FUNCTIONAL DIFFERENCE BETWEEN THE TWO VARIANTS OF RISING-FALLING INTONATION IN SPONTANEOUS JAPANESE MONOLOGUE

Kikuo Maekawa
Dept. Corpus Studies, National Institute for Japanese Language and Linguistics
kikuo@ninjal.ac.jp

ABSTRACT

Penultimate non-lexical prominence, or PNLP, is a variant of rising-falling boundary pitch movement (L%HL%) in Tokyo Japanese. Analysis of Corpus of Spontaneous Japanese revealed hitherto unknown function of the PNLP. It occurred basically only once near the end of an utterance flanked by deep clause boundaries. This finding suggests strongly the interpretation that native speakers of Japanese use PNLP to predict the end of a long utterance. Pilot analysis was also conducted to examine the factors governing the occurrence of PNLP.

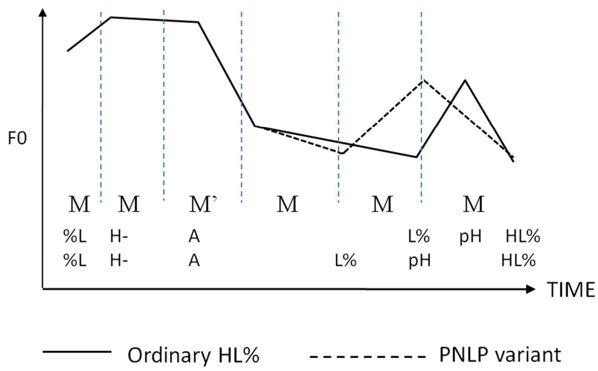
Keywords: PNLP, BPM, CSJ, delimitative function

1. INTRODUCTION

As is well known, Tokyo (Standard) Japanese is coloured by the use of various boundary pitch movements, or BPM, including rising (L%H%), rising-falling (L% HL%), and rising-falling-rising (L%HLH%) among others [1]. What we call here a penultimate non-lexical prominence, or PNLP, is a temporal variant of the L%HL% BPM. In ordinary L%HL%, the rise and fall of F0 both occur in the last mora of an accentual phrase (AP, hereafter), while in the PNLP variant, F0 starts to rise in the beginning of the penultimate mora, reaches its peak at the boundary between the penultimate and last morae, then starts to fall toward the end of the AP [2]. Figure 1 shows a schematic comparison between the ordinary and PNLP variants of L%HL%. The utterance consists of six morae that are represented by the symbol "M", the third one being accented (hence an apostrophe). The mora boundaries are shown by dotted lines. X-JToBI tone annotations [3] below the morae show schematically the time alignments of tones in the two variants. The symbol "pH" stands for the peak of the BPM.

PNLP was first reported by Hatsutaro Oishi in his 1959 paper [4], in which he classified various means of realizing phonetic emphasis into *maedakagata* ("early-peak type") and *atodakagata* ("late-peak type"); the former type realizes emphasis by expanding the pitch range of emphasized element, while the latter realizes prominence by means of BPM. Needless to say, PNLP belonged to the latter type. Oishi's analysis was mostly concerned with the functional difference between the two types of emphasis mentioned above, and no attention was paid to the difference between the ordinary and PNLP variants of the rising-falling BPM. Subsequent studies on the PNLP were no more successful than Oishi's in this respect [5-8]. The aim of the present study is to find out the difference, if there is any, based on the analysis of a spontaneous speech corpus.

Figure 1. Schematic comparison of ordinary and PNLP variants of the L%HL% BPM.



2. DATA

2.1. The CSJ-Core

CSJ-Core, i.e. the X-JToBI annotated part of the Corpus of Spontaneous Japanese (CSJ) was used for analysis [9]. Only the monologue samples consisting of 56 academic presentation speech (or APS) that included 18 female and 38 male speakers, and 53 simulated public speaking (or SPS) that included 53 female and 53 male speakers were analysed. Table 1 shows the number of relevant BPMs observed in the current data. The last row of the table shows that more than 26% of total AP was marked by BPM (L%H% or L%HL%), and 10% of the L%HL% were PNLP variant. The table also shows that PNLP appeared more frequently in APS rather than SPS.

Table 1. Number of BPM in the data.

