified sentence and adverbial clause boundaries into three types according to sentence or clause final morphemes as shown in Table 1. Japanese is a SOV language and the ends of sentences are marked by sentence final forms. The ends of clauses are marked by connective particles or certain conjugations of verbs, auxiliary verbs, or adjectives. As shown below, adverbial clauses always precede main clauses.

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

Among adverbial clauses, C type clauses are regarded as more independent of main clauses than B type clauses in terms of syntactic constraints [11], [12], [13]. Therefore, we assumed boundaries after C type clauses as stronger than boundaries after B type clauses.

The presentations examined contained discourse boundary tags. Each of three labellers marked paragraph boundaries at some of the sentence or clause boundaries when they recognised a topic shift. Discourse boundaries were chosen from all the paragraph boundaries through discussion by four discourse analysts, based on the theory of Grosz and Sidner [14]. Most discourse boundaries coincided with sentence boundaries. We call sentence boundaries which are also discourse boundaries E type boundaries and the other sentence boundaries D type

boundaries. We assumed E type boundaries as stronger than D type boundaries.

Second, we computed the rate of boundaries with filled pauses (a prosodic cue) or with conjunctions (a linguistic cue) for each type of boundaries in each presentation. We compared the mean rates of boundaries with filled pauses or with conjunctions among four boundary types.

Table 1: Classification of adverbial clauses

Type	Connective	Meaning, usage
B	~to, ba, tara, nara	when ~, if ~
	~te (wa)	~ and
	~te kara	since ~, after~
	~te mo	even if ~
	~yoo ni	so that ~
	adverbial form	~ and
C	~kara, node	as ~ (reason)
	~noni, ke(re)do	though ~
	~ga	although ~, ~but
	~si	~ and (list similar actions or features)
	~de	~ and
	~masite, desite	~ auxiliary verb expressing politeness + and
sentence	~yo, ne, to	sentence-final particles
(D or E)	~desu, masu, ta, n	sentence-final forms

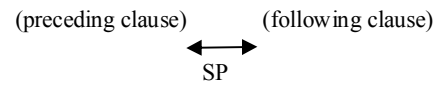
Third, we measured the durations of conjunctions and silent and filled pauses at four types of boundaries in the following four patterns. The durations measured are schematized in Figure 1.

- 1) When there was neither a filled pause nor a conjunction between two consecutive clauses, the silent pause duration between the end of the previous clause and the onset of the following clause was measured (the value can be zero).
- 2) When there was a filled pause between two consecutive clauses, silent pause duration between the end of the previous clause and the onset of the voiced part of the filled pause and the duration of the filled pause (from the onset of the voiced part of the filled pause up to the onset of the following clause) were measured.
- 3) When there was a conjunction between two consecutive clauses, silent pause duration between the end of the previous clause and the onset of the conjunction and the duration of the conjunction (from the onset of the conjunction up to the onset of the following clause) were measured.
- 4) When there were both filled pauses and conjunctions between two consecutive clauses, the silent pause durations between the end of the previous clause and the onset of the conjunction or the voiced part of the filled pause, and the durations of the conjunction and the filled pause were measured.

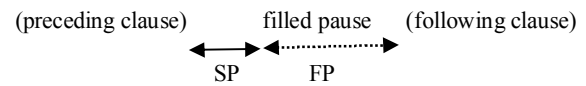
We took the median of the durations in each category for each presentation, and compared the mean values over presentations among four types of boundaries. In this paper, we analyzed the durations of silent and filled pauses and conjunctions of the first three patterns.

CON: conjunction; FP: filled pause; SP: silent pause

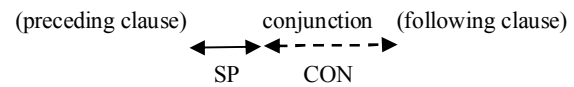
1) Boundaries without filled pause or conjunction (pattern 1)



2) Boundaries with filled pauses (pattern 2)



3) Boundaries with conjunctions (pattern 3)



4) Boundaries with filled pauses and conjunctions (pattern 4)

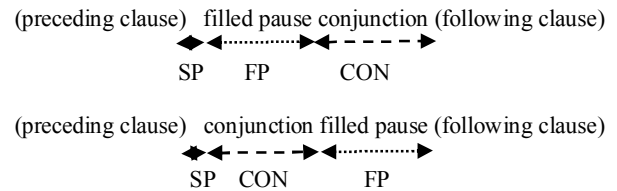


Figure 1: Four patterns of clause boundaries at which durations of conjunctions and silent and filled pauses were measured

3. Results

3.1. The rates of filled pauses and conjunctions

Table 2 shows the mean number of each boundary type.

