

A Study of the Correlation Between Durational Mistakes and Pitch Pattern in English Speaker's Production of Japanese

Chiharu Tsurutani
Griffith University

<Introduction>

When Japanese language teachers hear students pronouncing Japanese poorly, they tend to correct the mistakes by transcribing them and pointing out the differences as follows:

daigak <u>ku</u>	->	daigaku
subarash <u>i</u>	->	subarash <u>ii</u>

However, corrections such as these do not produce a long term effect unless the fundamental problem is solved, which is the acquisition of Japanese rhythm. The rhythm of a language is ruled by its prosodic features, such as accent pattern, segmentation units and intonation. What is happening in the above example is that the speaker is using English prosody on a Japanese sentence. English listeners use different prosodic parsing strategies, such as syllabic segmentation and feet, and put the stress accent on certain vowels, none of which are appropriate to Japanese prosody. In this paper I will focus on the durational mistakes of the geminate consonant and consider the cause of these mistakes in relation to pitch pattern by examining students' speech data. Before examining the actual data, I will discuss the differences in English and Japanese prosody in terms of their prosodic units and accent pattern. At this stage the discussion will focus on the problems at a word level.

1. Prosodic unit - syllable structure in English and Japanese

The following diagram shows how English speakers respond to the Japanese word "hon".

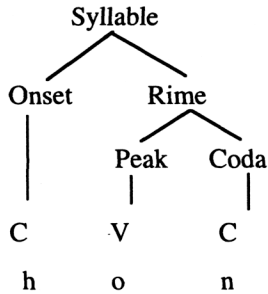


Figure 1: Syllable structure of English as exemplified with the Japanese word /hon/

The major boundary within a syllable comes between onset and rime, and the peak is more closely related to coda than to onset. The basic concept of syllable can be well explained by the Sonority Sequencing Principle (SSP). Sonority is a relative loudness to other sounds with the same length, stress and pitch. (Ladefoged:1982) The scale of the sonority is marked highest in vowels and lowest in obstruents, as follows:

vowels

glides (semi vowels,[y][w])

liquids (r,l)

nasals (m,n,n)

obstruents (fricatives and stops)

Most primitive syllable inventories are CV, VC, V, CVC (C= consonant. V= vowel).

A vowel is usually regarded as a peak of sonority in a syllable, that is the core of a syllable. In most languages of the world vowels are the only possible syllabic nuclear (Trubetzkoy:1969). The following are possible patterns of syllable structure in English:

Table 1: Syllable structure

Word	Structure	Onset	Rime	
			Nucleus	Coda
I	VV		aɪ	
it	VC		ɪ	t
go	CV	g	o	
pit	CVC	p	ɪ	t
grasp	CCVCC	gr	æ	sp
strands	CCCVCC	str	æ	nz

SSP further explains the possible sequence of consonant clusters in Onset and Coda. For example, in the onset and coda of the word, *grasp*, [g] and [p] are less sonorant than [r] and [s] respectively. This means that the sequence of consonants in a syllable is also ruled by a sonority hierarchy, and explains why the consonant clusters, such as "nt" or "ls", in onset position are not possible in English. However, some problems, such as "strands" remain. In the sequence of "str", [s] is higher in the sonority hierarchy than [t] and it violates SSP. A similar case occurs in the plural suffix [s] in "sticks". Giegerich (1992) treats these as appendices and distinguishes them from core syllables. Therefore, if we exclude the appendices as exceptions, a core syllable is associated with a single sonority peak. This sonority cycle also correlates with articulatory action.

For Japanese speakers the syllable structure as shown in Figure 1 is hard to accept for a number of reasons. First, the internal boundary of 'hon' comes between the peak "o" and coda "n". In careful speech "hon" is pronounced as two units "ho" and "n". This concept is largely based on the fact that in the Japanese orthography Kana, the corresponding patterns are limited to CV, V, CCV (the second consonant is limited to

the glide [j]), N (nasal coda), and Q (geminate consonant). The CV type is the simplest structure of a syllable and is called a mora. Therefore syllabification in Japanese has been widely identified with moraic segmentation by native speakers and some scholars. Kindaichi was one of the scholars who insisted that there is no other syllable structure than the mora in Japanese.