Register	L%H%	L%HL%		Total AP
		Ordinary	PNLP	
APS	19,691	2,953	616	74,973
SPS	9,939	6,677	375	75,609
Total	29,630	9,630	991	150,582

2.2. Clause boundaries

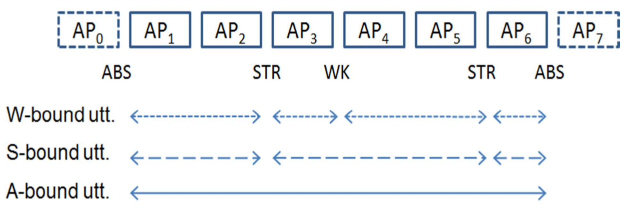
It is widely acknowledged that segmentation of spontaneous speech into "sentence" is very difficult. Instead, segmentation into clauses is known to be much easier [10]. For this reason, the CSJ (including the CSJ-Core) was annotated in terms of clause boundary labels (or CBL). All syntactic clauses in the CSJ were analysed into 49 classes, and reclassified into three major CBL classes. The absolute (or ABS) boundary applies for the cases where a clause ends with typical sentence ending forms like /desu/, /desita/, /masu/, /masita/ etc. The strong (or STR) boundary applies for the clauses ending with /ga/, /keredo/, /kedo/, /si/ etc. The weak (or WK) boundary applies for the clauses ending with /tara/, /to/, /nara/, /node/, /te/, /toiu/, etc.

ABS boundary is the strongest boundary; nearly all syntactic dependency relationships are reset at this boundary. STR boundary is the second strongest, and the WK is the weakest [11]. In the current data, 10224, 7202, and 24405 AP were classified as ABS, STR, and WK boundaries respectively. Remaining 108769 AP are not concerned with any clause boundaries (See Table 1).

2.3. Utterance segmentation by means of CBL

Speech data can be segmented into chunks by means of CBL. In this paper, three types of "utterances" are defined by means of CBL. Chunks of APs whose edges are flanked both by ABS boundaries and having no internal ABS boundary will be called A-bounded utterance. Similarly, S-bounded utterances are the chunks flanked either by ABS or STR boundaries and having no internal ABS/STR boundaries. Lastly, W-bounded utterances are the chunks flanked by whatever CBL and having no internal CBL. Figure 2 compares various utterances derived from the same data. The data consists of eight APs, and CBLs are shown below, whenever available, near the end of each AP. Arrows below the CBL show resulting segmentations of the target AP sequence (i.e., AP_1 to AP_6). Dotted, broken, and real arrows stand respectively for W-bounded, S-bounded, and A-bounded utterances. As seen from the figure, there is a hierarchical relationship among these utterances.

Figure 2. Utterance segmentation by CBL.



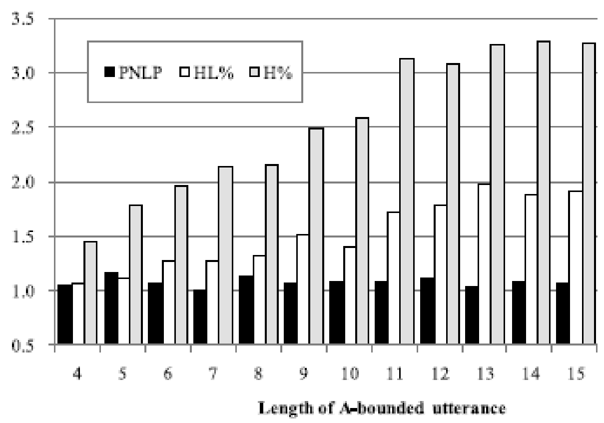
3. ANALYSIS

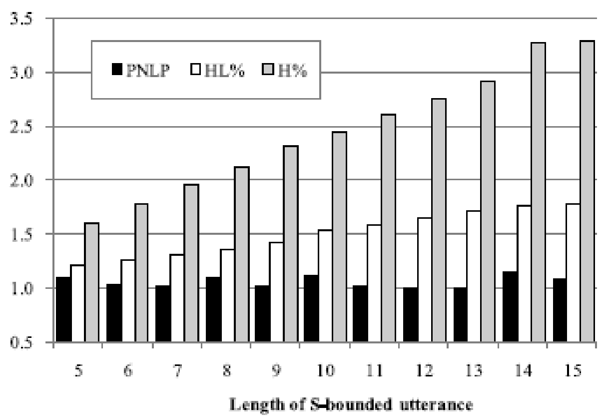
3.1. Comparison of occurrence frequency

In this section, occurrence frequencies of three BPMs are to be compared across utterance types. Figure 3 shows mean frequencies of the L%H% (abb. H% hereafter), ordinary L%HL% (abb. HL%), and PNLP version of L%HL% as a function of the length of utterance. The utterance length is measured as the number of AP in an utterance. Note that, in the computation for this figure, only the utterances containing at least one designated BPM were counted. Accordingly, the least possible mean frequency is 1.0.

