

Revisiting Vowel Coalescence in Japanese

Koji Ono

1. Introduction

Vowel coalescence, which is widely observed in Japanese both diachronically and synchronically, is a phonological process in which two different adjacent vowels merge into a monophthong, for example, [sugoi] → [suge:], [omae] → [ome:], and [samui] → [sami:]. As Kindaichi (1976) points out, the immediate cause of this phenomenon is assumed to lie in the strong tendency for Japanese to avoid any vowel sequence within a morpheme. This tendency presumably dates as far back as the 8th century, and in fact, Kindaichi (1976) claims that no vowel sequence can be found in the oldest writings like *Kojiki* (Records of Ancient Matters) (A.D.713), the *Nihon Shoki* (Chronicles of Japan) (A.D.720) or the *Man'yooshuu* (Collection of a Myriad Leaves) (A.D.759).

As convincing as Kindaichi's (1976) explanation for vowel coalescence appears to be, two questions remain unanswered; that is, what vowel sequence can cause coalescence, and what phonetic form can be created as a result? Since the former is related to the input of coalescence and the latter to the output, the main purpose of this paper is to specify the properties of each level and set constraints for them.

2. A Feature-based Analysis

2.1. *A Brief Outline of Kubozono (1999)*

Typical examples of vowel coalescence in Japanese are given in (1) and (2), where the underlined segments to the left of the arrow refer to the input of coalescence and those to the right, the output. Hence, by coalescence we mean that two input vowels change into a specific monophthong, long or short.

- (1) a. [ai] → [e(:)]
 nagaiki → nageki 'grief' amai → ame: 'sweet'
- b. [au] → [o:]
 auta → o:ta 'met(past)' awaumi → aumi → o:mi 'O:mi (place name)'
- c. [eu] → [o:]
 tefu → teu → tyo:: 'butterfly' kefu → keu → kyo: 'today'
- d. [oi] → [e:]
 sugoi → suge: 'great' hidoi → hide: 'terrible'
- (2) a. [ae] → [e:]
 omae → ome: 'you' kaeru → ke:ru 'to return'

- b. [ao] → [o]
hataori → hattori ‘Hattori (person’s name)’
- c. [eo] → [o(:)]
meoto → myo:to ‘couple’ miteoku → mitoku ‘to see’
- d. [oe] → [e:]
sokoe → soke: ‘there’ dokoe → doke: ‘where’
- e. [iu] → [u:]
riu → ryu: ‘dragon’ iu → yu: ‘to say’
- f. [ui] → [i:]
samui → sami: ‘cold’ kayui → kaii: ‘itchy’
- g. [ea] → [a(:)]
miteageru → mitageru ‘I will see it for you’
orewa → orea → ora(:) ‘I’
- h. [oa] → [a]
(soNnako)towa → (soNnako)toa → (soNnako)ta: ‘such a thing’
- i. [ei] → [e:]
rei → re: ‘courtesy’ kirei → kire: ‘pretty’
- j. [ou] → [o:]
tou → to: ‘to ask’ omou → omo: ‘to think’

These examples demonstrate the diversity and complexity of the results of vowel coalescence and seem to preclude the formulation of rules or constraints that cover all cases in (1) and (2). Kubozono (1999), however, addresses these difficulties with his epoch-making rule of coalescence:

(3) Vowel Coalescence Rule (Kubozono 1999:103)

[αhigh, δlow, εback] [ζhigh, βlow, γback] → [αhigh, βlow, γback]

Simplicity, above all, is the distinguishing feature of this rule. In fact, all we have to do to produce a correct coalescence form is to combine the first vowel’s [high] with the second vowel’s [low] and [back]. To demonstrate how (3) actually works, let us first look at the Japanese vowel system, which consists of five members, each distinguished by the three features [high], [low] and [back]. Since this rule depends on phonetic features, it may well be called a ‘feature-based’ rule.

