Phonological and Phonetic Effects of Minor Phrase Length on f0 in Japanese

Elisabeth Selkirk, Tahahito Shinya and Shigeto Kawahara

Department of Linguistics
University of Massachusetts, Amherst
{selkirk; tshinya; kawahara}@linguist.umass.edu

Abstract

The Minor Phrase (MiP, aka accentual phrase) is the prosodic constituent that immediately dominates the prosodic word (PWd) in the prosodic structure hierarchy; it may consist of one or more PWd. In Japanese all MiPs are marked by an initial LH rise. This paper examines the scaling of the initial rise in single-word MiPs in Japanese as a function of the syllable/mora length of the word constituting the MiP, the position of the MiP with respect to edges of prosodic major phrase (MaP), and the composition of MiP in terms of lexical accent. These rises are found to be subject to two types of scaling: (i) *local, edge-based* scaling, specifically the upward "resetting" of f0 seen at the left edge of MaP (aka intermediate phrase) [1, 2 3], and (ii) *global, lookahead-based* scaling, in this case the upward scaling of the f0 of MiP-initial peaks as a function of the overall length of the MiP in terms of syllables/moras. Word length also turns out to have an indirect influence on local, edge-based scaling in Japanese, since it can be shown that word length has an effect on the number and distribution of major and minor prosodic phrases in the phonological representation as well.

1. Background and Findings

For Japanese, it has been shown that the left edge of a prosodic MaP, which typically appears at the left edge of a major syntactic phrase, is the context for a significant upward "resetting" of pitch [5, 6, 1, 2, 4]. [2] and [4] explicitly compare the realization of LH rises at the left edges of MiPs which coincide with the left edge of MaP to MiP-initial rises that are medial in MaP and find the former to be significantly higher. (These results are replicated in the present experiment.) Such an effect of prosodic organization on f0 realization is *local* and *edge-based* insofar as the phonetic realization mechanism can produce these effects in conditions defined locally, in proceeding from left to right in the utterance.

Some have claimed that the overall word or syllable count of an utterance must be a factor in f0 realization [7]. The existence, or even possibility, of this sort of *global, lookahead* effect has been contested [8], and no lookahead effect has to our knowledge been reported for Japanese. In our recent research, however, we have found a syllable or mora-counting lookahead effect at the level of the MiP in Japanese—the longer the anticipated MiP, the greater the initial rise, and the higher the initial peak.

The paper also reports an effect of the syllable/mora count of MiPs that is phonological in kind: words consisting of 3 moras contrast with words consisting of 5 and 7 moras in their ability to stand alone as MiPs in the larger prosodic structure. So word length here too does have an (indirect) effect on the phonetic realization of f0, insofar as it influences the prosodic structure of the output phonological representation, which is input to the phonetics.

2. The experiment

In our experiment 3 female native speakers of Tokyo Japanese from the UMass community in their mid-twenties (KK, SS, RH) read sets of Japanese sentences which provided the relevant linguistic contrasts.

2.1. The experimental materials

Experimental materials involved 5-word sentences in which the third noun varied in length from 3 to 7 syllables/moras; the syllables were all monomoraic CV. All other nouns in the sentence were consistently 5 moras long. N3 appeared in two distinct syntactic contexts: either as the third and final noun of a subject noun phrase (a) or as the sole noun of a post-subject dative noun phrase (b). In both types of sentence, N1 and N2 form (part of) the subject noun phrase; N4 is an accusative noun phrase and is followed by the Verb. Since earlier studies have shown that the left edge of a syntactic maximal projection coincides in prosodic structure with a MaP edge [2, 4], the different sentence types were expected to show the prosodic phrasing indicated with curly brackets.

(1) Subject-final N3

```
[[[[ N1-no]  N2-no ]  N3-ga ]  [[ N4-o ] Verb] ]
 {                          } {                 } = MaP
```

(2) Dative N3

```
[[[ N1-no ] N2-ga ] [ [ N3-ni  ] [[ N4-o ]  Verb]]]
 {                 } {          }{                } = MaP
```

We also varied the lexical accent status of N3 and N2. Cases in which N2 and N3 were accented were contrasted with cases in which they were unaccented. All the remaining words were unaccented. Accent on N2 should lead to a catathesis of N3 if it is contained within the same MaP, but absence of catathesis if N3 initiates a MaP on its own. [6, 2].

These syntax/length/accent combinations yield twelve distinct sentence types. For each type we used two different token sentences consisting of different lexical items. Each token sentence was given 6 repetitions, for a total of 12 pronunciations of each sentence type.