Table 2: The mean number of each boundary type

Type	B	C	D	E
mean N	56	62	37	12

Figure 2 illustrates the rates of filled pauses and conjunctions at four types of boundaries. The rate of filled pauses increases gradually whereas the rate of conjunctions increases linearly with the boundary strengths. A one-way repeated measures analysis of variance (ANOVA) showed a significant effect of boundary type on the rate of filled pauses, $F(3, 69) = 17.63, p < .001$. Paired comparisons (alpha adjusted by Bonferroni) revealed that the rate of filled pauses at type E boundaries was significantly higher than the rate at any other type of boundaries, B vs. E: $t(23) = 5.48, p < .001$; C vs. E: $t(23) = 3.85, p < .005$; D vs. E: $t(23) = 4.62, p < .001$. The rate of filled pauses at C type boundaries was significantly higher than the rate at B type boundaries, $t(23) = 4.80, p < .001$. There was no significant difference between the rates at D type and B type boundaries, or D type and C type boundaries, B vs. D: $t(23) = 1.64, p = .69$; C vs. D: $t(23) = 1.94, p = .39$. The rate of filled pauses at clause and discourse boundaries corresponded to the boundary strengths. However, the rate of filled pauses at sentence

boundaries was not significantly different from the rates at clause boundaries.

Regarding the rate of conjunctions, a one-way repeated measures ANOVA showed a significant effect of boundary type, $F(3, 69) = 90.49, p < .001$. Paired comparisons (alpha adjusted by Bonferroni) revealed that the stronger the boundary, the higher the rate of conjunctions. The rate at E type boundaries was significantly higher than the rate at any other type of boundaries, B vs. E: $t(23) = 14.18, p < .001$; C vs. E: $t(23) = 11.64, p < .001$; D vs. E: $t(23) = 5.63, p < .001$. The rate at D type boundaries was significantly higher than the rates at C type and B type boundaries, B vs. D: $t(23) = 8.81, p < .001$; C vs. D: $t(23) = 4.98, p < .001$. The rate at C type boundaries was significantly higher than the rates at B type boundaries, $t(23) = 7.37, p < .001$. The rate of conjunctions more clearly corresponded to the boundary strengths than the rate of filled pauses did.

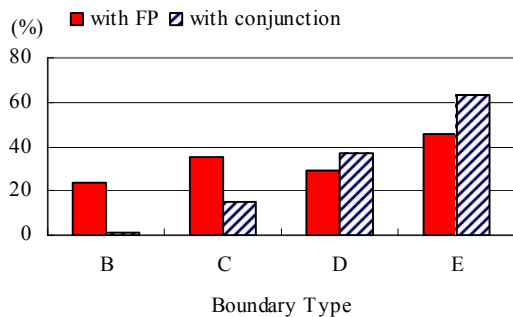


Figure 2: The rate of boundaries with filled pauses (FP) and with conjunctions

3.2. The duration of silent pauses

Figure 3 illustrates the durations of silent pauses when there was neither filled pause nor conjunction (the left bar at each boundary type) and when there was a filled pause (the right bar). A two-way repeated measures ANOVA (boundary type * with or without filled pause) showed a main effect of boundary type, $F(3, 42) = 10.44, p < .001$. There was no significant main effect of with or without filled pause condition, $F(1, 14) = 2.47, p = .14$. There was no significant interaction between the two factors either, $F(3, 42) = .940, p = .39$.

Paired comparisons (alpha adjusted by Bonferroni) revealed that the silent pause durations at E type boundaries were significantly longer than the durations at any other type of boundaries, B vs. E: $t(14) = 4.21, p < .001$; C vs. E: $t(14) = 3.87, p < .01$; D vs. E: $t(14) = 42.66, p < .05$. The pauses at C type boundaries were significantly longer than the pauses at B type boundaries, $t(14) = 2.17, p < .05$. The pauses at D type boundaries tended to be significantly longer than the pauses at B type boundaries, $t(14) = 2.11, p = .05$. There was no significant difference between the durations at C type and D type boundaries, $t(14) = 1.71, p = .11$.

