An Acoustic Analysis of the Rise-Fall: A Case of Nonsense Utterances

An Acoustic Analysis of the Rise-Fall: A Case of Nonsense Utterances

Nobuo Yuzawa

Keywords: Rise-Fall, acoustic features, nonsense utterances

1. Introduction

The Rise-Fall tone is presented as an English tone in Roach (2009), along with the Fall, the Rise, the Fall-Rise, and the Level. The distribution of these five tones, however, is not equal, with the Fall being the most frequently used and the Rise-Fall being the least frequently used. This does not mean, however, that in-depth research into the Rise-Fall is unnecessary. In addition to Roach (2009), others have done research into this tone, such as Kingdon (1958), Schubiger (1958), Halliday (1967), Crystal (1969), Gunter (1972), O'Connor and Arnold (1973), Brazil et al. (1980), Ladd (1996), Gussenhoven (1984), Couper-Kuhlen (1986), and Cruttenden (1997), but their research was not conducted acoustically.

The purpose of the present study is to explore key acoustic features of the Rise-Fall. To achieve this purpose, it may be better to use structurally and phonologically simple utterances, such as monosyllabic and bi-syllabic utterances consisting of the same sonorant consonants. It may also be useful to examine whether the position of a consonant in the tonic syllable affects acoustic features of this tone. In bi-syllabic utterances, a trochaic stress pattern may be more useful than an iambic stress pattern to see what the F0 will be like after the tonic syllable.

2. Data

Nonsense utterances are used as acoustic data for analysis in this paper for two reasons. First, it is difficult to find data in the English lexicon that meet the conditions mentioned above. Second, they are semantically neutral. Lexical words may give a speaker certain special connotations attributed to them, which in turn may affect his or her performance even unconsciously while recording. The segments chosen here are /ɑ:/, /n/, and /m/. All of them are sonorants, and no F0 gaps will be made unless the

signal level becomes very low at the end of an utterance. Since there is only one kind of vowel in this study, vowel height factors are not relevant to the F0 measurement.[1]

The following eight nonsense utterances were chosen for the present study: 'A', 'An', 'Na', 'Nan', 'Ama', 'Anma', 'Nama', and 'Nanma'. The underlined syllable indicates a tonic syllable. No special mark is used to represent the Rise-Fall in this paper. The 'A' utterance is the most basic type, consisting of a single vowel /ɑ:/. The 'An' and 'Na' utterances are used to see how a consonant before and after the tonic syllable affects the F0 contours. The 'Nan' utterance is a combined case of the 'An' and 'Na' utterances. The 'Ama', 'Anma', 'Nama' and 'Nanma' utterances are parallel to the 'A', 'An', 'Na' and 'Nan' utterances with an addition of 'ma' in the second syllable. They will be useful in examining how the F0 peak moves rightward in bi-syllabic utterances when they are spoken with the Rise-Fall.

A male British phonetician read each of these utterances five times in a sound-treated room, and the data were recorded with the recording attributes of 11 kHz, 16-bit and mono. No speaker, whether trained or untrained, reads the same utterance identically. A single recording per utterance may not be enough to find out key features of the Rise-Fall. This is the reason why each utterance was read five times. In this paper, these five examples of each utterance are collectively named a set (e.g. the 'A' set). PRAAT was used as the speech analysis software package.

3. Analysis

The total duration is measured to examine whether there is any noticeable difference in duration among the eight sets. Table 1 shows the results.

Sets	#1	#2	#3	#4	#5	Mean
A	654	663	657	670	721	673
An	651	642	674	669	632	654
Na	686	670	676	696	682	682
Nan	692	648	724	724	642	686
Ama	803	708	683	663	669	705
Anma	744	761	698	757	722	736
Nama	759	727	743	743	789	752
Nanma	738	719	747	815	841	772

Table 1 Duration of the utterances in the eight sets

[1] The vowel height in the oral cavity tends to affect the F0 measurements. High vowels tend to have higher F0 values than low vowels. Intensity and duration are also affected by the vowel height. High vowels have less intensity than low vowels, and high vowels are generally shorter than low vowels.

The symbol '#' at the top of the table indicates the order of the five examples appearing in the recording in each set. The unit in the table is milliseconds. Figure 1 shows the mean duration among the eight sets.