2.2. Procedure and Analysis

Each of the 3 speakers had two recording sessions on different days in a sound-attenuated booth. In each session there were three separately randomized repetitions of the entire set of sentences, which consisted of experimental sentences and filler sentences. Each sentence was written on an index card in the Japanese *kana-kanji* orthography. Speakers were asked to read the sentences one at a time as naturally as possible. The sentences were recorded onto a CD and digitized at a 11.025 KHz sampling rate and 16 bit quantization level. *PitchWorks* (Scicon R & D) was used for f0 measurement.

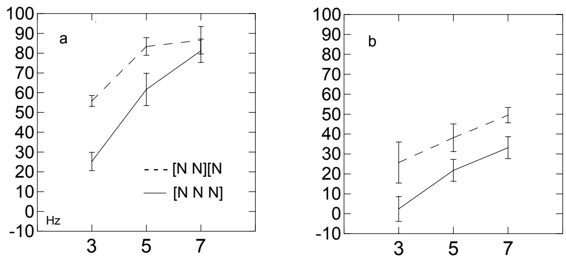
In an accented word the initial low point and the peak of the accent-related fall were measured. In an unaccented word, which may or may not show an initial rise, the f0 at the start of the vocalic nucleus of the initial syllable was measured and the f0 at the peak of the initial rise (if there was one) or the f0 at the beginning of the second syllable (if there was not).

A two-way ANOVA was performed on the accented data for each speaker with type of sentence and word length as the independent variables. *T*-tests were used for the unaccented data.

3. The Syntax-Based Phrasing Effect

3.1 Accented cases

For all speakers, there is a sizable rise at the left edge of the accented dative N3, for all lengths of N3. This is arguably due to the expected upward reset of pitch at the left edge of the MaP which, according to the syntax-phonology interface alignment constraint **AlignL (XP, MaP),** aligns the left edge of the dative XP with the left edge of MaP [2], cf. (2). (XP refers to syntactic maximal projection.) This large rise in the dative case contrasts with a significantly lower rise for the length-matched subject-final case when N3 is 3 or 5 moras long, for all speakers: (p<.0001 (3mora or 3μ) & p<.0003 (5μ) for K; p<.0001(3μ) & p<.0001 (5μ) for R; p<.0001 (3μ) & p<.0001 (5μ) for S). When N3 is 7 moras long, though, it shows a significantly lower value in the subject-final case only for speaker R (p<.0001). In the Figure 1 below, the top line shows the initial rise values (H3-L3) for N3 in the dative, the lower line shows the values for N3 in the subject.



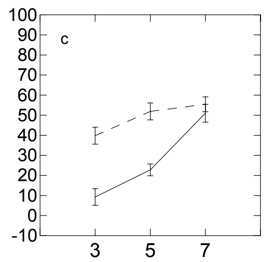
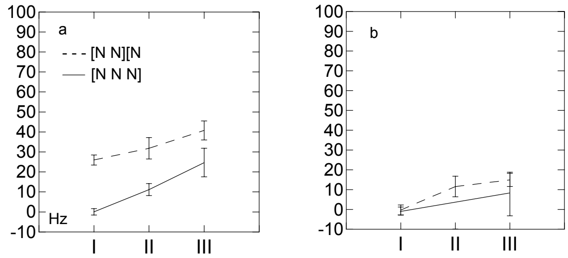


Figure 1
Mean values of the H3-L3 rise at N3 for the accented cases.
(a) Speaker K, (b) Speaker R, (c) Speaker S.
Numbers 3, 5 and 7 stand for mora count of the MiP.

For speakers K and S, the 7 mora subject-final N3 is seemingly as high as in the dative case (p=.013 for K; p=.105 for S . In further experiments with the same speakers [9] we have determined that this apparent neutralization in the 7μ cases is due to an exceptionally large phonetic effect of the 7μ length in the subject-final case and not to a length-driven phonological reorganization.

3.2 Unaccented cases

Figure 2 shows the graphs for the unaccented cases. For speakers K and S, when N3 consists of 3, 5 or 7 moras, the initial rise of N3 in the dative context is always significantly greater than in the subject–final context (p<.0001 (3μ) & (p<.0001(5μ) & p<.0007 (7μ) for K; p<.0001(3μ) & p<.0001(5μ) & p<.0001 (7μ) for S.). This, again, is a straightforward effect of the difference in MaP organization in the two syntactic contexts: an initial upward "reset" scaling found at the left edge of MaP in the dative.



