



Phrase-final rise-fall intonation and disfluency in Japanese: - A preliminary study -

Jumpei Kaneda

Graduate School of Human Science, Kobe University, Japan

Abstract

In Japanese conversations, rise-fall intonation with vowel lengthening often occurs on the final syllable of a phrase. This phrase-final rise-fall (PFRF) is a new type of intonation first reported in the 1960's. Researchers consider PFRF intonation a discourse marker which functions to sharpen the phrase boundary and retain the utterance turn, but other phrase-final intonation such as phrase-final lengthening (PFL) can have a similar pattern. PFLs are recognized as a type of disfluent speech with similar characteristics to PFRFs in terms of final-lengthening and having discourse functions. Also from reports about the spontaneity of speech, we assume that PFRFs would have a relation with disfluency, as well as with PFLs. To examine this assumption, this paper attempts to show the co-occurrence relation between PFRF and disfluency in the same utterance. The results show that PFRFs and PFLs have a relation to posterior disfluent units and suggest that both indicate speech planning strategies. Further, this paper speculates that a difference between PFRF and PFL is a difference in the purposes of speech planning: the latter represents ongoing linguistic editing while the former indicates adjusting the utterance according to the interlocutor's reaction. Disfluencies accordingly occur as effects from processes of speech planning.

1. Introduction

In Japanese spontaneous casual speech, it is often observed that the final syllable of an utterance non-final phrase (more precisely “intermediate phrase” [1]) is pronounced with a rise-fall intonation and the final vowel is more or less lengthened. This phenomenon has been called *shiriagari* “final-rising” intonation [2] or *shoko-cho* “rise-fall intonation” [3], but this paper calls this intonation “**phrase-final rise-fall** (later **PFRF**)”. An example is shown in Figure 1 below.

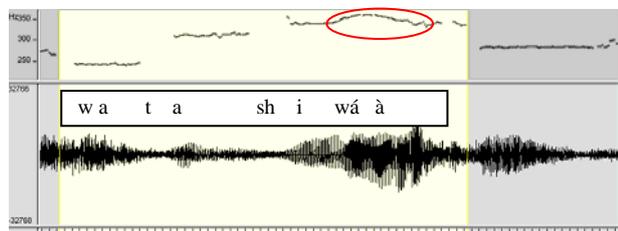


Figure 1: Pitch contour of *watashi-wa* “as for me” with a phrase-final rise-fall intonation (see the circled area)

Some researchers say a PFRF is used in order to attract the interlocutor's attention, and others say it has discourse functions (see Section 2.2). Although PFRF is not successfully explained yet, we can say that a PFRF shows at least a modality toward the interlocutor.

A PFRF always occurs on the end of a phrase, not a word. In Japanese, a head-final language with postpositions, *watashi-wa* is a single phrase formed by a noun *watashi* “me” and a topic marker *wa* as head of the phrase. A PFRF thus, in this

case, occurs on the final syllable, the particle *wa*. Meanwhile, a PFRF seems to be a “postposition (particle) prominence” for the purpose of emphasizing the postposition, but we claim it is not. The reason is because a noun phrase with only a monosyllable postposition particle such as *wa*, *ga* and *o* containing a PFRF would indicate postposition prominence, but this is different from the case of a two-syllable postposition particle as illustrated below.

- (1) Koobe-kara (ki-mashi-ta.)
(I came) from Kobe.
a. koobe-kara
b. koobe-karà

Example (1) shows a contrast in intonation and meaning. Bold letters indicate that the syllable is pronounced with prominence (high pitch) and *koobe* has a lexical accent on the first syllable. (1)a is an example of postposition prominence, that emphasizes the meaning of the postposition particle when contrasting with other particles (not “to” but “from”). In this case the particle *kara* is accented on its first syllable independently from *Koobe*, but never on the second. This is phonologically different from (1)b, in which a PFRF appears on the last syllable. As for a difference in meaning, (1)a sounds emphasized on the particle while (1)b does not. In addition, a PFRF can appear on zero-particle NPs, adverbials, and VPs (not in the sentence-final position); this is another reason for differentiating PFRFs from postposition prominence. A PFRF can thus be said to be an intermediate phrase-level intonation phenomenon.