In both A-bounded and S-bounded utterances, there were striking difference between the PNLP and other two BPMs. In the case of H% and ordinary HL%, mean occurrence frequency increased monotonically as a function of utterance length, while in PNLP, mean frequency stayed nearly unchanged at slightly above 1.0, suggesting that PNLP occurred, in the vast majority of cases, only once in an utterance.

Figure 3. Mean occurrence frequency of BPMs as a function of utterance length in A-bounded utterance (top) and S-bounded utterance (bottom)





3.2. Location of PNLP

The result reported in the previous section suggested strongly the interpretation that PNLP had a sort of culminative (a.k.a. peak-marking) function [12] at the level of utterance (i.e., the domain larger than ordinary sentences; see the discussion below), while ordinary HL% and H% do not have such a function. To see the validity of this interpretation, occurrence locations of PNLP and other BPMs were examined across utterance types.

Figure 4 shows the relative (%) occurrence frequency of PNLP in A-bounded utterances as a function of the location of the AP in which a PNLP occurred. Note only the utterances having PNLP were analysed for this figure. The top panel stands for utterances consisting of 5-10 APs, and the bottom panel stands for utterances of 11-15 AP long. If we take example of the utterance consisting of six APs (indicated by open rectangle in the top panel), the frequency was low in the beginning, increased gradually as utterances got longer, reached the maximum value in the penultimate (fifth in this case) AP, and decayed abruptly to zero at the end. The same tendency was observed in all utterances regardless of utterance length. This suggests strongly that the optimal location of PNLP is the penultimate AP of an A-bounded utterance.

Figure 4. Relative frequency (%) of PNLP as a function of AP location in A-bounded utterances. Cases of 5-10 APs (top) and 11-15 APs (bottom).

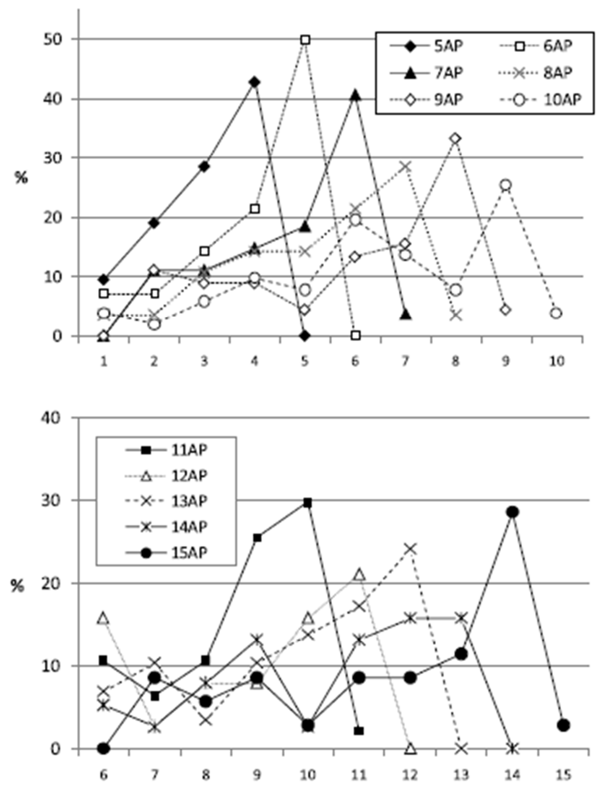


Figure 5. Relative frequency (%) of PNLP as a function of AP location in S-bounded (top) and W-bounded (bottom) utterances.