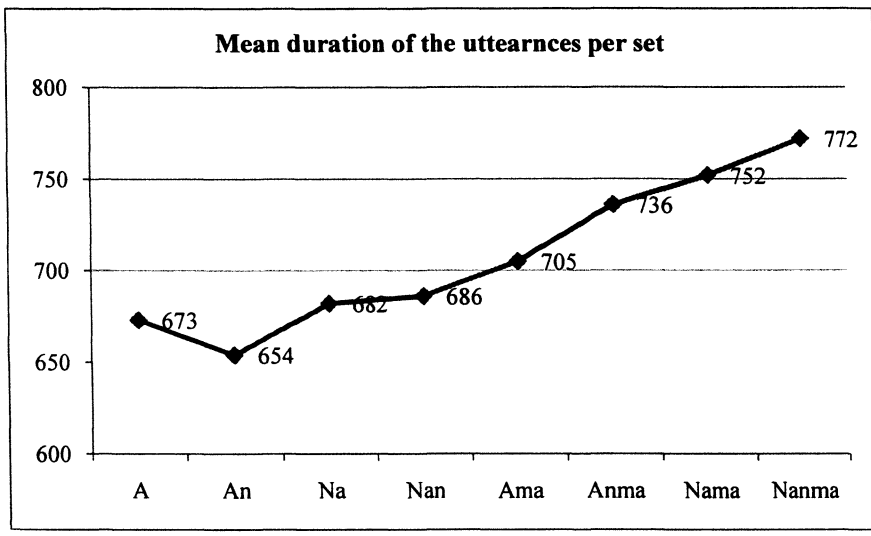


Figure 1 Mean duration of the five utterances per set

From this figure, two features are found. First, the bi-syllabic utterances require more time than the monosyllabic ones. More importantly, however, the bi-syllabic utterances are not twice as long as the monosyllabic ones. The bi-syllabic utterances seem to be squashed in such a way as not to make themselves too long in relation to the monosyllabic ones. An ANOVA test reveals that there is a significant difference in the duration of utterances between the monosyllabic utterances and the bi-syllabic ones ($F = 6.92$, $df = 7/32$, $p < 0.001$).

Source	SS	DF	MS	F	Appx P
Total	101351.9	39			
Treatment	61033.9	7	8719.13	6.92	<.001
Error	40318.0	32	1259.94		

Table 2 ANOVA table[2]

The second feature is that the position of the consonant /n/ in the tonic syllable seems to affect the duration of the utterances consisting of the same phonemes: between the 'An' set and the 'Na' set, and between the 'Anma' set and the 'Nama' set. The

[2] 'SS', 'DF' and 'MS' in Table 2 mean 'sums of squares', 'degree of freedom', and 'mean square', respectively. Treatment and Error are also written as 'within groups' and 'between groups', respectively.

utterance with the post-vocalic /n/ in the tonic syllable (i.e. the 'An' and 'Anma' sets) is shorter than the one with the pre-vocalic /n/ (i.e. the 'Na' and 'Nama' sets). This feature is found more noticeably in the monosyllabic utterances.

As its name implies, the Rise-Fall is shaped like a mountain. To describe such a shape, three points are minimally required: (1) a point where the left slope of the mountain begins, (2) a point corresponding to the peak, and (3) a point where the right slope of the mountain ends. In this paper, these points are named *Point 1*, *Point 2* and *Point 3*, respectively. Figure 2 illustrates how these points are assigned in a real example.

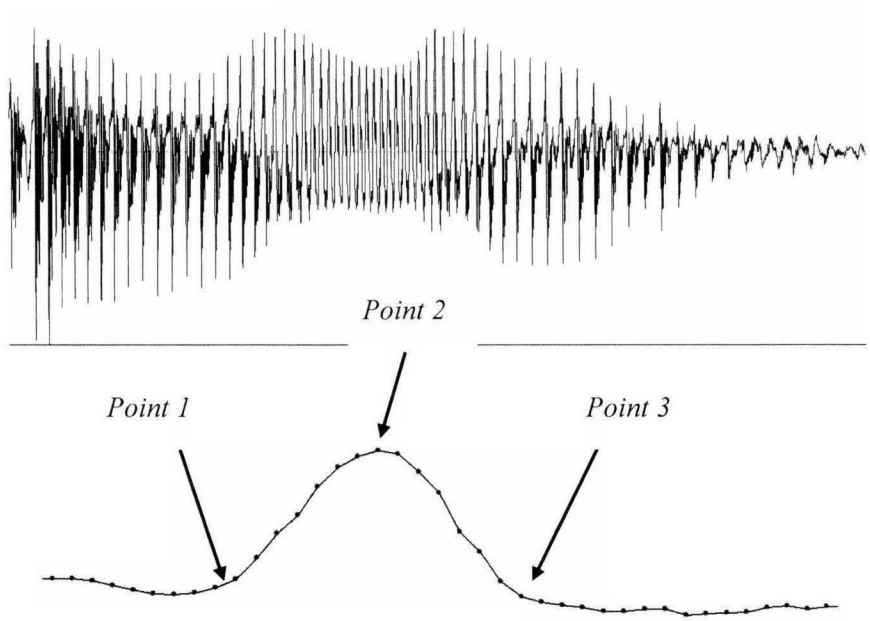


Figure 2 Three points in the Rise-Fall

The example in this figure is taken from the first example of the 'A' set. In order to indicate the location of these three points, duration from the utterance-onset is measured in relation to the total duration of each utterance in the form of a percentage. Because the duration of each utterance is different, it is pointless to compare the location of these points in absolute values. Table 3 shows the horizontal values and the mean values of the three points.