2. Background

2.1. History of PFRF

PFRF is a comparatively new phrase-level intonation phenomenon. According to Akinaga (1966) PFRFs existed in the 1960's at the latest [4], being used mostly by young women and children in the Kanto district [3]. Today the use of PFRF has spread throughout Japan as well as among male speakers. Teachers or lecturers also tend to use PFRFs since PFRFs are considered to be effective to help pupils understand when explaining things.

On the other hand, PFRFs are stigmatised by some older or conservative speakers of Japanese saying that it sounds rude, flippant or childish [2][3], and even among the younger generation, speakers who frequently use PFRFs are sometimes evaluated negatively [5].

2.2. Discourse functions of PFRF and their problem

Apart from sociolinguistic features mentioned in the previous section, some researchers affirm that PFRFs have pragmatic or discourse functions. In this section we see two functions: clarifying phrase boundaries (demarcating) and retaining the speaker's turn. As for the former, Sugito (1983) affirms PFRF's demarcating function [2]. The turn-keeping function is explained in Inoue (1997); PFRFs indicate continuity of the utterance and prevent the other speaker from interrupting [3].

The notion that a PFRF is a discourse marker is, however, incorrect. The above two functions are not inherent to PFRFs. With a normal (no rise-fall) phrase-final intonation, pauses in sentence-medial position can clarify the boundaries of phrases and keep the utterance turn. Postposition prominences (see Section 1) have these two functions, also. This implies that PFRFs originate from boundary (extra-)linguistic factors other than discourse. A possible extra-linguistic phenomenon is **disfluency**.

2.3. Disfluency and PFRFs

There is another phrase-final intonation, **phrase-final lengthening** without falling intonation (referred to as **PFL**). Figure 2 below is an example of PFL. The final vowel is lengthened and the pitch stays high.

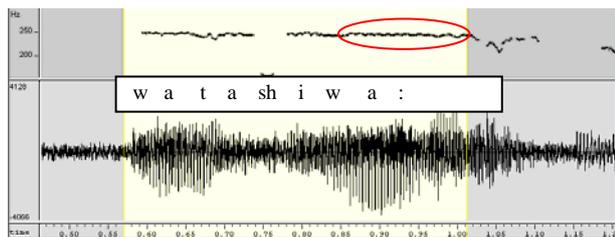


Figure 2: Pitch contour of “*watashi-wa* (as for me)” with a phrase-final lengthening (see the circled area).

A PFL intonation has two characteristics: one is a representation of disfluency, the other, a discourse function. It is thought that a PFL appears when the speaker hesitates to utter the following part of the utterance (cf. [6]). That is to say, a PFL occurs at the moment speech planning involving word selection, calculation, or recalling something is not completed. In this sense, phrases with a PFL intonation are similar to **fillers** such as *eeto* “let me see”, *anoo* “uhm” and *nanka* “like.” As for discourse functions, a PFL also has both demarcating and turn-keeping functions, mentioned in Section 2.2. Phrase boundaries are defined by a final vowel lengthening and the utterance turn does not move because the utterance is still incomplete.

A PFRF is similar to a PFL with regard to two points. First, both have some sort of final vowel lengthening. Second, the two discourse functions, demarcating and turn-keeping, are associated with both intonation patterns. The similarity implies that a PFRF intonation tends to appear in disfluent speech as well as does a PFL. For supporting this implication, the National Institute of Japanese Language shows that spontaneity of speech correlates with co-occurrence of disfluencies (e.g. fillers and fragments of word) and PFRF intonation [7]. In other words, in highly spontaneous speech such as casual conversations, disfluencies and PFRFs co-occur frequently. Although these data do not show that disfluency has a direct relation with PFRF, they at least suggest a possible relation between them.

