Child Phonological Development in Vowel Sequences and some Proposals concerning Japanese VV Classification

Martin GORE *

Received 30 October, 2008

Abstract. After a review of perceptual experiments, this paper makes proposals concerning the classification of vowel sequences in Japanese, focusing firstly on children's perceptual development, and secondly on mature phonology. It is hoped that the proposals presented will allow (1) an impartial framework within which to compare groups of sequences (some of which have been proposed in the literature as diphthongs), (2) a theoretically neutral reclassification that does not rely on the presupposition of syllable as a phonological unit, and (3) an understanding of why the concept of diphthong is attractive but not essential in the analysis of a mora-timed language.

Keywords: Phonetics; perception; vowel sequence; syllable / mora; constraint-based; optimality theory; phonological development; Japanese; child phonology

1. Introduction

There have been occasional but repeated suggestions in the literature that some vowel sequences in Japanese may qualify for membership of a special category called "diphthong" (e.g. McCawley 1968; Haraguchi 1996; Kubozono 1999). These assertions have sometimes contradicted each other, and this has been analyzed in detail by Gore (2006), who concluded that though there was little justification for the diphthong in mainstream Japanese phonology, there was something to be said for considering a few VVs as unbreakable units either morphophonologically, or perceptually, especially at an early developmental stage of child phonology.

In the following sections we adopt a view suggested by Jusczyk (1997) that the young child's ranking of constraints may change subtly as the child develops. This would potentially allow perception of a VV as a diphthong at one developmental level, while disallowing it at another. This approach would go some way to explaining why some researchers find the concept of the diphthong admissible even in a

* Professor, Faculty of Education, Kagoshima University

phonology that is nominally based on the mora.

2. Background: experiments concerning the perception of VVs

Two types of perceptual tests have shown interesting characteristics in the segmentation of vowel sequences in Japanese and English. In /ai/ tail-catching tests (Gore 2006, 5.3), preschoolers showed mixed single-diphthong and dual-monophthong perception of the /ai/ sequence, with a tendency to perceive /(C)ai/ as a single (indivisible) unit. In a different type of test focusing on cross-language perception at a later stage of development, L2 English /au/ was often segmented moraically, whereas /ai/ tended to be perceived as a single unit (Gore 2006). These various tests, taken in combination, suggest that the accepted (phonologically mature) process of moraic segmentation is noticeably less dominant or pervasive in the specific case of /ai/ (i.e. this sequence can be perceived as an indivisible unit). We have speculated (Gore 2006) that this may be due to factors on two levels, morphophonological and acoustic: (1) the existence of /ai/ in Japanese as a common (SJ) morphophonological unit and (2) a universal bias in favour of phonetic /ai/ in any language as an archetypal and stable (prominent followed by non-prominent) combination of sounds.

3. Proposal concerning theoretical approach

We now propose a theoretical framework for viewing the above phenomena based on the principle of constraint-ranking. Traditionally, in terms of the Principle and Parameters Theory (Chomsky 1981; Blevins, 1995), mora-timed languages had (-Complex Onset) and (-Complex Coda), syllable-timed languages had (+Complex Onset and +Coda), and stress-timed languages had (+Complex Onset and +Complex Coda). Optimality Theory (OT) (Prince and Smolensky 1996, 2004) shows this syllable complexity by the relative ranking of faithfulness constraints such as ONS or COD- (also known as Onset and NoCoda). According to OT, human languages share a common set of universal constraints on well-formedness. Interlinguistic differences are considered to arise from the permutations of constraint-ranking. Such differences are seen not just between languages but also between different levels of phonological development within the same language (Allen and Hawkins 1978; Jusczyk 1997; Jusczyk, Smolensky and Allocco 2002). Our view assumes that the rhythm of the mother tongue helps the child to establish the ranking of phonological constraints, and that both the perception of rhythm and the ranking of constraints may undergo a slight alteration as the child develops phonologically.

Taking our results together (Gore 2006), we may say Japanese preschoolers' ordering of two constraints on syllable structure (as proposed in OT) changes subtly but significantly at a certain stage of development. The two constraints concerned here are the basic syllable structure constraints, ONS (a syllable must have an onset) and -COD (a syllable cannot have a coda) (Prince and Smolensky 2004 p.106).

Both of these constraints are violated in certain contexts, for instance in English *an*, and Japanese *aN*, both of which violate both constraints (there is no onset but there is a coda). (For the sake of the following argument, we include Japanese N and V2 in the category of Coda.) In both Japanese and English, in a non-initial syllable in a word, the Onset constraint is rarely violated: words such as English *re.act* are rare. English VVs are more likely to be diphthongs; but Japanese mora/kana segmentation allows Onset violation to occur frequently, and VV to become codified as V.V without becoming diphthongized. We have suggested that, for Japanese children, V.V segmentation is conceptually reinforced by learning kana, and that this conception predominates after a certain age (Gore 2005, 2006).