Sets	Points	#1	#2	#3	#4	#5	Mean
A	*Point 1*	25.6	23.1	19.9	23.9	24.2	23.3
	Point 2	44.8	41.1	40.8	43.5	46.2	43.3
	Point 3	60.3	63.6	61.0	64.8	71.1	64.2
An	*Point 1*	24.3	21.0	21.9	18.5	23.1	21.8
	Point 2	43.2	40.7	42.1	41.4	41.6	41.8
	Point 3	62.2	57.3	62.0	62.4	61.3	61.0
Na	*Point 1*	31.6	31.8	30.8	32.5	27.8	30.9
	Point 2	46.7	49.5	49.8	48.3	47.0	48.3
	Point 3	68.7	70.8	69.2	70.0	70.0	69.7
Nan	*Point 1*	33.2	31.6	32.6	31.0	28.8	31.4
	Point 2	51.6	47.9	50.1	48.5	49.1	49.4
	Point 3	69.4	65.0	68.4	65.6	65.9	66.9
Ama	*Point 1*	31.5	31.4	31.3	26.9	29.0	30.0
	Point 2	50.3	50.2	51.9	43.1	51.6	49.4
	Point 3	66.0	69.7	67.9	63.8	65.5	66.6
Anma	*Point 1*	38.2	38.4	33.4	23.9	28.2	32.4
	Point 2	53.4	52.8	54.1	51.3	53.8	53.1
	Point 3	69.7	69.1	70.4	71.4	77.4	71.6
Nama	*Point 1*	36.2	40.8	39.2	38.3	32.6	37.4
	Point 2	53.0	56.5	58.6	57.8	52.6	55.7
	Point 3	69.6	72.0	75.8	73.0	72.1	72.5
Nanma	*Point 1*	30.0	30.9	30.6	47.0	46.6	37.0
	Point 2	59.0	56.5	61.2	64.1	60.2	60.2
	Point 3	79.2	80.2	80.4	80.4	78.8	79.8

Table 3 Location of *Points 1*, *2* and *3* (absolute values)

The mean of the five horizontal values per point is shown in the column 'Mean'. To show more clearly the difference of each value from the mean, Table 4 is provided.

Sets	Points	#1	#2	#3	#4	#5
A	*Point 1*	-2.3	0.2	3.4	-0.6	-0.9
	Point 2	-1.5	2.2	2.5	-0.2	-2.9
	Point 3	3.9	0.6	3.2	-0.6	-6.9
An	*Point 1*	-2.5	0.8	-0.1	3.3	-1.3
	Point 2	-1.4	1.1	-0.3	0.4	0.2
	Point 3	-1.2	3.7	-1.0	-1.4	-0.3
Na	*Point 1*	-0.7	-0.9	0.1	-1.6	3.1
	Point 2	1.6	-1.2	-1.5	0.0	1.3
	Point 3	1.0	-1.1	0.5	-0.3	-0.3
Nan	*Point 1*	-1.8	-0.2	-1.2	0.4	2.6
	Point 2	-2.2	1.5	-0.7	0.9	0.3
	Point 3	-2.5	1.9	-1.5	1.3	1.0
Ama	*Point 1*	-1.5	-1.4	-1.3	3.1	1.0
	Point 2	-0.9	-0.8	-2.5	6.3	-2.2
	Point 3	0.6	-3.1	-1.3	2.8	1.1
Anma	*Point 1*	-5.8	-6.0	-1.0	8.5	4.2
	Point 2	-0.3	0.3	-1.0	1.8	-0.7
	Point 3	1.9	2.5	1.2	0.2	-5.8
Nama	*Point 1*	1.2	-3.4	-1.8	-0.9	4.8
	Point 2	2.7	-0.8	-2.9	-2.1	3.1
	Point 3	2.9	0.5	-3.3	-0.5	0.4
Nanma	*Point 1*	7.0	6.1	6.4	-10.0	-9.6
	Point 2	1.2	3.7	-1.0	-3.9	0.0
	Point 3	0.6	-0.4	-0.6	-0.6	1.0

Table 4 Location of *Points 1*, *2* and *3* (differences from the mean)

It appears that the value of each point is almost the same. There are nine cases where the difference from the mean is 6 or over: (1) #5, *Point 3* in the 'A' set, (2) #4, *Point 2* in the 'Ama' set, (3) #2, *Point 1* in the 'Anma' set, (4) #4, *Point 1* in the 'Anma' set, (5) #1, *Point 1* in the 'Nanma' set, (6) #2, *Point 1* in the 'Nanma' set, (7) #3, *Point 1* in the 'Nanma' set, (8) #4, *Point 1* in the 'Nanma' set, and (9) #5, *Point 1* in the 'Nanma' set. Among these cases, only the first case is related to *Point 3*, and only the second case is related to *Point 2*. The remaining cases are all concerned with *Point 1* in the 'Nanma' set. The reason for such a great difference in each of these cases will be examined in the next section.

In order to examine the more general tendency of these eight sets, attention may need to be paid to the mean point location, the result of which is shown in Figure 3.