2.4. Hypothesis

Based on the similarity of PFRF and PFL, and an expectation of a relation among spontaneity, disfluency and occurrence of PFRF shown in the previous section, this paper sets up a hypothesis that the PFRF intonation is a representation of disfluency. It is, however, difficult for this paper to fully verify this hypothesis, but we can test the expectation mentioned in the bottom of Section 2.3 that PFRFs have a relation with disfluency. If their relation is recognized, it will partially support our main hypothesis. This paper will thus test the relation between PFRF and disfluency by making use of the corpus of spontaneous dialogues as described below.

3. Method

3.1. Corpus

From a fifty-hour spontaneous dialogue corpus recorded during October and November 2004, utterances from 5 female informants between the ages 27 and 45 with a total amount of 15 hours were selected.

3.2. Method

The samples were utterances containing a pronominal expression *watashi* (sometimes pronounced with the first consonant dropped like *atashi*) “me” as topic or argument, but not as predicate. The samples were classified into 3 categories according to phrase-final intonation patterns and checked to see if there were co-occurrences of disfluency units such as otiose pauses and fillers in the same utterance. The co-occurrence check consisted of two parts, (i) presence of excessive pauses or fillers immediately before or after *watashi(-wa)* and (ii) co-occurrence of these disfluency factors in an anterior or posterior position to *watashi(-wa)* in the same utterance. Our hypothesis is that if the co-occurrence rate of disfluencies with PFRF intonation is significantly higher than that with normal intonation, and at the same time the co-occurrence with PFLs is also significantly higher than that with normal intonation, we can say that a PFRF is related to disfluency.

3.3. Reasons for collecting samples containing *watashi*

There are three reasons why utterances with *watashi* were chosen. First, *watashi* tends to appear in the topical position as it does in other languages (cf. [8]). If a representation of disfluency appears in the anterior position of an utterance, it may cause another disfluency representation in the rest of the utterance because of ongoing speech planning. We expect that more co-occurrences of disfluent elements like fillers will be observed if a PFRF is a representation of disfluency.

Second, when *watashi* appears in a zero-particle (bare) form or topic-marked form (followed by *wa*) it rarely has salience. Meanwhile, case-marked forms such as *watashi-ga* (subject), *watashi-o* (object) etc. are likely to be focused and pronounced with an emphatic prominence [9]. To avoid interference from focal prominence, we excluded case-marked forms from the target samples.

The third reason for adopting the utterances with *watashi* was because it is lexically unaccented unlike its masculine-only counterpart *boku*, which has an accent on the first syllable. *Watashi* is a general form indicating first person singular but since in casual speech it tends to be used by women, we chose only women for subjects.

3.4. Classification

The samples were classified by phrase-final intonation patterns on *watashi(-wa)* phrases into 3 categories, Normal, PFL and PFRF. Classifications were performed according to auditory perception by the author. PFLs had phrase-final syllable prolongation and no falling intonation, while PFRFs were those with rise-fall intonation on *watashi(-wa)*. The rest of the samples were labelled as Normal.

For disfluency elements, we focused on pauses and fillers. The criteria of pause as disfluency were (i) perceptible silent section inside an utterance and (ii) it sounded longer than the expected length. Judgments of pause were done according to auditory impressions of phrase-final intonation patterns. Fillers were units with no linguistic content (in a traditional sense) like *eeto*, *anoo*, *nanka*. Interjections or discourse makers like *un* “yes”, *e* (surprise), *a* (finding), *hee* “I see” etc. were not treated as fillers.

4. Results

4.1. Basic data

Before testing the relation between phrase-final intonations and disfluency, we analyze non-disfluent types of phrase-final intonation. Distribution of each phrase-final intonation pattern is shown in Table 1 below.