Some empirical research which supports the moraic segmentation is available. Kubozono(1989) found a unique pattern of substitution errors which occurred by splitting the syllable into CV/V or CV/N. They involve the split of complex (ie. long or diphthongal) vowels as well as of geminate consonants:

- a) zju.u -go **pa.a** - se.n - to (15%)
 -> zju.u - go **pa.n** - se.n - to
- b) ko - **ma.t** - te i - ru (be in trouble)
 - . ko - **ma.n** - te i - ru

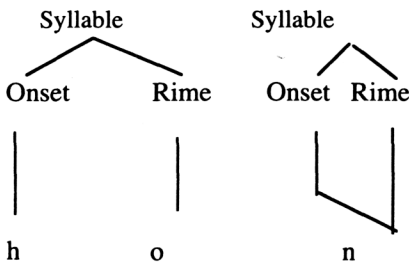
Otake(1993) conducted a target detection test to investigate the most natural segmentation unit in Japanese. He found that CV targets in CVNVCV words were detected as easily as in CVCVCV words, suggesting that CV (mora) is the segmentation unit of speech in Japanese. On the other hand, the same type of study conducted with French listeners who would use a syllabic segmentation showed the predominance of CV target detection in CVCVCV and CVN in CVNVCV. However, there are some researchers who acknowledged that "hon" is commonly pronounced as one syllable and recognized the existence of both mora and syllable structure in Japanese. Arisaka (1940) called "hon" a phonetic syllable, and the moraic segmentation "ho-n", present in careful speech, they called a phonological syllable. The empirical researches mentioned above also have the possibility that the result can be affected by the orthography in native speaker's perception and the type of mora. Another

experiment conducted by Otake in 1992, using CVV syllables, showed a different result. In this experiment 40 Japanese natives were asked to indicate the most natural segmentation point in 8 words which included CVV syllable. 70% of the subjects preferred CVV to CV. Otake himself regards mora as a sub-syllabic unit and admits that morae and syllables coexist.

Nowadays the concept of explaining Japanese prosody using new phonological units is gaining ground in Japanese phonology. One of these theories is the bi-moraic theory. It states that Japanese has a different foot structure from English, and consists of two morae which work as a basic phonological unit.

Whether the bi-moraic foot should be counted from right to left or from left to right in odd-numbered moraic words is still the subject of much discussion, but is beyond the scope of this paper.

Yoshida (1990) claimed that "hon" is analysed as 2 syllables from the viewpoint of Government phonology. This theory denied the structure "hon" as one syllable based on Coda Licensing Principle, which says post-nuclear rhymal positions (eg:n) must be licensed by the following onset. Therefore 'n' has to be the onset of the next syllable followed by an empty nuclei. Then he proposes that Domain-final empty nuclei are not licensed in Japanese. As a result the syllable structure of "hon" is presented as follows:



In this way the need for morae disappears.

However, much work still remains to be done before the concept of syllable is accepted universally.

Here, I will present the moraic segmentation of the Japanese word "hon":

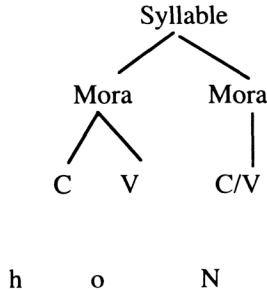


Figure 2 : Moraic segmentation of "hon"

The basic structure of a mora is (C)V. In all, over 60% of all syllables are of the CV type. CCV, N(moraic nasal), Q(moraic obstruent) are called special morae. Since every mora is regarded as having equal duration, a problem for English speakers arises. In the next section the actual prosodic organization in each language will be discussed.

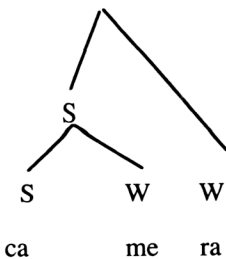
2. Prosodic organization in English and Japanese

2-1. Accent pattern in English and Japanese

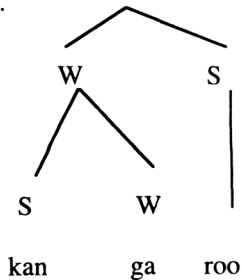
English has a stress accent, and the English stress is a marker of the rhythmic impulse. Stressed syllables are assumed to have higher pitch, greater intensity and longer duration, (Nagano-Madsen:1992). Of the three elements, loudness (intensity) used to be regarded as the major component (Jones:1918,1969). However, a number of experimental studies have indicated that the duration ratio (Fry:1955) or pitch

(Boliger:1958) is a more effective cue than intensity for English listeners. If you observe the pitch pattern and duration of the stressed syllable, it is found that the stressed syllable is not much louder than the rest, but does show a noticeable pitch movement and longer duration. The syllables on which the stress falls regularly have a branching rhyme, VV, VC, VVC and VVCC. A syllable which has a branching rhyme is called a heavy syllable, as opposed to light syllables which do not have a branching rhyme. When a word contains more than one heavy syllable, the stress placement is ruled out by the notion of Metrical Phonology. In this notion the prominence relations are shown by binary nodes, S (strong) and W (weak) (see below). Since strong and weak are relational definitions to the sister node, words consisting of more than two syllables must be broken up into binary structure. The following are some examples.

a.



b.



Stress falls on "ca" in "camera" and "roo" in "kangaroo".

The foot assignment rule in nouns will be discussed in detail later in this paper.