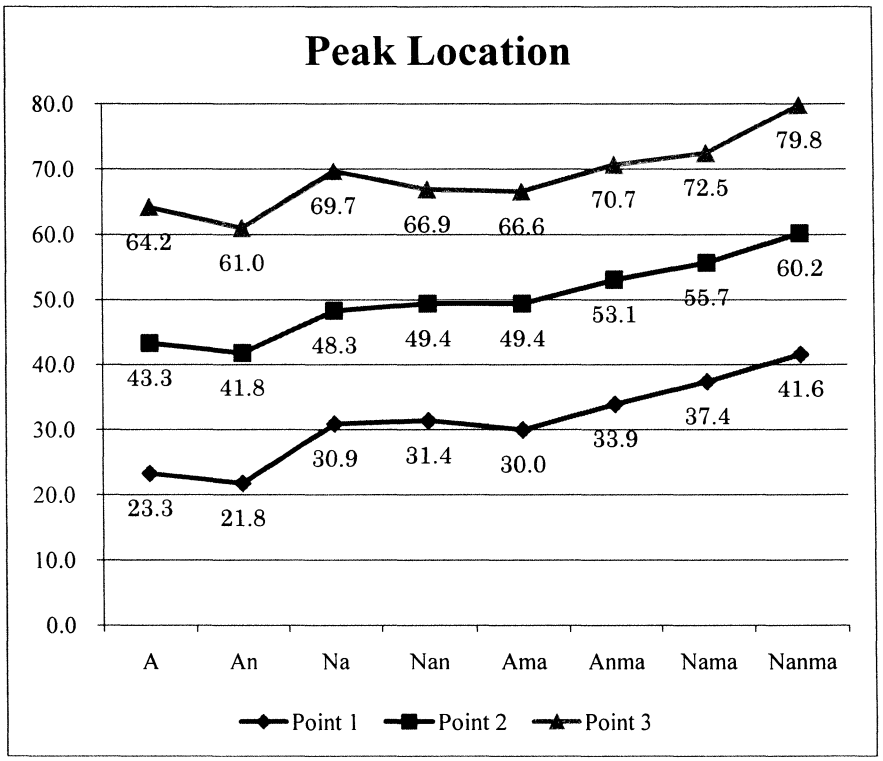
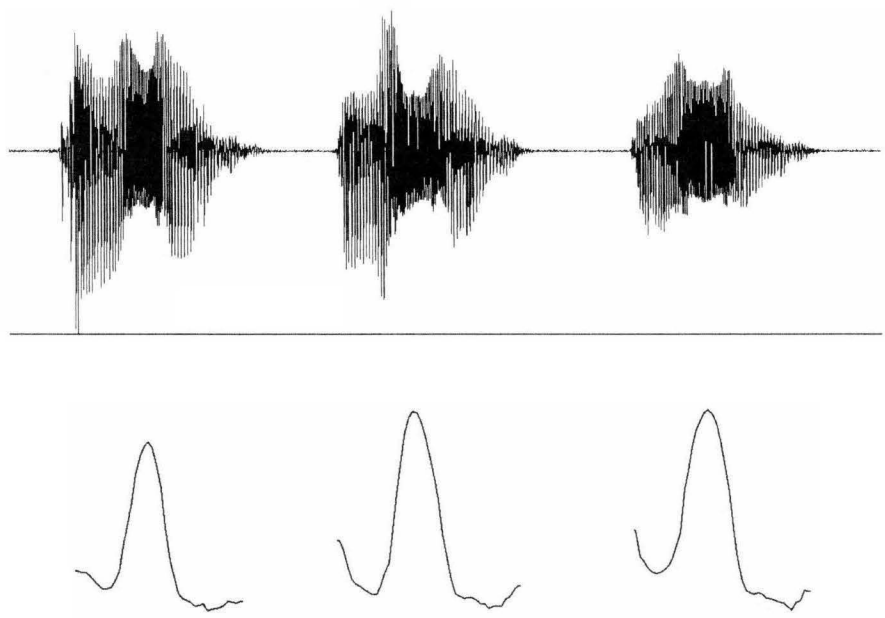


Figure 3 Mean location of the three points

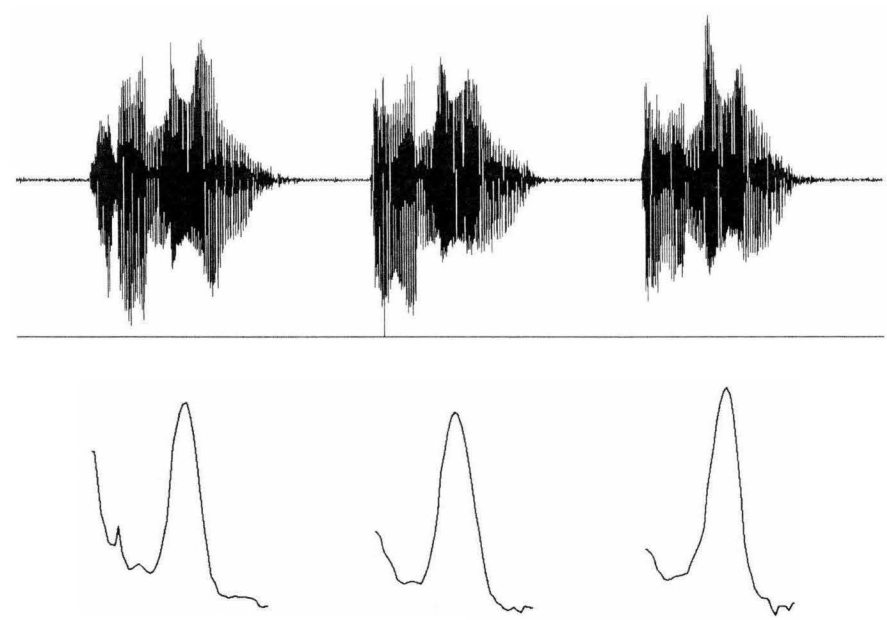
Three features are found in this figure. First, all three points tend to increase their values proportionally from the monosyllabic utterances to the bi-syllabic utterances. Second, the nasal consonant before and after the vowel tends to influence the location of the three points. It tends to move these points forward when placed post-vocalically and delay them when placed pre-vocally. This feature is more noticeable in the monosyllable utterances than in the bi-syllabic utterances. Third, all three points have an equally approximate range of 20%: 20-40% for *Point 1*, 40-60% for *Point 2*, and 60-80% for *Point 3*. 40-60% may be regarded as the peak range for the Rise-Fall.