(4) [i]: [+high, -low, -back] [u]: [+high, -low, +back]
[e]: [-high, -low, -back] [o]: [-high, -low, +back]
[a]: [-high, +low, -back]

The representation given in (4) and the rule in (3) enable us to get the correct derivations of (1) and (2). Three are exemplified below:

- (5) a. [ai] → [e] (= (1a))
 [-high, +low, -back] + [+high, -low, -back] → [-high, -low, -back]
 b. [ao] → [o] (= (2b))
 [-high, +low, -back] + [-high, -low, +back] → [-high, -low, +back]
 c. [ea] → [a] (= (2g))
 [-high, -low, -back] + [-high, +low, -back] → [-high, +low, -back]

At first glance these derivations do not appear to raise any problems with respect to vowel coalescence, but it should be noted that Kubozono's (1999) feature-based analysis involves a serious problem in itself, which will be discussed in detail in the next section.

2.2. Problems with a Feature-based Analysis

As suggested above, the most serious drawback of Kubozono's rule is the 'overgeneralization' caused by (3). More specifically, the rule creates not only well-formed phonetic forms, as we have seen in the previous section, but also several ill-formed ones. Consider the following derivations Kubozono (1999) regards as adequate but which are not attested in Japanese.

- (6) [ie] → *[i]: [+high, -low, -back] [+high, -low, +back] → * [+high, -low, -back]
 [io] → *[u]: [+high, -low, -back] [-high, -low, +back] → * [+high, -low, +back]
 [ue] → *[i]: [+high, -low, +back] [-high, -low, -back] → * [+high, -low, -back]
 [uo] → *[u]: [+high, -low, +back] [-high, -low, +back] → * [+high, -low, +back]

It is obviously impossible to distinguish difference between the correct derivations in (4) and the incorrect ones in (6) by relying solely on Kubozono's coalescence rule (3), because it allows any combination of features if its conditions are met. We claim that the problem of overgeneralization discussed here is due mainly to the simplified explanation made by a feature-based analysis. In other words, the coalescence process is so complex that it is difficult to account for it through phonetic features alone. In the next section, we will propose that factors other than phonetic features must be taken into consideration in order to explain the complicated behavior of coalescence.

Besides the problem of overgeneralization, there is another difficulty with an analysis relying solely on phonetic features; that is, no logical basis can be found in (3) for combining the first vowel's [high] with the second vowel's [low] and [back]. In fact, there are six combinational possibilities, namely the first vowel's [low] plus the second vowel's [high] and [back], the first vowel's [high] and [back] plus the second vowel's [low], and so on. The question here is why Kubozono (1999) chooses the one combination shown in (3). Are there any specific reasons for this, or is it just a

coincidence? Kubozono himself (1999, 2005) offers no answers to these questions.

3. An Analysis Based on Sonority Hierarchy

3.1. *Sonority Constraint*

One thing that Kubozono (1999) neglects in analyzing vowel coalescence in Japanese is the fact that the resulting forms are classified into two types: Those that are completely different from the two input vowels as in (1) and those that are the same as in (2). Where do these forms come from, and what causes their development? Clarification of these questions is our immediate objective, and to achieve it, the first thing we need to do is to make explicit the reason why the derivations in (6) are unacceptable.

Consider all the examples in (7) which are not subject to vowel coalescence; the first four are from (6) and the last two (i.e. [ia] and [ua]) are introduced here:

(7) [ie] [io] [ue] [uo] [ia] [ua]

These examples indicate that the first vowels are less sonorous than the second from the view point of the sonority degree based on the universal sonority hierarchy show in (8a). (8b), which is based on (8a), illustrates the concrete difference in degree of each sequence in (7):

(8) a. high vowel [i, u] < mid vowel [e, o] < low vowel [a]
 b. $i < e$ $i < o$ $u < e$ $u < o$ $i < a$ $u < a$

Bearing in mind that the examples in (7) do not undergo coalescence, we can conclude that coalescence is not allowed when the first vowel is less sonorous than the second; conversely, coalescence takes place when the first vowel is either more sonorous than or the same as the second vowel. The examples in (2) and (3), as shown in (9), demonstrate the validity of this conclusion. Furthermore, contrary to our prediction, coalescence still occurs in [ea] and [oa] even though the second, not the first, vowel is more sonorous. We will discuss these examples in detail later.

(9) $a > i$ $a > u$ $e > u$ $o > i$ $a > e$ $a > o$ $e = o$ $o = e$ $i = u$ $u = i$
 $e < a$ $o < a$ $e > i$ $o > u$

The observations given above allow us to formulate the following constraint (cf. Ono 2001, Kusters (2004)):

(10) Sonority Constraint

Coalescence of two vowels takes place when the first vowel is more sonorous than or the same as the second vowel.