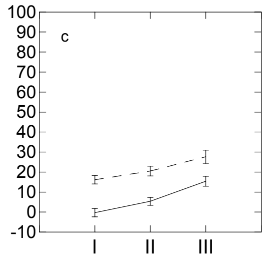


Figure 2
Mean values of rise at N3 for the unaccented cases: (a) Speaker K, (b) Speaker R, (c) Speaker S.
I, II and III stand for observed patterns (I: no initial rise (IR), II: slight IR and III (bigger IR)

For speaker R, however, there was no significant difference in the values of the initial rise for the two contrasting 3-mora cases (p=.753(3μ)). Indeed, in both the dative and the subject-final case the rise with 3-mora N3 was virtually nil, which means that in both cases there not only is there no MiP edge, but no MaP edge either. Rather, R's 3 mora dative forms part of the same MiP as the final noun in the preceding subject. As for R's 7μ cases, both show initial rises, but though the mean of the 7μ dative is higher, the difference between the dative and the subject-final case is not statistically significant (p<.055 (7μ) for R), probably due to the extreme variability seen in the subject-final case. As for the R's 5 mora case, for reasons explained below, data that can be compared is not available.

3.2 Summary

The above data provides clear evidence for upward pitch resetting at the left edge of the prosodic constituent MaP in Japanese, which in these cases marks the left edge of a syntactic maximal projection (XP). Evidence that the context for resetting is not defined directly in terms of syntactic structure comes from R's 3 mora dative cases, which *lack* the upward resetting that would be expected if the context were defined directly in terms of the syntax, as the left edge of a syntactic maximal projection.

4 Phonological Length Effects

4.1 Unaccented cases

4.1.1 The subject-final case: a constraint against short MiPs

N3 is always in the same major phrase (MaP) as the preceding N1 and N2 when N3 is in subject-final position, as in the structure (1) above. Our results show that whether or not unaccented N3 groups with the preceding N2 in the same minor phrase (MiP) depends on the length of N3. For all speakers, a subject-final unaccented 3 mora N3 fails to show any initial rise—it always joined with the preceding unaccented N2 in the same MiP. 7 mora units, by contrast, always showed an initial rise--they were never in the same MiP as what precedes. 5 mora cases showed both presence and absence of initial rise, so both types of organization.

(3) Length-based MiP phrasing of unaccented N3 in Subject MaP:

a. N3=3μ or 5μ: { (N1-no N2-no N3-ga $)_{MiP}$ $\}_{MaP}$
b. N3=7μ or 5μ: { (N1-no N2-no $)_{MiP}$ (N3-ga $)_{MiP}$ $\}_{MaP}$.

In these MaP-internal subject-final cases, there is a clear preference for a 3 mora noun to *not* stand on its own as a MiP. We can express this constraint on the size of MiP as **$*(3\mu)_{MiP}$: "3-mora MiP are not allowed"**.

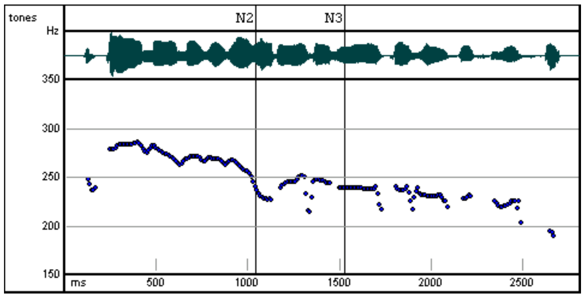
4.1.2 The dative case: MiP length constraint may override XP-MaP alignment constraint

When an unaccented N3 is a dative noun phrase with 5 or 7 moras, it will always stand on its own as a MaP (and hence a MiP) as shown by the presence of initial lowering with significant upward pitch reset. This phonological representation, shown in (4a), is expected, since it respects both the MiP length constraint and the syntax-phonology alignment constraint Align-L(XP, MaP). When unaccented dative N3 is 3 moras long, however, these two constraints are in conflict, and our speakers show different resolutions of the conflict. Speakers K and S parse the 3-mora dative as a MaP, as in (4a). This structure violates the length constraint on MiP (since a MaP always contains a MiP), but it satisfies the XP-MaP alignment constraint. For speaker R, however, the 3-mora unaccented dative N3 forms part of the same MiP and MaP with the preceding subject-final N2, giving (4b), which is a structure identical to (3a). In this case, the alignment constraint is violated, but the MiP length constraint is satisfied.