4. Discussion

To examine acoustic features of the Rise-Fall in more detail, the first three examples of the '<u>A</u>' and '<u>A</u>ma' sets are chosen, as shown in Figure 4. Only these three examples are chosen here because of space limit.



(a)



(b)

Figure 4 Analysis of three examples ((a) '<u>A</u>' set and (b) '<u>A</u>ma' set)

The top part in the figure shows the waveform and the bottom part shows the F0. The F0 scale ranges from 50 to 250 Hz. The F0 contours displayed in this figure are not completely the same, but share the following five important common features.

The first feature is that the F0 contours have a mountain-like shape with a pointed peak. In some cases, however, the top of the mountain is flat, as in the fifth example of the 'A' set, whose acoustic analysis is shown in Figure 5.

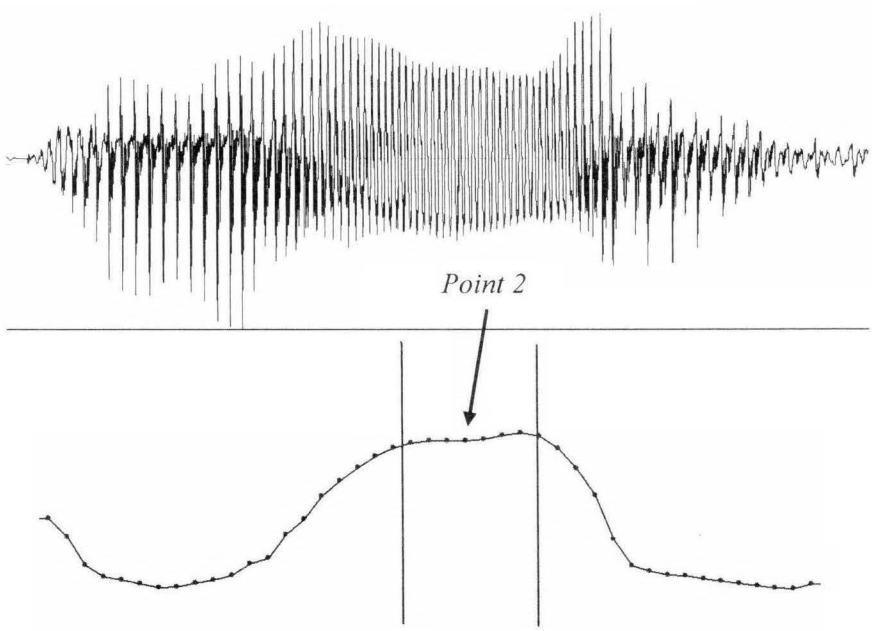


Figure 5 Flat top

In such a case, the peak is measured at the mid-point of the flat top. As shown in the figure, the flat top is delimited by the two vertical lines, and then the peak is decided as the mid-point between the two crossing points. Because of this flat top, *Point 3* is located a little to the right, compared to the other four examples in this set.

The second feature is that the F0 contours begin with a sharp fall and end in a mild fall. Commonly found in the six examples shown in Figure 4 is the fact that there is not so much variation in the F0 contours after *Point 3*, but there are some noticeable changes before *Point 1*. In Figure 4, this feature is most noticeable in the first example of the 'Ama' set. In fact, this feature is found among all the examples in the eight sets, as can be seen in Figure 6.