Table 1: Distribution of phrase-final intonation

Type	Normal	PFL	PFRF	Total
<i>watashi</i>	143	22	8	173
<i>watashi-wa</i>	71	23	19	113
Total	214	45	27	286

These data show that PFRFs are fewest among the 3 categories (9.4%). Focusing on the rate of *watashi-wa*, categories PFL and PFRF are significantly higher than Normal (PFL $p < 0.05$, PFRF $p < 0.001$), though PFL and PFRF are not significantly different. Relations between the presence of particle *wa* and phrase-final intonation will be discussed later in this section.

Also we observed co-occurrences of interjectional phrase-final particles like *ne* after *watashi(-wa)* which appear very frequently in colloquial Japanese. Table 2 shows the distribution of particles among intonation patterns.

Table 2: Distribution of phrase-final particle among intonation patterns

PF-particle	Normal	PFL	PFRF	Total
None	207	37	18	262
<i>Ne</i>	6	6	8	20
Others	1	2	1	4
Total	214	45	27	286

Comparing the categories None and *Ne*, PFL and PFRF are significantly different from Normal category (PFL $p < 0.005$, PFRF $p < 0.001$), i.e., the particle *ne* is likely to appear with either PFL or PFRF intonation patterns. As is seen in Table 1, there were 8 cases in which non *wa*-marked *watashi* was pronounced with a PFRF, while 6 of them were accompanied by phrase-final particle *ne*. This means *watashi* is less likely to appear without a phrase-final particle (2 out of 8) when pronounced with a PFRF intonation. Adding *wa*-marked *watashi-wa* with a PFRF (19 samples), only 2 samples of 27 (7.4%) appeared without postpositional morpheme (*wa*, *ne* and *sa* in Other category). For PFLs, 17 out of 45 samples were of *watashi* without any particle. This fact may be a hint for finding differences between PFL and PFRF. Further discussion is continued in Section 5.

4.2. Disfluency-related data

In this section, we report the main analyses, and the possibility of a relation between phrase-final intonation patterns with occurrences of disfluency units such as otiose pauses and fillers defined in Section 3.4. First, we examined the presence of idling pauses or fillers immediately before or after *watashi(-wa)*. Table 3 shows the distribution of co-occurrences of fillers **right before** *watashi(-wa)* and Table 4 for pauses. Samples in which *watashi(-wa)* comes first in the **utterance** were excluded (41 out of 286).

Table 3: Co-occurrences of FILLERS just BEFORE *watashi(-wa)*

Fillers	Normal	PFL	PFRF	Total
-	152	23	13	188
+	30	16	11	57
total	182	39	24	245

Table 4: Co-occurrences of PAUSES just BEFORE *watashi(-wa)*

Pauses	Normal	PFL	PFRF	Total
-	130	31	20	181
+	52	8	4	64
Total	182	39	24	245

With regard to the co-occurrence rate of fillers (Table 3), categories PFL and PFRF are significantly higher than Normal (both $p < 0.001$); with regard to pauses (Table 4), no significant difference was observed. Next, co-occurrences of Fillers **immediately after** *watashi-wa* are shown in Table 5, those of Pauses in Table 6. 17 utterances in which *watashi(-wa)* came after the predicate were out of the subjects.

Table 5: Co-occurrences of FILLERS just AFTER *watashi(-wa)*

Fillers	Normal	PFL	PFRF	Total
-	164	26	16	206
+	35	17	11	63
Total	199	43	27	269

Table 6: Co-occurrences of PAUSES just AFTER *watashi(-wa)*

Pauses	Normal	PFL	PFRF	Total
-	181	28	14	223
+	18	15	13	46
Total	199	43	27	269

Unlike the immediately anterior position (Table 3, 4), for the co-occurrence rate of both Fillers and Pauses, PFL and PFRF were significantly higher than the category Normal (Fillers: PFL $p < 0.005$, PFRF $p < 0.01$; Pauses: both $p < 0.001$). In sum, in the situation of co-occurrence with PFLs and PFRFs, Fillers are more strongly related than Pauses, and disfluency units are more likely to co-occur after *watashi(-wa)* than before.