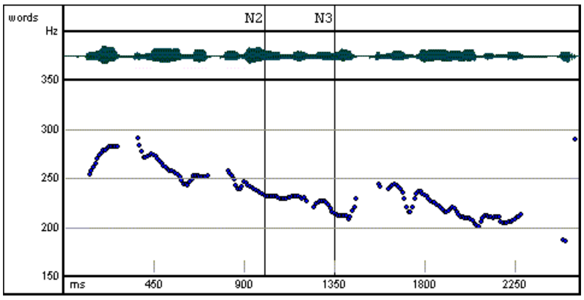
(4) Interaction of Phonological Length and Interface Alignment in the Unaccented Dative Case:

a. 3μ, 5μ, 7μ (K,S): {(N1-no N2-no$)_{MiP}$ $\}_{MaP}$ {(N3-ga$)_{MiP}$ $\}_{MaP}$.
5μ, 7μ (R)
b. 3μ (R): {(N1-no N2-no N3-ga $)_{MiP}$ $\}_{MaP}$

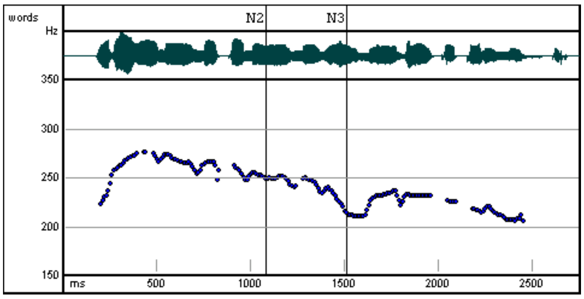
We can explain this difference by assuming that the constraint $*(3\mu)_{MiP}$ overrides the XP-MaP alignment constraint in R's internalized grammar, an assumption we can express with the optimality theoretic ranking **$*(3\mu)_{MiP}$ >>Align L (XP, MaP).** As for speakers K and S , the alignment constraint would instead dominate the length constraint--**Align L (XP, MaP) >> $*(3\mu)_{MiP}$** --forcing a 3 mora noun to stand as a MaP (and MiP) on its own. Figure 3 shows pitch tracks of (i) R's 5μ dative (4a), contrasting with (ii) the 3μ dative (4b), and (iii) the virtually indistinguishable 3μ subject-final (3a). The left edge of N3 is found at the cursor preceded by the label N2.



[N1 N2] [N3] [N4 V] with 5 mora N3 -- speaker RH



[N1 N2] [N3][N4 V] with 3 mora N3 – speaker RH



[N1 N2 N3][N4 V] with 3 mora N3 -- speaker RH

Figure 3*: Representative pitch tracks from speaker RH showing the presence/absence of initial rise at left edge of N3 as a function of mora length and syntactic context*

4.2 Accented cases

4.1.1 *The dative case*

Unsurprisingly, in the accented dative case, speakers K and S treat the 3 mora N3, as well as the 5 and 7 mora dative N3 as a MaP on its own. This is shown by the significantly greater initial rise found in dative as compared to the length-matched cases of N3 in subject-final position (p<.0001 for K; p<.0002 for R; p<.0001 for S). Interestingly, in the accented case, speaker R fails to show the neutralization of the 3 mora dative with the 3 mora subject-final N3 that she showed in the unaccented case. Instead the 3 mora dative N3, as well as the 5 and 7 mora N3, stand on their own as a MaP, as evidenced by the substantial initial rise exhibited. This indicates that a yet further prosodic constraint must outrank the minimal size constraint **$*(3\mu)_{MiP}$** in R's grammar. The relevant constraint is presumably the one requiring that every accent coincide with the head of a MiP, **Accent to ΔMiP,** and which is independently responsible for ruling out MiPs containing more than one lexically accented word [3]: **Accent to ΔMiP >> $*(3\mu)_{MiP}$**. This same ranking would be found in the grammars of the other two speakers as well. It produces the representation in (5a) in the dative and (5b) in the subject case:

(5) Dative and Subject-final phrasings in the Accented case respect both XP-MaP alignment and Accent to ΔMiP:

a. All Dative: $\{(N1\text{-}no)_{MiP}(N2\text{-}no)_{MiP}\}_{MaP}\{(N3\text{-}ga)_{MiP}\}_{MaP}$.
b. All Subject-final: $\{(N1\text{-}no)_{MiP}(N2\text{-}no)_{MiP}(N3\text{-}ga)_{MiP}\}_{MaP}$

4.2 The subject-final case

The constraint Accent-to-ΔMiP predicts that two accented words in sequence would be parsed into separate MiP, and this prediction is borne out in the subject final cases of N3. In contrast to the unaccented case, here all speakers show a tangible initial rise—hence MiP break—at the left edge of subject-final accented N3, even when N3 = 3μ. In other words, for all speakers, the constraint Accent-to-ΔMiP is guaranteeing that the subject final accented N3 be a MiP, whatever its mora count, suggesting that the ranking Accent to ΔMiP >> $*(3\mu)_{MiP}$ is indeed true for all speakers.