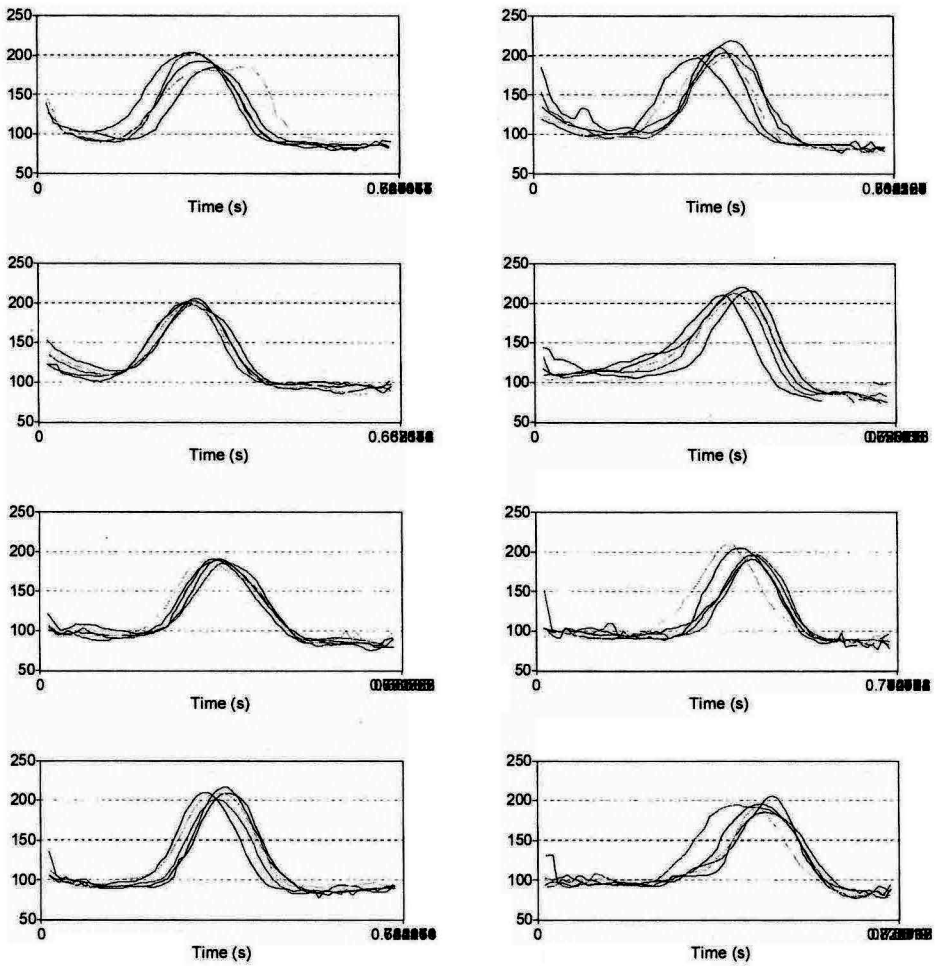


Figure 6 The superimposed F0 contours per set

This figure shows the superimposed F0 contours of the five examples in the eight sets. On the left column, from top to bottom, the F0 contours for the '<u>A</u>' set, the '<u>An</u>' set, the '<u>Na</u>' set, and the '<u>Nan</u>' set are described. On the right column, from top to bottom, the F0 contours for the '<u>A</u>ma' set, the '<u>An</u>ma' set, the '<u>Na</u>ma' set, and the '<u>Nan</u>ma' set are described. The numbers on the horizontal axis, which mean duration, are not clearly shown because all five examples in each set are not identical in duration. It appears from this figure that this second feature is noticeable in the '<u>A</u>' set, the '<u>An</u>' set, the '<u>A</u>ma' set, and one example in the '<u>An</u>ma' set, where the tonic syllable does not begin with the consonant /n/. Something else that may need to be pointed out for the second feature is that all five examples in each set show almost the same type of F0 contour though the speaker said after the recording that he intentionally tried to speak these examples differently. These close matches look more noticeable in the monosyllabic sets rather than in the bi-syllabic sets.

The third feature is that the utterance-initial part is higher in pitch than the utterance-final part. As seen in Figure 6, this feature seems to be more noticeable when there is no onset in the tonic syllable. In short, in an utterance spoken with the Rise-Fall, the tonic syllable tends to be higher pitched word-initially when there is no onset than when there is an onset.

The fourth feature is that the peak in the bi-syllabic utterances is located in the second syllable. This feature may not be understood well simply from Figure 4 because spectrographic information is deleted to clarify the F0 in black and white printing. To compensate for the weakness of this figure, Figure 7 is presented.

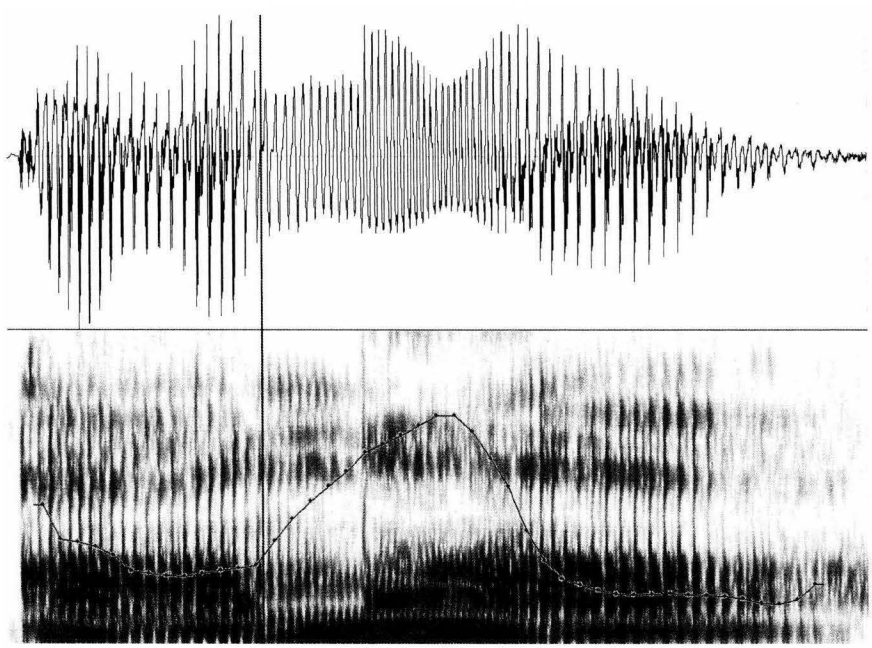


Figure 7 Analysis of the fifth example of the '<u>A</u>ma' set

This figure shows the acoustic analysis of the fifth example of the '<u>A</u>ma' set. At the bottom are included both the F0 contour and the spectrogram of this example. The vertical line indicates the boundary between the first syllable /ɑ:/ and the second syllable /mɑ:/. Here the peak is located in the vowel of the second syllable. One exception is found in the fourth example of the '<u>A</u>ma' set, where the peak is located on the consonant. Its acoustic analysis is shown in Figure 8.