Finally, we examined co-occurrences of disfluent units (Pauses and Fillers) in an anterior or posterior position to *watashi(-wa)* in the same utterance. Table 7 shows the results for anterior positions, Table 8 for posterior positions. Table 7 excludes samples with the utterance-initial *watashi(-wa)*, and Table 8 those with *watashi(-wa)* after the predicate.

Table 7: Co-occurrences of disfluent units FORMER than *watashi(-wa)*

Fillers	Normal	PFL	PFRF	Total
-	108	18	9	135
+	74	21	15	110
Total	182	39	24	245

Table 8: Co-occurrences of disfluent units LATTER than *watashi(-wa)*

Pauses	Normal	PFL	PFRF	Total
-	128	12	7	147
+	71	31	20	122
Total	199	43	27	269

The co-occurrence rate of disfluent units **anterior** to *watashi(-wa)* (Table 7) was significantly higher in the PFRF category than in Normal ($p < 0.05$), but PFL was not significantly different from Normal, while for **posterior** positions (Table 8) the rate of disfluency units which co-occurred is significantly higher in both PFL and PFRF than Normal (both $p < 0.001$). Similar to the immediately prior and post positions (Tables 3 to 6), disfluent units posterior to *watashi(-wa)* are significantly more likely to co-occur with PFLs and PFRFs than those of anterior.

In conclusion, PFLs and PFRFs were observed to have a co-occurrence relation with posterior disfluent units rather than

anterior. In this sense, we can say PFRF intonation has a relation with disfluency similar to that with PFLs. This implies that PFRFs represent disfluencies in a way similar to PFLs.

5. Discussion

In Section 4, PFRFs as well as PFLs have been shown to relate to disfluency. We will review similarities of PFRF and PFL. As mentioned above, both intonation patterns function to retain the turn in discourse. In addition, when talking to the speaker her/himself, i.e. soliloquies, both PFLs and PFRFs seldom occur. The filler *anoo*, which was mentioned to be similar to PFL intonation in section 2.3, is also unsuitable for soliloquies [10]. Sadanobu & Takubo (1995) explained that *anoo* is “used when the speaker makes a linguistic editing in the mind” and “functions to maintain the speaker-interlocutor interface (1995:79).” PFRFs and PFLs are similar to *anoo*, and in this sense they indicate speech planning in process.

There remains, however, another problem: what is the difference between PFRF and PFL? In Section 4.1, a difference between PFRF and PFL is that the former is likely to co-occur with the particle *wa* or *ne*, not in bare form. When a particle appears after a noun, they can close as one phrase, while bare nouns without postposition are not always independent phrases. From this notion, PFRFs can appear only when the phrase is formed while PFLs do not have such a constraint; they can appear on an unclosed noun phrase. As evidence, our corpus has one sample in which bare *watashi* is pronounced with its final vowel lengthened and after that the particle *wa* appears with a PFRF, like *watashi-i-wâa*. The phrase is not closed yet at the moment of lengthening of the third syllable *shi* and on the fourth syllable *wa* the phrase is closed and can be pronounced with a PFRF intonation. In this sense a PFL is not exactly a phrase final intonation.

Another difference between PFRF and PFL is reported by Ichikawa (2005). Analysis of the relation between phrase-final intonation and the interlocutor’s reactions shows that PFRFs cause the interlocutor to nod or do an *aizuchi* (supportive response) at a higher rate than do PFLs [11]. This means a PFRF can give the interlocutor a chance to react more than a PFL does. We saw that PFRFs and PFLs resemble the filler *anoo*. Sadanobu & Takubo (1995) classified the linguistic function of *anoo* into two operations: “searching for the name” and “considering an appropriate expression”(ibid) [10]. The former is a purely linguistic operation while the latter is a pragmatic one. Applying this idea to phrase-final intonation, I speculate that PFLs are representations of ongoing linguistic editing, such as word selection, remembering things or calculating, while PFRFs, like the above, appear when the speaker monitors her/himself to adjust the utterance according to the interlocutor’s reaction or the discourse situation. These operations require a considerable amount of work for the brain. When the working memory or buffer runs short, disfluencies will occur, fillers and otiose pauses will interrupt linguistic information, and the speech speed therefore will decrease. Also, the speaker’s attention to the interlocutor’s reactions will retard the tempo of speech with presence of PFRF intonation.