Similar to the rate of filled pauses, the duration of silent pauses between two consecutive clauses positively corresponded to the boundary strengths of clauses. However, the

pauses at sentence boundaries were not significantly longer than the pauses at B type or C type clause boundaries.

Figure 4 illustrates the durations of silent pauses when there was neither filled pause nor conjunction (the left bar) and when there was a conjunction (the right bar). As the rate of conjunctions at B type boundaries was low (1%), we compared the pause durations among the other three boundary types. A two-way repeated measures ANOVA (boundary type * with or without conjunction) showed a main effect of boundary type, $F(2, 30) = 5.86, p < .01$. There was no significant main effect of with or without conjunction condition, $F(1, 15) = 3.73, p = .07$. There was no significant interaction between the two factors either, $F(2, 30) = .99, p = .38$.

Paired comparisons (alpha adjusted by Bonferroni) revealed that the silent pause durations at E type boundaries were significantly longer than the durations at C type boundaries, $t(15) = 3.08, p < .05$. There was no significant difference either between C type and D type or D type and E type boundaries, C vs. D: $t(15) = 1.71, p = .32$; D vs. E: $t(15) = 1.91, p = .23$.

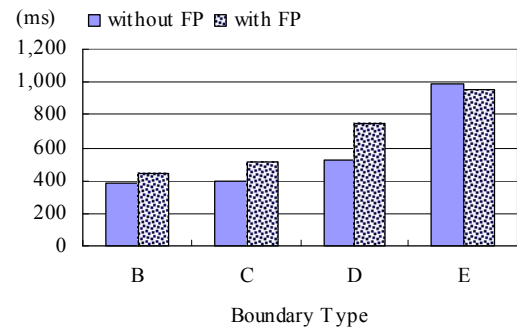


Figure 3: The duration of silent pauses with and without filled pauses followed at four types of boundaries

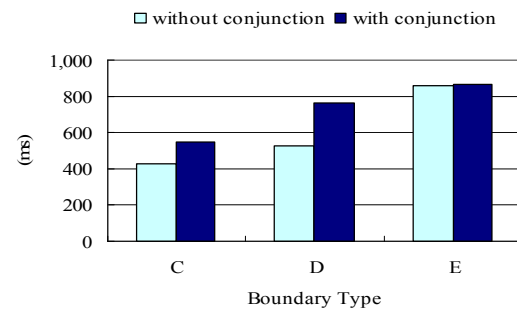


Figure 4: The duration of silent pauses with and without conjunctions followed at three types of boundaries

3.3. The durations of filled pauses and conjunctions

Table 3 illustrates the durations of filled pauses at four types of boundaries. A one-way repeated measures ANOVA showed no significant effect of boundary type, $F(3, 57) = 3.96, p = .06$, although the filled pauses at E type boundaries tended to be longer than the others.

As the rate of conjunctions at B type boundaries was low (1%), we excluded them from the analysis and compared the durations of conjunctions at the other three boundary types.

Table 4 shows the mean durations of conjunctions at three types of boundaries. A one-way repeated measures ANOVA showed no significant effect of boundary type, $F(2, 38) = 2.18, p = .13$.

Table 3: *The duration of filled pauses at four types of boundaries*

Boundary type	B	C	D	E
Filled pause (ms)	274	248	278	804

Table 4: *The duration of conjunctions at three types of boundaries*

Boundary type	C	D	E
Conjunction (ms)	211	160	226

4. Discussion

The rates of filled pauses and conjunctions both corresponded to the boundary strengths. However, the conjunction rates showed clearer correspondence to the boundary strengths than the filler rates. The filler rate at D type boundaries was not as high as expected from its boundary strength. Linguistic cues (conjunctions) rather than prosodic cues (filled pauses) seem to play a more significant role in signalling boundary strengths.