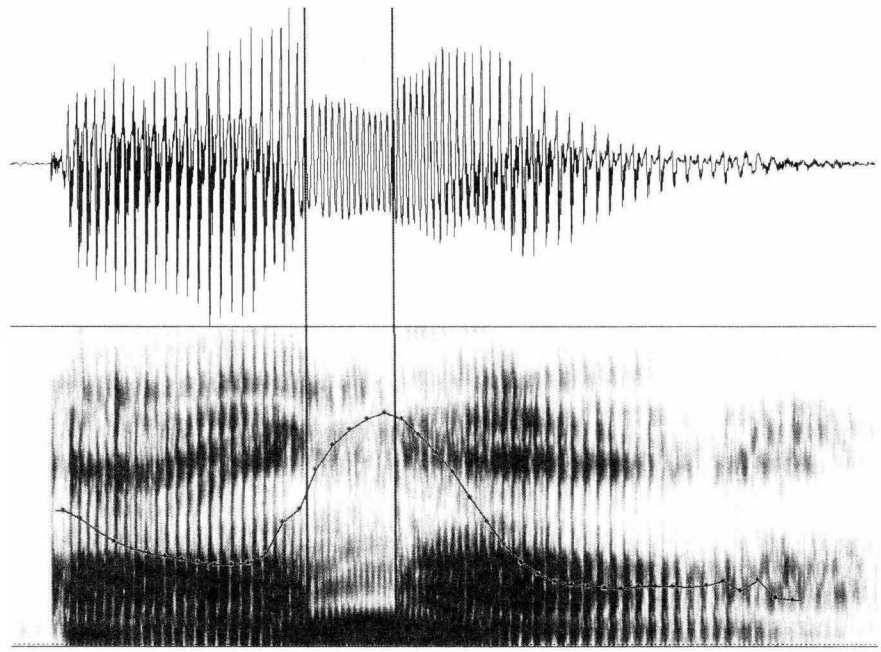


Figure 8 Analysis of the fourth example of the 'Ama' set

The two vertical lines delimit the intervocalic consonant /m/. As this figure shows, the peak is located at the end of this consonant. In contrast, in the other bi-syllabic utterances, the peak is located on the second vowel. This explains, as seen in Tables 3 and 4, why the peak is located much earlier in this example than in the other four examples. What is more important here is that in both cases the peak is located in the second syllable, and this is the feature commonly found among all the bi-syllabic examples used in this paper. Not every initial consonant in the second syllable, such as /p/ and /f/, however, may be able to function as the /m/. It is assumed that such a consonant needs to have a vowel-like quality. One important feature found in a vowel and the /m/ is that both of them are sonorants.

The fifth feature is that when there are two nasals between the tonic syllable and the second syllable, where the first nasal belongs to the tonic syllable and the second nasal belongs to the second syllable, *Point 1* may be located somewhere between the two nasals. As seen in Tables 3 and 4, *Point 1* in the 'Nanma' set is located very differently in all five examples. All of them have a difference of 6 or more from its mean: 7.0, 6.1, 6.4, -10.0 and -9.6. These five values, however, may be divided into two categories: the first three (all positive values) and the last two (all negative values). In fact, the first three examples have a similar F0 contour, and the same is true of the last two examples. This may be explained clearly by comparing the third and fourth examples acoustically, as shown below.

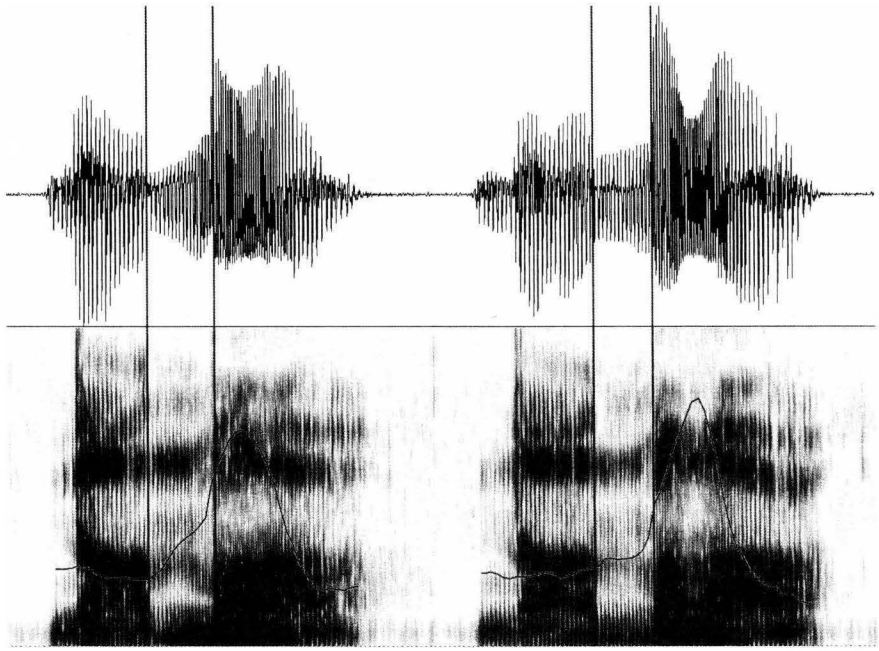


Figure 9 Analysis of the third and fourth examples of the 'Nanma' set

The two vertical lines delimit the intervocalic nasal consonants /nm/. As this figure shows, the location of *Point 1* is noticeably different. This point is located on the first consonant /n/ in the left panel, but on the second consonant /m/ in the right panel. The same feature found in the left panel is also found in the first two examples of the 'Nanma' set, and the same feature found in the left panel is also found in the fifth example of this set. This may explain clearly why there are many differences in the location of *Point 1* from its mean. It seems that when there are two sonorants like nasals between the two vowels in bi-syllabic utterances, the peak may come somewhere between the two consonants. In fact, the same features are also found in the 'Anma' set, where *Point 1* is located on the /m/ in the first two examples, on the /n/ in the last two examples, and between the two nasals in the third examples.