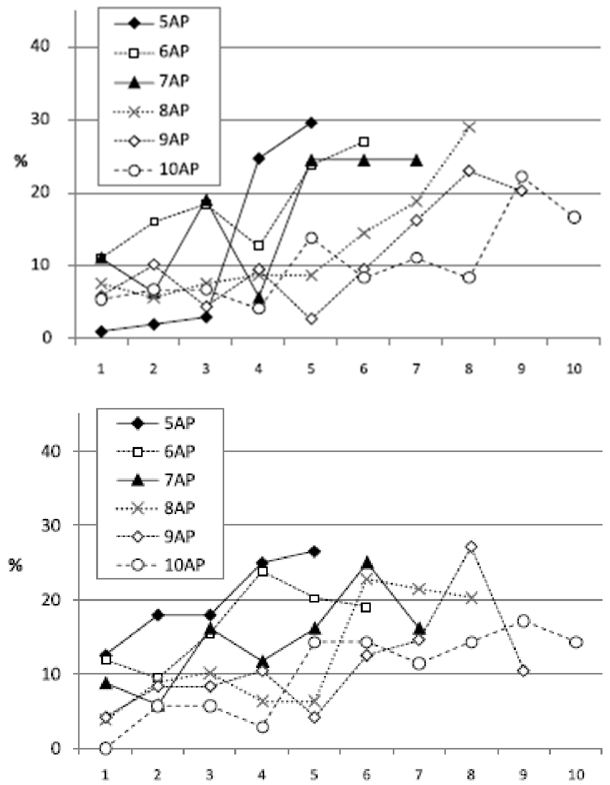
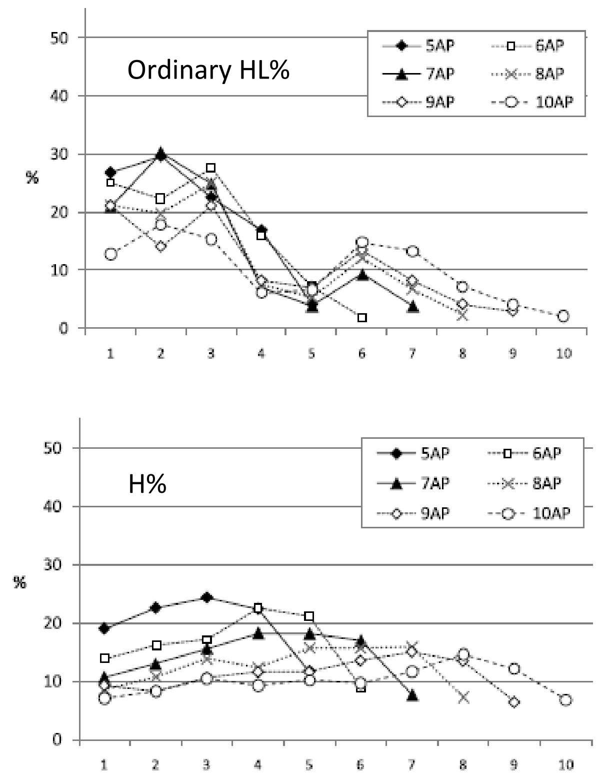


Figure 6. Relative frequency (%) of ordinary HL% (top) and H% (bottom) as a function of AP location in A-bounded utterances.



At this point, it is very important to note that this pattern was observed exclusively in the A-bounded utterances. As shown in Figure 5, S-bounded and W-bounded utterances did not show anything comparable to the pattern observed in Figure 4.

Lastly, Figure 6 examined the behaviour of other BPMs. The top and bottom panels stand respectively for ordinary HL% and H%. As can be seen from the figure, their frequency distributions were dissimilar to the one observed in Figure 4; the distribution of ordinary HL% had two peaks, and, that of H% was almost uniform.

4. DISCUSSIONS

It turned out by the analyses reported in section 3 that PNLP occurs, most typically, shortly before the end, and almost never at the end, of an A-bounded utterance. This property, which could not be found neither in ordinary HL% nor H%, suggested strongly that PNLP had the function of predicting the end of an A-bounded utterance, i.e. a sort of "delimitative" function [12] at the utterance level. On the other hand, however, it is important to note that the use of PNLP is not obligatory in Japanese intonation, unlike the nuclear tones in English [13]. PNLP marked only a small subset of A-bounded utterances. In this respect, it is of much interest to pursue the question what kind of A-bounded utterances are likely to be marked by PNLP. It is not an easy question, but there are at least two issues to be examined here.

First, A-bounded utterances in spontaneous monologue tend to be very long compared to ordinary written sentences. In the current data, actually, there were 204, 64, 40, and 14 A-bounded utterances that were 20, 30, 40, and 50 AP long respectively. Therefore, it is natural to expect higher occurrence of PNLP in longer utterances; prediction of the end of an utterance will help listeners greatly in the task of sentence processing.

Figure 7 is a scatter plot of utterance length (measured by the number of AP) and relative frequency of PNLP. The range of utterance length was limited so that there were at least 100 A-bounded utterances for a given utterance length. As predicted, there is clear positive correlation (r=0.852) between the two variables.

Second, while analysing the data, it seemed to be the case that there was higher chance of finding topic boundary after an A-bounded utterance marked by a PNLP than elsewhere. A pilot text analysis was conducted to check the validity of this impression. Figure 8 compares the occurrence rates of conjunctions (like /de/, /sore de/, /sore kara/, /mata/ etc.) after the A-bounded utterance boundaries with and without PNLP. If there was one or more conjunctions in the first three words (compounds were counted as one) of the utterance that immediately followed the boundary in question, the boundary was counted as a boundary having conjunction. As predicted, conjunction had higher occurrence rate after the PNLP boundary than elsewhere. But at the same time, the absolute difference remained less than 10%. This is probably too weak an evidence to show that PNLP was used mostly to predict a topic boundary. There should be other factors than topic boundary prediction behind the occurrence of PNLP.