In Japanese both heavy and light syllables can be stressed, although in the stressed heavy syllable only the first mora will be stressed.

The accentuation of Japanese words has a clear difference from English. In Japanese, accent is not located on a particular syllable, nor pronounced more strongly than other

parts of a word. Instead, a change of pitch, which occurs on the border of a mora, is regarded as the Japanese accent unit. However, the strength of the pronunciation has to be kept even throughout a word, unless pragmatic circumstances require an emphasis on the word, and then only the pitch changes. The following examples are the basic pitch patterns of Japanese in 4 mora- words:

1) Flat	ko-u-ko-u	LLLL (HHHH)
2) Head high	bi-n-bo-u	HLLL
3) Middle high	a-sa-ga-o	LHLL
	ka-ra-ka-sa	LHHL
4) Tail high	o-to-u-t-o	LHHH

*H = high pitch mora, L = low pitch mora

The pitch rises only once in a word, and some researchers (Vance 1987) conventionally call the last high pitch mora the accented mora. Therefore the pitch pattern 1) is called an accentless word. There is no regulation in the assignment of high pitch, except in the second mora of a long syllable, such as kou or kan. The pitch change occurs quite clearly on the mora boundary.

2-2. Rhythmic organization

Pike(1945) categorized rhythms into two groups, stress-timed and syllable-timed. English is said to be a stress-timed language, with about the same duration between one primary stressed syllable and the next. On the other hand, Japanese belongs to the syllable-timed group, and is specifically called a mora-timed language. Each mora is supposed to have equal duration. This consistency of the equal duration of isocronic units is still very controversial. However, careful speech at the word level reflects this

tendency relatively well. That is, Japanese words of the structure CV-V-CV-V and CV-CV-CV-CV have roughly the same length. It is assumed that English speakers ignore the temporal unit, ie mora, and capture these words as 2-syllable and 4-syllable word respectively.

The plausible mistakes in English native speakers of Japanese will be assigning the longer duration as well as high pitch on the stressed mora. As a result, the Japanese word "oosaka" = accentless word (LHHH) can be pronounced as "osaaka" (LHL) in exaggeration. Further discussion has to be done based on the actual examples of mistakes.

3. Background of the Experiment

In a Japanese second language classroom, the following mistakes are often heard:

Correct word	Mistake
a) <u>mot<u>te</u>iru</u>	<u>moteiru</u> (have)
b) <u>daigaku</u>	<u>daigaku</u> (university)
c) <u>hito</u>	<u>hitto</u> (person)
d) <u>nakade</u>	<u>nakkade</u> (in)
e) <u>iss<u>hu</u>ukan</u>	<u>ishukan</u> (one week)
f) <u>itakotoga</u>	<u>itakotoga</u> (have been)
g) amega <u>futtanoni</u>	amegafutanoni (although it rained)

h) nihongoga hetanihongogaheta (not good at
Japanese)

This list was compiled by transcribing mistakes from a tape recorded for an oral test by first year students at Griffith University. The students have all studied Japanese for 4~5 years at high school. Observing the list, two points may be noticed:

- 1) /Q/ is inserted, when the pitch changes from High to Low.(b,c,d,h)
- 2) /Q/ is deleted (not heard properly) in flat pitch patterns (a,e,f,g)

Let me analyse the possible causes for these types of mistake.

- 1) Insertion of /Q/ when pitch changes from High to Low.

The learners seem to put a stress on the high pitched syllable. By putting a stress, the learners tend to lengthen the vowel but they are aware of the isocrony of mora and stop lengthening the vowel by putting glottal stop. An alternate explanatiuon might be the case that native listeners hear the extra length on a stressed vowel and interpret it as /Q/. It is observed that English speakers are not good at flat pitch patterns and tend to put stress on the penultimate mora (=syllable),ie, oosaka, watashi (I), sukiyaki, sake, karaoke, karate, yasumi (holiday).

- 2) Deletion of /Q/ in flat pitch patterns

In examples a) e) and g), the pitch change of the Japanese word occurs in the middle of a syllable again. The learners seem to have difficulty in producing /Q/ in flat pitch pattern.

Japanese native speakers hear incorrect insertion or deletion of /Q/ in foreigners' pronunciation in the following phonetic environment:

Correct form		Mistake
CV-CV(HH)	->	CV-Q-CV(HHL)
CV-Q-CV(HHH)	->	CV-CV(HH)

Two factors are listed as a cause of this type of mistake.

i) Speaker's misproduction

Since the flat accent pattern is not familiar for English speakers, they place HL pitch movement and as a result of this accentuation, /Q/ is inserted after the vowel of high pitch.

ii) Listener's misperception

When a lax vowel is heard before a voiceless consonant in the pitch pattern HL of CVCV sequence, it is interpreted as CV-Q-CV. A similar phenomenon can be seen in loan word formation in Japanese. When an English word is taken into Japanese, /Q/ is inserted, which does not exist in the original word. An exact rule which explains all loan word formation is still to be formulated. However, the following factors which render /Q/ insertion to English loan words are worth noting.