5. Phonetic Length Effects

5.1 Unaccented cases:

As seen above, most cases of *subject-final* 5-mora N3 and all cases of 7-mora N3 do show an initial rise and are hence MiP. A comparison of the 5-mora and 7-mora rises shows that the rise in the 7-mora case is significantly higher than that in the 5-mora case for speakers K and S (K: $p < .002$*; S: $p < .0001$**). Because these values are still lower than the values seen in MaP initial position in the dative, we are led to conclude that this effect of length is due to pitch scaling which *looks ahead and counts* the number of syllables/morae in the MiP. (For speaker R there is no 5μ unaccented data available, thus no comparison with 7μ, since she treated one token set of 5 mora N3 here as accented nouns and the other did not have MiP status, showing no initial rise.)

We saw in (4) that for K and S, all three lengths of dative N3 stand on their own as MaP. For both speakers, the difference in the initial rise of the shortest, 3μ, and the longest, 7μ, is highly significant (K: p<.0001; S: p<.0001), though the difference between 7μ and 5μ was significant only for S (K: p=.012; S: p<.0002) and the difference between 3μ and 5μ is not significant for either (K: p=46; S: p=.007).

5.2 Accented cases:

In the accented dative cases, cf. (5a), all speakers show a significant difference between the rises in the 3μ and 7μ cases, showing there is a definite overall length anticipation trend.

	5 vs. 3	**7 vs. 5**	**7 vs. 3**	**μ lengths**
Kinuyo	p**<.0001**	p=.120	p<**.0001**	
Satoko	p=**.0002**	p=.531	p<**.0001**	
Ruiko	p=.033	p=.044	p<**.0001**	

The ANOVA results for K and S also show a significant difference in initial rise between 3μ and 5μ but not between 5μ and 7μ, suggesting a possible ceiling effect in the lookahead effect. In sum, then, there is a visible lookahead effect of MiP length seen in the accented dative cases as well.

Turning to the subject-final accented cases, with the structure in (5b), with one exception they show highly significant length-based differences on all comparisons.

	5 vs. 3	**7 vs. 5**	**7 vs. 3**	**μ lengths**
Kinuyo	p<**.0001**	p<**.0001**	p<**.0001**	
Satoko	p<**.0001**	p<**.0001**	p<**.0001**	
Ruiko	p<**.0001**	p=.015	p<**.0001**	

Clearly the syllable/mora length of the word constituting a MiP has a significant effect on its f0 scaling.

6. Conclusion

This paper has shown that there is a phonetic word-length anticipation effect on the realization of the f0 rise found at the left edge of minor phrase (MiP) in Japanese. This phonetic effect could be established by carefully controlling for the phonological prosodic structure representation which is input to phonetic interpretation. That prosodic structure is influenced by word length as well-- and also by the surface syntactic structure-- through a ranking of constraints on the surface prosodic structure.

* Acknowledgements: The work reported on this paper was supported by NSF grant BCS-0004038 to Elisabeth Selkirk.

7. References

[1] Pierrehumbert, J.B.; Beckman, M. E., 1988. *Japanese Tone Structure*. Cambridge, Mass: MIT Press.
[2] Selkirk, E.; Tateishi, K., 1991. Syntax and downstep in Japanese, In *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda,* C. Georgopoulos; R. Ishihara (eds.) Dordrecht: Kluwer, 519-43.
[3] Sugahara, M., 2003. *Downtrends and Post-FOCUS Intonation in Tokyo Japanese.* Doctoral Dissertation, University of Massachusetts, Amherst.
[4] Selkirk, E.; Sugahara, M.; Shinya, T., 2003 Degree of Initial Lowering in Japanese as a Reflex of Prosodic Structure Organization. *Proc. 15th ICPhS Barcelona.*, 491-494.
[5] Fujisaki, H.; Hirose, K., 1984 Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *J. Acoust. Soc. Jpn. (E),* 5(4), 233-242.
[6] Poser, W. J., 1984. The phonetics and phonology of tone and intonation in Japanese. Doctoral Dissertation, MIT.
[7] Cooper, W.; Sorenson, J., 1981 *Fundamental Frequency in Sentence Production..* Heidelberg: Springer.
[8] Ladd, D.R. 1987 'Metrical' Factors in the Scaling of Sentence-Initial Accent Peaks. *Phonetica* 44: 238-245.
[9] Shinya, T.; Selkirk, E.; Kawahara, S., (2003) Rhythmic Boost and Recursive Minor Phrase in Japanese. Ms., University of Massachusetts at Amherst.