Results in Gore (2006) show (in agreement with those of Hatano and Inagaki 1992) that preschoolers' "syllables" have both codas and onsets, i.e. in OT terms -COD is ranked low and ONS is ranked high. At a later stage, which seems to coincide with the learning of kana, VV and VN tend to be segmented moraically as V.V and V.N. Adult moras (in standard, central dialects) normally have no coda and may also have zero onset, i.e. -COD is ranked high and ONS is ranked low. As young children grow up, they learn that the old segmentation is "wrong" or "childish." The new, kana-consistent version is the generally accepted, mature segmentation strategy.

Thus, in Japanese, we can interpret the above as showing that the young child's relative rankings of -COD and ONS reverse with increasing age and/or kana acquisition. In short, the preschool analysis allows a diphthongal post-nuclear ending that violates -COD; whereas the moraic CV analysis ranks -COD higher, moving the final V into a separate unit which violates ONS. Thus, for the mature child, and throughout mature phonology, -COD >> ONS.

However, results (Gore 2006, 3.3 and 3.4) suggest this crossover occurs in a piecemeal fashion as some VVs seem to be moraically segmented at an earlier stage than others; and that traces or echoes of undivided perception can also surface later in rhythmical or repetitive forms such as songs and jingles.

4. Proposal concerning stages of phonological development

The table below **[1]** brings together several related arguments and presents them in terms of perceptual development. We propose three stages of phonological development, the first two relate to all possible VVs. The third includes a subsidiary mode that allows morpheme-internal Vv (with a less prominent second vowel) as an undivided rhythmic unit.

[1] Developmental stages in VV perception

Stage 1 Mixed perception (till approximately age 6)

(C)V.V / **(C)V.v** and **(C)Vv** / **(C)VV**

Stage 2 Moraic division of all sequences

-COD >> ONS **(C)V.V / (C)V.v**

Stage 3 Acquisition of SJ units reactivates (C)Vv traces, and legitimizes a conceptually undivided form. This operates rhythmically as a subsidiary mode.

Dominant moraic mode *(all sequences)*	Subsidiary syllabic mode
-COD >> ONS ***(C)V.V [(C)V.v]***	*ONS >> -COD* ***(C)Vv***

Stage 1 shows mixed (divided and undivided) perception. By Stage 2, moraic (divided) perception dominates the perception of all sequences, conceptually reinforced by kana orthography; in this stage the second vowel is perceived separately as if it were a CV even though it has no Onset (NoCoda ranks higher than Onset). Stage 3 finalizes the dominant moraic mode, but allows a subsidiary syllabic mode for SJ-type morpheme-internal sequences (mostly /ai/ and /ei/) in songs, jingles, poems, some long names and other rhythmical or repetitive forms (Gore 2006, 3.9; 3.11). In this syllabic mode, the morphophonological (C)Vv unit functions rhythmically as a unitary syllable (Onset ranks high, allowing NoCoda to be violated), and the second vowel is not prominent, regardless of whether it would normally (in the moraic mode) be accented or not.

A period (.) between two vowels indicates a perceived moraic division (not necessarily morphemic or phonetic, but always moraic), thus the absence of a period indicates (infant) ignorance or (adult) suspension of moraic segmentation procedures.

The table below **[2]** gives examples of Japanese words or phrases in the developmental stages.

[2] Examples of words in three developmental stages

Stage 1	Syllabically divided examples		
	[aŋ]	[mo: nai]	[atta]
	(more food!)	(all gone!)	(there it is!)
Stage 2	Moraically divided examples		
	[a.N]	[mo.o.na.i]	[a.tt.ta]
Stage 3	Dominant moraic mode	Subsidiary syllabic mode	
[a.N]	[mo.o.na.i] [a.tt.ta]	[aN] [mo: nai]	[at.ta]

5. Proposal concerning reclassification of VVs in adult Japanese phonology

In (mature) Japanese phonology, since there is a potential differential or step between vowels at the mora boundary within the bimora, we can show four theoretical types of vowel sequence succinctly by using upper and lower case V/v as follows: (1) V^1V^1; (2) V^1V^2; (3) V^1v^1; (4) V^1v^2. The HIGH-low shape indicates the step (which spectrographically may be visible in several ways: pitch (F0), F1, F2 or intensity);

what is important is that there is a rapid stepwise movement at or encompassing the halfway point. The first two VV types will have approximately level perceived sonority. V^1V^1 can be found across a morpheme boundary ($V^1\#V^1$); e.g., /ja#aku/ "evil," and occasionally morpheme-internally, e.g. /toori/ "road." V^1V^2 is found equally within and across ($V^1\#V^2$) morphemes; e.g., /kao/ "face" and /o#agari/ "come in." V^1v^1 represents many instances of the so-called "long vowel," which usually occurs within a morpheme and with weakening of the second element (V^1v^1); this can be an allophone of V^1v^2 (/ee/=/ei/; /oo/=/ou/). V^1v^2 is the quasi-syllabic bimora that some call a diphthong, and that Gore (2006) has discussed. However, as has been reported (Gore 2006, Vance 1987, Kubozono 2001), the identical sound may occur both within (V^1v^2) and across ($V^1\#v^2$) morphemes. Also, the identical morpheme or sequence of morphemes may be realized both as VV and as Vv, depending on speech rate, context, prosodic effects, and dialect.