Figure 7. Scatter plot of utterance length and relative frequency of PNLP.

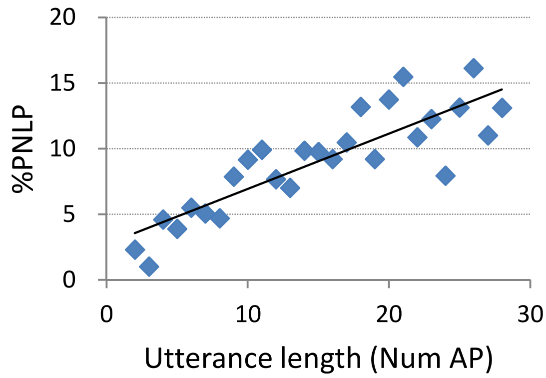
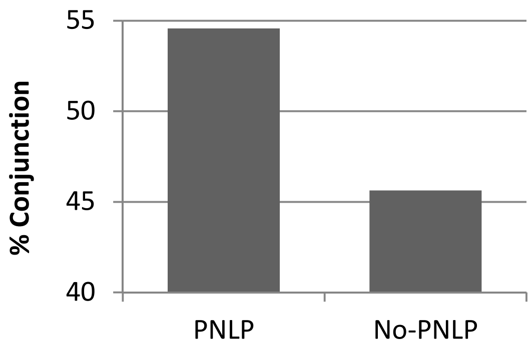


Figure 8. Occurrence rates of conjunction after the end of an A-bounded utterance with and without PNLP.



5. CONCLUDING REMARKS

The present analysis revealed crucial functional difference between the ordinary and PNLP variants of rising-falling (L%HL%) BPM in Tokyo Japanese. The difference consisted in that the PNLP had culminative and delimitative functions at the level of A-bounded utterance. As for the factors governing the occurrence of PNLP, however, the current analysis remained still in a preliminary stage. Further investigation on the discourse factors favouring / disfavouring PNLP is needed.

6. REFERENCES

[1] Venditti, J., K. Maekawa, and M. Beckman. 2008. "Prominence marking in the Japanese intonation system". *The Oxford Handbook of Japanese Linguistics*, Oxford Univ. Press, pp.456-512.

[2] Maekawa, K. "Phonetic shape and linguistic function of penultimate non-lexical prominence". *Journal of the Phonetic Society of Japan*, 15 (1), pp.1-13, 2011 (In Japanese).

[3] Maekawa, K., H. Kikuchi, Y. Igarashi and J. Venditti. "X-JToBI: An extended J_ToBI for spontaneous speech". *Proc. ICSLP2002*, pp.1545-1548, Denver, 2002.

[4] Oishi, H. "Purominensu ni tsuite". *Kotoba no Kenkyū*, 1, pp.87-102, Kokutistu kokugo kenkyūjo, 1959.

[5] Maruyama, T. and M. Taniguchi. "Bun no shōten kōzō to kyokushoteki takuritsu". *KLS*, 22, pp. 18-28, Kansai gengo gakkai, 2002.

[6] Ezaki, T. "PNLP no seikiyouin". *Proc. 20th Annual Convention of the Phonetic Society of Japan*, pp. 171-176, 2006.

[7] Tagashira-Taniguchi, M. "PNLP no onsei tokuchō". *Onsei Gengo*, 4, pp.67-80, 2008.

[8] Taniguchi, M. PNLP to yobareru onchō henka no jittai chōsa". *Gengo Bunkagaku*, 17, pp.253-262.

[9] Maekawa. K. "Corpus of Spontaneous Japanese: Its Design and Evaluation". *Proc. ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition (SSPR2003)*, Tokyo, pp.7-12, 2003.

[10] Takanashi, K., et al. "Identification of 'sentence' in spontaneous Japanese". *Proc. ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition*, 183-186, Tokyo, 2003.

[11] Maruyama, T., K. Takanashi, and K. Uchimoto. "Setsutan'i jōhō". In Nt'l Language Research Inst. (ed). Nihongo Hanashikotoba Kōpasu no Kōchikuhō, Nt'l Language Research Inst., pp. 255-322, 2006. (http://www.ninjal.ac.jp/products-k/katsudo/seika/ corpus/csj_report/).

[12] Trubetzkoy, N. *Principles of Phonology*. Translated by C. Baltaxe. Univ. California Press, 1963.

[13] Cruttenden, A. *Intonation*. Cambridge Univ. Press. 1986.