Note that this constraint not only specifies the circumstance in which coalescence occurs but gives a principled explanation of why the sequences in (6) are not subject to it.

3.2. Vowel Power Constraint

In the preceding section we presented a constraint whose function is mainly to restrict the possible combinations of input vowels. Let us now turn to another constraint, whose function is to restrict the possible resulting vowels (i.e. outputs) of coalescence.

In order to formulate such a constraint, we first assume that each vowel has its own power that is determined in accordance with the sonority hierarchy given in (8a): We claim that the more sonorous a vowel is, the more power it has. Furthermore, we posit that the Japanese high vowels ([i] and [u]) are far less sonorous than the other three vowels ([a], [o] and [e]). Since the latter assumption is more crucial to our analysis, we will provide independent evidence to support it. First, a candidate of vowel insertion must be either [i] or [u] when words or morphemes are borrowed from foreign languages. Most Japanese syllables end with vowels to formulate a basic pattern of CV (C = consonant and V = vowel), but if borrowed words do not fit this pattern (in other words, if one syllable involved in those words ends with CVC), a vowel must be inserted after the second consonant so as to create the native phonetic pattern of CVCV. In such cases, [i] or [u] is regularly chosen. This is illustrated in (11); examples in (11a) are from Chinese and those in (11b) from English. Note that the underlined vowels in each case do not exist in the original languages.

- (11) a. eki (駅, 益...) seki (積, 籍...) kaku (核, 格...) satu (冊, 札...)
 b. buku 'book' ke:su 'case' bi:ti 'beach' ke:ki 'cake'

Preservation of the original pronunciation of the borrowed word is highly desirable. Therefore, even if some vowel is introduced to convert them into native words, it must be one that is the least damaging to the original pronunciation. In other words, the vowels chosen must be the least corrupting and conspicuous, and as far as Japanese is concerned, [i] and [u] are such vowels.

In addition to being the least corrupting, [i] and [u] are also much less powerful than the others, especially when they are flanked by voiceless consonants. More specifically, they are apt to be devoiced when such consonants stand between them. Note that the underlined vowels in (12) undergo devoicing:

- (12) hikari 'light' sikeN 'examination' kusa 'grass' tuki 'moon'

Since [a], [o] and [e] generally do not undergo devoicing, (12) indicates that [i] and [u] are readily affected by the context, in this case by flanking voiceless consonants in this case. This implies that they are less powerful and less independent than the other three vowels.

Finally, let us consider euphony (*onbin*), which deletes an onset velar (usually,

[k]) preceding [i] or [u]. Note that other vowels do not participate in this process. For example, the underlined velar in (13) is removed by euphony:

(13) kaki + ta → kaita ‘wrote’ arigataku → arigatau → (arigato:) ‘thank you’

The direct cause that triggers euphony has not yet been determined, but one thing for sure is that Japanese [i] and [u] are less powerful as the nucleus of a syllable than the other vowels, so that by deleting the onset ([k] in (13)) that precedes them, they can easily become part of the preceding syllable; [ka] and [ta] in (13) turn to [kai] and [tau], respectively.

Based on the evidence given above, it seems correct to assume that [i] and [u] can be definitely differentiated from the other vowels in terms of vowel power. Specifically, there are good reasons to assume that [i] and [u] are much less powerful than [a], [o] and [e]. This assumption can be represented numerically, as illustrated in (14) below, where [a] is considered the strongest with the value 4, while [i] and [u] are the weakest with the value 1. What is crucial here is that the numerical difference between [i, u] and [e, o] is larger than that between [e, o] and [a], though the arrangement of each vowel is the same as that of (8) (cf. Ono (2004, 2005)):

(14) Hierarchy of Vowel Power

$$\begin{array}{cccc} i, u < e, o < a \\ 1 & 3 & 4 \end{array}$$

By establishing (14), we are now in a position to postulate a constraint that determines what form is appropriate as an output for coalescence. Let us first examine the vowel sequences in (2a-f), where the coalescence forms are the same as the second vowels of the sequences. As (15) shows, the difference in vowel power between the first vowel (VP1) and the second (VP2), namely, VP1 – VP2, is +1 or 0.