Since morpheme boundaries normally have no phonetic correlate (Lehiste 1972) and are independent of perceived divisions of the VV (Gore 2006), there are 2 independent variables, and 4 possible contingencies here. And since there is no single coherent definition of diphthong in Japanese (Gore 2006), it is helpful to classify Japanese bimoraic VVs not as diphthong/non-diphthong, but according to whether there is a morpheme boundary (#), a high-low pitch accent (Vv) (in Tokyo Japanese), and/or a perceived division (.) between the two moras. Theoretically there are eight possible combinations of these:

1)	+ morpheme boundary (#)	+ perceived division (.)	+ accent (Vv)
2)	+ morpheme boundary (#)	+ perceived division (.)	- accent
3)	+ morpheme boundary (#)	- perceived division	+ accent (Vv)
4)	+ morpheme boundary (#)	- perceived division	- accent
5)	- morpheme boundary	+ perceived division (.)	+ accent (Vv)
6)	- morpheme boundary	+ perceived division (.)	- accent
7)	- morpheme boundary	- perceived division	+ accent (Vv)
8)	- morpheme boundary	- perceived division	- accent

In each morphemic and perceptual category the same set of phonetic contrasts is available (VV and Vv), but VV-internal breaks at slow speed can only occur at the morpheme boundary (+morpheme boundary). A suggested exemplification of the above categorization is given in **[3]** below for non-identical VVs.

[3] Morpheme boundary (#) and perceived division (.) in non-identical VVs

+ morpheme boundary; + perceived division (+accent Vv *left;* -accent VV *right*)

V#.v haka#.iʃi; ha#.iʃa V#.V mana#.ita

+morpheme boundary; -perceived division ("single-note" and "single-mora")

V#v na#i; nasa#i; kudasa#i; ita#i; V#V *zero*

-morpheme boundary; +perceived division

V.v *zero* V.V genda.ikko*

-morpheme boundary; -perceived division

Vv haiʃa (“loser”); gendaizin; gendaikko VV gendai; daizi

VVs sung on a single note are “**- perceived division.**” See discussion of Vance (1987) below.

*/gendaikko/ appears in two accentual categories: in careful Tokyo pronunciation the prominence of [i] when followed by silence (Q) may enhance divided perception (Gore 2006, 3.7.12; 3.9.1).

This table proposes an economical and impartial framework within which to compare groups of sequences in morphemic categories without conflating first vowel prominence with morphemic internality (some of which have been proposed in the literature as diphthongs (for review, see Gore 2006, 3.7)), and to understand more clearly why they have been put forward and why categorization as “diphthong” adds little or nothing to the accepted units of the mora and the bimora, i.e. is not phonologically significant in adult, central Japanese. This is helpful since there have been many suggestions that presupposition of the syllable (as separate from the mora) in adult central Japanese is not justified (Kindaichi 1957; Han 1962, 1994; Vance 1987; Gore 2005) even in the case of /ai/, though this sequence may constitute a morphophonological entity (Gore 2006).

One group, [minus morpheme boundary; minus perceived division /ai/ Vv], corresponds to the diphthong proposed by Kubozono. We propose that this type is better characterized as a “quasi-syllabic morpheme-internal Vv bimora,” a subset of the morphophonological VV unit we have described as “unbreakable” (Gore 2006 p. 92).

Another group with an interesting set of properties is the one defined as [plus morpheme boundary, minus perceived division, V#v]; this group includes diphthong candidates proposed by Vance (1987) that are not accepted by Kubozono (2001b p.61, 2003) owing to their dimorphemicity. This group is interesting because, even though there is a morpheme boundary between the vowels, if the sequence is Vv, then there tends to be no perceived division (at least in Standard Japanese, according to the above researchers). This group includes the /ai/ “syllables” that are often sung on a single note despite being dimorphemic (and digraphic), e.g. kudasa#i, and nasa#i, and many children’s perceptions of /a#i/ adjectives, e.g. ita#i (“ouch”) and negative –na#i. As we have noted (Gore 2008 3.9.3) these words have /a#i/ in the final position, tend to appear phrase-finally, and have /ee/ and /e/ as allophones. These characteristics seem largely independent of dialect. This sequence, /a#i/, be classed as a “quasi-syllabic