The duration of silent pauses corresponded to boundary strengths. The stronger the boundary, the longer the silent pauses tended to be. This finding is consistent with [4]. However, silent pauses at D type boundaries were not significantly longer than the pauses at clause boundaries. It is possible that weak sentence boundaries are not much stronger than C type clause boundaries in terms of the content of speech.

The durations of filled pauses showed no correspondence to boundary strengths except for the case with E type boundaries. Filled pauses at E type boundaries tended to be longer than the others. Once filled pauses are uttered, their durations may be affected by factors other than the boundary strength itself. The degree of planning difficulties of the following parts of speech can be one of the factors.

Research on Dutch filled pauses showed that the F0 values of filled pauses tended to be higher at stronger boundaries than at weaker boundaries, although their durations did not significantly differ between the two boundary types. Differences may be found in F0 values of filled pauses or conjunctions according to boundary strengths in Japanese as well as in Dutch.

5. Conclusions

We investigated the correspondence between boundary strengths and some linguistic and prosodic features at syntactic and discourse boundaries in Japanese spontaneous monologues. The rate of filled pauses and conjunctions as well as the silent pause durations corresponded to boundary strengths. Therefore, they can be cues to boundary strengths. However, the durations of filled pauses and conjunctions hardly showed any correspondence to boundary strengths.

As a next step, we plan to investigate the F0 values of filled pauses and conjunctions at four types of boundaries. We also plan to examine the correspondence between the prosodic features of boundaries and the complexity of the following speech.

6. Acknowledgements

This study was partly supported by the 21st Century COE Program in Electronics for Future Generations, The University of Tokyo, and The HAKUHO FOUNDATION.

7. References

- [1] Chafe, W., 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, ALEX Publishing Corporation, Norwood, New Jersey, 1980.
- [2] Koiso, H., Shimojima, A., and Katagiri, A., "Collaborative Signaling of Informational Structures by Dynamic Speech Rate," *Language and Speech* Vol. 41, No. 3 - 4, 323 - 350, 1998.
- [3] Shriberg, E., Stolcke, A., Hakkani-Tür, D., and Tür, D., "Prosody-based automatic segmentation of speech into sentences and topics", *Speech Communication*, 32, 127-154, 2000.
- [4] Yoneyama, K., Fon, J., and Koiso, H. "Durational and prosodic patterning at discourse boundaries in Japanese spontaneous monologs", *Proc. 15th International Congress of Phonetic Sciences: 2637-2640*, Barcelona, Spain, 2003.
- [5] Swerts, M., Wichmann, A. and Beun, R. Filled pauses as markers of discourse structure, *Proc. ICSLP 96: 1033-1036*, Philadelphia, PA, 1996.
- [6] Watanabe, M., Hirose, K., Den, Y. and Minematsu, N., Factors influencing ratios of filled pauses at clause boundaries in Japanese, *Proc. ISCA Tutorial and Research Workshop on Experimental Linguistics*, 253-256, Athens, Greece, 2006.
- [7] Schiffrin, D., *Discourse markers*, Cambridge University Press, Cambridge, 1987.
- [8] Watanabe, M., Fillers and Connectives as Discourse Segment Boundary Markers in an Academic Monologue in Japanese, *Bulletin of International Center* Vol.12, 107-119, International Center, The University of Tokyo, 2002.
- [9] The National Institute for Japanese Language, The Communications Research Laboratory, *The Corpus of Spontaneous Japanese*, 2004.
- [10] The National Institute for Japanese Language, Homepage of *The Corpus of Spontaneous Japanese*, http://www2.kokken.go.jp/~csj/public/6_1.html, retrieved in April, 2006.
- [11] Minami, F., *Gendai Nihongo no Kouzou* (The Structure of Modern Japanese), Taisyukan Syoten, Tokyo, 1974 (in Japanese).
- [12] Takubo, Y., Syntactic structure and contextual information, *Nihongogaku*, 6 (5), 37-48, 1987 (in Japanese).
- [13] Watanabe, M., Den, Y., Hirose, K., and Minematsu, N., "Clause types and filled pauses in Japanese spontaneous monologues", *Proc. 8th ICSLP: 905-908*, Jeju Island, Korea, 2004.
- [14] Grosz, B., and Sidner, C., "Attention, Intentions, and the Structure of Discourse." *Computational Linguistics*, 12, 3, 175-204, 1986.