6. Conclusion and ending remarks

This paper has shown the relation between a PFRF intonation and disfluency and tried to elucidate the difference between the two phrase-final intonation types, PFRF and PFL. The study has problems mainly in its method; judgments of prosodic features such as pitch form and length of pause were

performed by the auditory impression of a single listener. In future studies, we will utilize more objective methods such as quantitative analysis and multiperson judgments. In addition, utterances of male speakers will also be examined.

7. Acknowledgements

This work is sponsored by (i) Grants-in-Aid for Scientific Research A 16202006 from the Ministry of Education, Culture, Sports, Science and Technology, and by (ii) Strategic Information and Communicative R&D Promotion Programme 041307003 from Ministry of Internal Affairs and Communications.

8. References

- [1] Pierrehumbert, Janet & Mary Beckman. 1988. *Japanese Tone Structure*. Cambridge, Massachusetts: MIT Press.
- [2] Sugito, Miyoko. 1983. Nihongo-no Akusento-to Intoneshon (Accents and Intonations in Japanese). *Kotoba-to Onsei (Language and Speech)*. Agency for Cultural Affairs.
- [3] Inoue, Fumuo. 1997. Intoneshon no Shakaisei (Sociality of Intonation). Sugito, Kunihiro, Hirose & Kono(Eds.) *Nihongo Onsei 2 Akusento, Intoneshon, Rizumu to Pozu (Japanese Speech 2, Accent, Intonation, Rhythm and Pause)*, pp. 143–168. Tokyo: Sanseido.
- [4] Akinaga, Kazue. 1966. Nihongo-no Hatsuon –Intoneshon nado (Pronunciation of Japanese – Intonation and other features). *Koza Nihongo Kyoiku (Seminar in Japanese Language Education)*, vol. 2, pp. 48-60. Waseda University Laboratory of Language Education.
- [5] Hara (Sasaki), Kaori. 1993. The Acoustic Feature and Sensory Impression of What Is Called “High-Rising” Intonation in Japanese. *Studies in Language and Culture*, vol. 11, pp 61-71. Graduate School of Area and Culture Studies, Tokyo University of Foreign Studies
- [6] Nakagawa, Akiko & Toshiyuki Sadanobu. 2003. The Contrastive Study of “Disfluency” between Japanese and Chinese: Preliminary Research. Proceedings of *The 1st JST/CREST International Workshop on Expressive Speech Processing*. 21-22, February 2003. In Kobe University, Japan.
- [7] National Institute of Japanese Language, the. 1999-2003. *The Corpus of Spontaneous Japanese, Preliminary Analyses II*. http://www2.kokken.go.jp/~csj/public/j6_2.html
- [8] Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Cole (ed.) *Radical Pragmatics*, pp. 223-255. New York: Academic Press.
- [9] Kaneda, Jumpei. 2005. Unmarkedness of Zero-Particle NPs. *KLS 25*, pp. 315–325. Kansai Linguistic Society.
- [10] Sadanobu, Toshiyuki & Yukinori Takubo. 1995. The Monitoring Devices of Mental Operations in Discourse: A Case of “Eeto” and “Ano(o).” *Gengo Kenkyu (Studies in Linguistics)* vol. 108, pp. 74-93. The Linguistic Society of Japan.
- [11] Ichikawa, Akira. 2005. Multimodal Dialogue Corpus and Analysis of Speaker’s Nod. Report of *Realization of Advanced Spoken Language Information Processing from Prosodic Features*. http://www.gavo.t.u-tokyo.ac.jp/tokutei_pub/houkoku/corpus/ichikawa.pdf.