(15) [ae]: a – e = 1 [ao]: a – o = 1 [eo]: e – o = 0 [oe]: o – e = 0
[iu]: i – u = 0 [ui]: u – i = 0

Given this result, it seems reasonable to conclude that if VP minus VP2 is less than 2, the second vowel becomes the output of coalescence.

The situation turns out to be even more complicated, however, if we consider the examples in (2g, f), that is, [ea] and [oa]. Recall that both of them are also subject to coalescence. The problem is that contrary to expectation, their coalescence form (i.e. [a]) is the same as the second vowel in spite of the fact that VP1 minus VP2 here is –1 rather than +1, as shown in (16):

(16) [ea]: e – a = –1 [oa]: o – a = –1

This problem, however, ceases to exist once we notice that +1 in (15) and –1 in (16) constitute the absolute value 1, namely, |1|. We can thus propose a

constraint like (17) in order to account for (15) and (16) at the same time.

(17) Vowel Power Constraint

If the difference in vowel power between the input two vowels is $|1|$ or 0, then the second vowel will be an output.

It should be noted that unlike Kubozono's feature-based rule, this constraint (17) enables us to explain why each example of (2a-h) has the second vowel as an output.

It is true that (17) explains why [ea] and [oa] in (2g, h) or in (16) can combine into one vowel, but recall that there still remains an unsolved problem in connection with the Sonority Constraint (10); that is, [ea] and [oa] obviously violate it in that their second vowel (i.e. [a]) is more sonorous than the first (i.e. [e] and [o]), which is the opposite of what the Sonority Constraint requires.

To solve this problem, we need to discover a characteristic that [ea] and [oa] share — but not the sequences in (7) — and incorporate it into the Sonority Constraint as a collateral condition. If we recall that neither obeys the Sonority Constraint and only the former (i.e. [ea] and [oa]) are allowed to be coalesced, then it is necessary to discover the characteristic which makes us distinguish them. This characteristic can be obtained rather easily, though, when the differences in the vowel power between them are considered. The difference between [ea] and [oa] is -1 , as illustrated in (16), while that in (7) is -2 or -3 , as illustrated below.

$$(18) \quad \begin{array}{lll} [\text{ie}]: i - e = -2 & [\text{io}]: i - o = -2 & [\text{ue}]: u - e = -2 \\ [\text{uo}]: u - o = -2 & [\text{ia}]: i - a = -3 & [\text{ua}]: u - a = -3 \end{array}$$

It follows that vowel coalescence is impossible when the numerical values of VP minus VP2 is either -2 or -3 , but possible when it is -1 . Given this fact, we can propose a revised version of Sonority Constraint with a collateral condition concerning vowel power.

(19) Sonority Constraint (Revised)

Coalescence of the two vowels takes place when the first vowel is more sonorous or the same as the second vowel, with the provision that the difference in vowel power between them is more than -1 .

(19) makes it possible not only to permit [ea] and [oa] to merge into a single vowel but also to prevent the examples in (7) from coalescing.

The final examples to be examined in this section are [ei] and [ou] in (2i, j) whose coalescence forms are the first vowels (i.e. [e] and [o] respectively), not the second (i.e. [i] and [u] respectively) contrary to expectation. This can be attributed to the fact that the numerical value of VP1 minus VP2 is $+2$, as follows:

$$(20) \quad [\text{ei}]: e - i = 2 \quad [\text{ou}]: o - u = 2$$

The computation in (20) leads us to conclude that if VP1 minus VP2 equals $+2$, the

output will be the second vowel of the sequences. This conclusion can be added to the output constraint (21), which is a revised version of (17).

(21) Vowel Power Constraint (Final Version)

- a. If the difference in vowel power between the two vowels is $|1|$ or 0 , then the second vowel will be an output.
- b. If it is $+2$, then the first vowel will be an output.

Thus, by postulating (21), we can derive all the output forms in (2) including [ea] and [oa], which would otherwise be considered exceptions to our constraints.