It may be concluded safely from the above discussion that when bi-syllabic utterances are spoken with the Rise-Fall, the tonic syllable is low pitched and the peak is located in the second syllable. In similar data that the author has, where utterances are spoken with the Fall, the peak is always located in the tonic syllable. This may help to support the statement that one important key feature of the Rise-Fall is peak delay especially when this tone is contrasted with the Fall.

In this paper, monosyllabic and bi-syllabic utterances were used. It was learned from the data used in the present study that the peak location has a range of 40-60% from the utterance-onset. In Yuzawa (forthcoming), where real data spoken by six speakers were used, it was learned that this range needs to be extended to 30-60%.

This may be because real data have more complex segmental structures and different speakers may have different timing in speaking.

5. Concluding Remarks

In this paper, monosyllabic utterances and bi-syllabic utterances were used to examine acoustic features of the Rise-Fall. Three points were selected as key points in the F0 contour for this tone. They were named *Point 1*, *Point 2* and *Point 3* from the utterance-onset.

The present study found seven major features about acoustic characteristics of the Rise-Fall. First, the F0 contours for the Rise-Fall have a mountain-like shape, mostly with a pointed peak. There are cases in which the mountain top is flat. In such cases, the peak is equated with the mid-point of the flat top. This flat top may move the location of *Point 3* slightly to the right. Second, the F0 contour begins with a sharp fall to *Point 1* and ends in a mild fall from *Point 3*. This feature is more noticeable when the utterance begins with a vowel. Monosyllabic utterances have more similarity in the overall shape among the five examples in each set than bi-syllabic utterances. Third, the utterance-initial part is higher in pitch than the utterance-final part. This feature is more noticeable when there is no onset in the tonic syllable. Fourth, in bi-syllabic utterances, the peak comes in the second syllable. It is mostly located on the vowel, but when there is a pre-vocalic nasal in the second syllable, the peak may be located on this consonant. This peak delay is one important key feature of the Rise-Fall. Fifth, when there are two nasals between the tonic syllable and the second syllable (the first nasal belongs to the tonic syllable and the second belongs to the second syllable), *Point 1* may be located somewhere between the two nasals, which makes the location of this point very different. Sixth, all three points have an equally approximate range of 20%: 20-40% for *Point 1*, 40-60% for *Point 2*, and 60-80% for *Point 3*. 40-60% may be regarded as the peak range for the present utterances spoken with the Rise-Fall. This range, however, may need to be extended to 30-60% when real data are taken into account. Seventh, in utterances consisting of the same phonemes, the position of the consonant /n/ in the tonic syllable seems to affect not only the duration of these utterances but also the location of the three points. In the former, the utterances with the post-vocalic nasal in the tonic syllable tend to be shorter than the ones with the pre-vocalic nasal. In the latter, the three points tend to move forward in the utterances with the post-vocalic nasal, compared to the ones with the pre-vocalic nasal. These two features are more noticeable in the monosyllable utterances.

One future study related to the present research may include comparison of the Rise-Fall with the Fall in order to examine how similar and different they are acoustically.

References

Brazil, D. C., Coulthard, M. R. and Jones, C. (1980). *Discourse Intonation and Language Teaching*. London: Longman.

Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. London: Edward Arnold.

Cruttenden, A. (1997). *Intonation*. Second edition. Cambridge: Cambridge University Press.

Crystal, D. (1969). *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Gunter, R. (1972). 'Intonation and Relevance.' In D. Bolinger (ed.), *Intonation*, pp. 194-215. London: Penguin Books Ltd.

Gussenhoven, C. (1984). *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.

Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague: Mouton.

Kingdon, R. (1958). *The Groundwork of English Intonation*. London: Longman.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

O'Connor, J. D. and Arnold, G. F. (1973). *Intonation of Colloquial English: A Practical Handbook*. Second Edition. London: Longman.

Roach, P. (2009). *English Phonetic and Phonology: A Practical Course*. Fourth Edition. Cambridge: Cambridge University Press.

Schubiger, M. (1958). *English Intonation: Its Form and Function*. Tubingen: Niemeyer.

Yuzawa (forthcoming). 'Acoustic Features of the Rise-Fall in English: A Case of the Monosyllabic Tonic Element.' *English Phonetics* 14. The English Phonetic Society of Japan.

Abstract

The Rise-Fall tone is acoustically analysed using nonsense utterances to find out key features of this tone. These utterances consist of four types of monosyllabic utterances (e.g. 'A') and four types of bi-syllabic utterances (e.g. 'Ama'), all of which have the same phonemes: /ɑ, n, m/. Seven major features are found in the present study. One of them is that the peak in bi-syllabic utterances is located in the second syllable. It is mostly located on the vowel, but when there is a pre-vocalic nasal in the second syllable, the peak may be located on this consonant. Another is that the peak ranges from the 40% position to the 60% position form the utterance-onset.

(Associate Professor, Utsunomiya University)