word-internal dimorphemic Vv bimora." This /ai/ is especially likely to be viewed as a unit: factors enhancing this perception are, we suggest, (1) inherent a/i sonority differential (2) accentual prominence of first vowel (3) final position in word or phrase and (4) existence of single vowel allophone (in adjectives); factors detracting from this perception are (5) the morpheme boundary and (6) mora timing. Vance's view of the Japanese "diphthong" implies that when (1), (2) and (3) coincide, they win over (5) and (6), and the sequence is perceived as a "diphthong." Others (Kokugogaku Jiten 1978 p.719) have defined "diphthong" mainly in terms of (4) but do not treat the case of /ai/ (Gore 2006, 3.7.6).

6. Conclusion

The above tabulation, by presenting a unified view of morpheme (boundary), mora (division), perception of division, and vowel prominence, allows a theoretically neutral reference point for the definitions found in the literature. Some researchers say the term "diphthong" can apply to some /ai/ bimoras accented on the first vowel (in Tokyo Japanese citation forms) (Gore 2006, 3.7); there agreement ends and confusion begins. The morpheme-boundary group (Vance) are YJ, phrase-final, sometimes have single-vowel allophones, and can be sung on single notes. In contrast, the morpheme-internal group (Kubozono) has slightly more claim to linguistic relevance since they are a morphological unit, but in central Japanese this type of unit also contains VVs with non-diphthongal accent patterns. This type of unit occurs mostly in SJ morphemes (or in similar "SJ-type" sequences) and can be a single rhythmic unit in young children's perception and in dialects that are non-accent or "syllabic" (Gore 2006, 3.7.2). However, its linguistic significance in dialects is beyond the scope of the present paper.

References

Allen, G. & Hawkins, S. (1978). The development of phonological rhythm. In Bell, A. & Hooper, J.B. (eds.) *Syllables and Segments*. 173-185. Amsterdam: North-Holland.

Gore, M. (2005a). Language rhythm and syllable type. *Journal of the English Phonetic Society of Japan. 7*. pp.61-71.

Gore, M. (2005b). Rhythmic perception and the VV. *Journal of the English Phonetic Society of Japan. 8*. pp.77-85.

Gore, M. (2005c). Vowel sequences in Japanese; diphthongs in English. *Proceedings of the International Congress of Phoneticians of English. pp.154-163.*

Gore, M. (2006b). *English and Japanese Diphthongs and Vowel Sequences*. PhD thesis. Department of Applied Linguistics, University of Reading.

Gore, M. (2006c). Segmentation of English diphthongs by Japanese schoolchildren. *Journal of the English Phonetic Society of Japan.*

Han, M. S. (1962). The feature of duration in Japanese. *Onsei no kenkyuu* 10: 65-80.

Han, M. S. (1994). Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96 (1) 73-82.

Haraguchi, S. (1996). Syllable, mora and accent. In Otake, T. & Cutler, A. (eds.), Speech Research 12. *Phonological Structure and Language Processing*. Cross-linguistic Studies. Mouton de Gruyter.

Hatano, G. & Inagaki, K. (1992). *Yoozi no on'in ishiki – sono 2: shiritori wo tooshite* [Phonological awareness in Japanese

children 2: tail-catching]. 56th Annual Convention, Japanese Association of Psychology.

Inagaki, K. & Hatano, G. (1992). *Yoozi no on'in ishiki – sono 1: go no bunkai wo tooshite*. [Phonological awareness in Japanese children 1: word segmentation]. 56th Annual Convention, Japanese Association of Psychology.

Inagaki, K., Hatano, G. & Otake, T. (2000). The effect of kana literacy acquisition on the speech segmentation unit used by Japanese young children. *Journal of Experimental Child Psychology* 75: 70–91.

Jusczyk, P. (1997). *The Discovery of Spoken Language*. Cambridge MA: MIT Press.

Jusczyk, P., Smolensky, P., and Allocco, T. (2002). How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition* 10: 31-73.

Kindaichi, H. (1957). *Nippongo* [The Japanese Language]. Tokyo: Iwanami Shoten.

Kokugogaku Jiten (*Dictionary of Japanese Linguistics*). (1978). Kokugogakkai. Tokyodo Shuppan.

Kubozono, H. (1999). Mora and syllable. In Tsujimura, N., (ed.) *The Handbook of Japanese Linguistics*. 31-61. Oxford: Blackwell.

Kubozono, H. (2001). *Mora and Syllable*. Kobe University.

Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America* 51 (6): 1972-2024.

McCawley, J.D. (1968) *The phonological component of a grammar of Japanese*. Mouton, The Hague.

Ramus, N., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-293.

Vance, J.V. (1987). *An Introduction to Japanese Phonology*. Albany NY. State University of New York Press.