3.3. Vowel Distance Constraint

We have so far examined cases where a coalescence form is the same as either the first or the second vowel. Let us now move on to the examples in (1) where the output differs from both of the input two vowels. Here, too, we will evaluate the numerical values of the vowel power between the first and second vowel in (1):

(22) [ai]: $a-i = 3$ [au]: $a-u = 3$ [eu]: $e-u = 2$ [oi]: $o-i = 2$

Of these four, only [eu] and [oi] are problematic. Since their numerical value of VP minus VP2 is $+2$, they are mistakenly subject to the constraint (21b) and their first vowels (namely, [e] and [o]) would be regarded as the correct outputs of coalescence. But this is not true because their output is neither the first nor the second vowel. In fact, it is located between them, that is, [o] and [e], respectively. To understand the meaning of ‘located between’ accurately, let us look at the triangle of Japanese vowels:

(23)

i		u
	e	o
	a	

It is clear from this chart that [e] lies between [a] and [i] and [o] between [a] and [u], and likewise, [o] and [e] may well be considered to be located between [eu] and [oi] respectively.

Provided that the adjacent vowels in (23), i.e. [ae], [ao], [eo], [oe], [iu], [ui], etc., which are actually included in (2), have a minimum distance ‘Distance 1’ (D1), then the vowels sequences in (22) or (1) can be assumed to have ‘Distance 2’ (D2) because their constituent vowels are not adjacent to each other. We thus present the following output constraint for (1) on the basis of these assumptions.

(24) Vowel Distance Constraint

If the distance between input two vowels is $+2$, then the output will be a vowel located between them; otherwise it will be either the first or the second vowels.

So far we have proposed three constraints; the Sonority Constraint (19), the

Vowel Power Constraint (21) and the Vowel Distance Constraint (24). But they do not apply to vowel sequences simultaneously. In fact, there is an order of application: The Sonority Constraint applies first, the Vowel Distant Constraint next, and the Vowel Power Constraint last. The Sonority Constraint applies first because it determines whether the vowel sequence in question is qualified to coalesce in the first place. Once the constraint permits coalescence, the Vowel Distance Constraint applies next to ascertain whether the output lies between the two input vowels or not. Finally, after the VDC confirms that the output is not located between the two input vowels, the Vowel Power Constraint determines which vowel, the first or the second, is chosen as the output.

4. Conclusion

This paper has given a principled account of vowel coalescence in Japanese. In so doing, we first introduced Kubozono (1999), the only analysis using a clear and simple rule to explain it, and pointed out the critical problem it has, namely, the problem of overgeneralization. After demonstrating the limitation of Kubozono's (1999) feature-based analysis, we proposed an alternative, a sonority-based analysis, which is equipped with three constraints: The Sonority Constraint, the Vowel Power Constraint and the Vowel Distance Constraint. These constraints allow us not only to explain the intricate behavior of vowel coalescence but also to avoid the problem of 'overgeneralization' caused by Kubozono's approach.

REFERENCES

- Kindaichi, Kyouzuke (1976) *Nihongo no Hensen* (Historical Development of Japanese), Kodansya, Tokyo.
- Kubozono, Haruo (1999) *Nihongo no Onsei* (The Sound System in Japanese), Iwanami, Tokyo.
- Kubozono, Haruo (2005) "[ai]-[au] Asymmetry in English and Japanese," *English Linguistics* 22-1, 1-22, Kaitakusya, Tokyo.
- Kusters, Harold (2004) "Vowel Alternation in Modern Japanese and Sonority," *Phonological Studies* 7, 25-32, Kaitakusya, Tokyo.
- Ono, Koji. (2001) "Nihongo no Boinyuugoo ni Tuite (On Vowel Coalescence in Japanese)," *Imi to Katachi no Intaafaisu 2 (Interface of Meaning and Form 2): A Festschrift for Minoru Nakau on the Occasion of His Sixtieth Birthday*, ed. by Shyuuichi Takeda, Aturo Tubomoto, Yukio Hirose, Koichi Takeda and Nobuhiro Kaga, 885-896, Kuroshio, Tokyo.
- Ono, Koji. (2004) "Vowel Coalescence and Vowel Alternation in Japanese," *Journal of the Faculty of Culture and Education* 9-1, 107-115, Saga University.
- Ono, Koji. (2005) "Sonoritii to Nihongo (Sonority and Japanese)," *Journal of the Faculty of Culture*

and Education 10-1, 125-133, Saga University.

Faculty of Culture and Education

Saga University

email: onok@cc.saga-u.ac.jp