It occurs	after	1) lax vowel
		2) stress
	before	a voiceless consonant

On the contrary English listeners insert /Q/ unnecessarily when they write Japanese words, which is very likely the reflection of their mispronunciation. This phenomenon is called perceptual assimilation by some researchers (Takagi & Mann:1994). The actual cause has to be further examined from both directions, perception and production.

4. Experiment

1) Aim

The durational mistakes in English speakers' pronunciation are considered to derive from incorrect pitch assignment. This experiment aims at clarifying, by the use of spectrographic analysis, whether:

1. there is a correlation between pitch and duration in mistakes made by English speakers
2. the durational mistakes are caused by speakers' misproduction or listeners' misperception

2) Hypothesis

English learners tend to change the pitch on the syllable boundary and assign high pitch on the penultimate syllable in words with flat pitch pattern. It is easier for English speakers to pronounce words which have /Q/ on the border of pitch change High to Low, but not in flat pitch patterns.

3) Method

In order to eradicate the influence of orthography target words will not be presented in Kana. Therefore they are limited to familiar words. The students see the English word written on the paper and are asked the equivalent word in Japanese. Before the recording I checked whether students knew the exact word.

4) Target words

Words which were familiar to students and which had **a flat pitch pattern** were chosen.

	No. of morae	No. of syllables	
1) da-i-ga-ku	4	3	(university)
2) to-mo-da-chi	4	4	(friend)
3) ka-i-mo-no	4	4	(shopping)
4) he-ta	2	2	(not skilful)
5) u-ta	2	2	(song)
6) hi-to	2	2	(person)
7) ki-Q-to	3	2	(for sure)
8) i-Q-ta	3	2	(have been)
kotoga aru			

5) Subjects

22 students who have studied Japanese for 4~5 years participated in the experiment. Their level is late beginners to intermediate.

5. Findings

Out of 22 samples I chose 7 samples (student A~G) which clearly contain durational mistakes associated with assignment of the wrong pitch. I confirmed with 10 native Japanese teachers that they also perceived the mistakes, and then, using a spectrogram I analysed the data with which the majority agreed. In the table, number indicates how many people agreed with its incorrectness and the bold letter means that the word was pronounced in wrong pitch pattern.

	A	B	C	D	E	F	G
daigaku	0	0	0	1	0	1	10
tomodachi	0	7	7	4	0	0	10
kaimono	0	0	0	2	0	0	7
heta	8	2	1	4	8	10	10
uta	5	2	4	1	1	10	10
hito	1	4	1	1	1	9	10
kitto	1	0	10	5	3	0	0
itta	3	2	5	5	10	8	0

Table 2: The number of mistakes which occurred in flat pitch pattern words

I used student G's pronunciation and compared with my own pronunciation. The following is the phonetic transcription of the right and wrong pronunciation.

Correct

Mistakes

1	da i ga ku	daɪ ga kku
2	tomo da t fɪ	to m ora t fɪ
3	h ɛ ta	h ɛ tta
5	uta	u tta
6	hi to	hi tto
8	itta	it a

It is observed that insertion of geminate co-occurs with pitch change from H to L.

However the correlation between durational mistakes and pitch pattern was not observed. The following is the table of the occurrence of mistakes in word "hito" and "heta"

<hito>	CD	WD	<heta>	CD	WD
CP	4	1	CP	7	1
WP	13	4	WP	8	6

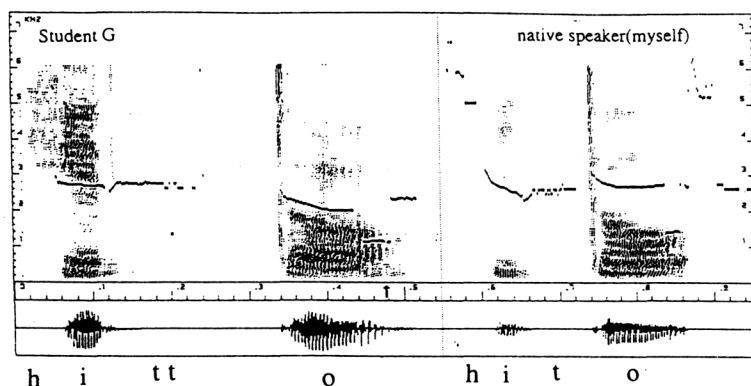
CP = Correct Pitch, WP = Wrong Pitch

CD = Correct Duration, WD = Wrong Duration

One student had a correct pitch and a wrong duration in both words. Obviously she did not have confidence to pronounce them and placed an unusual pause between all syllables. This has to be considered as a marginal case.

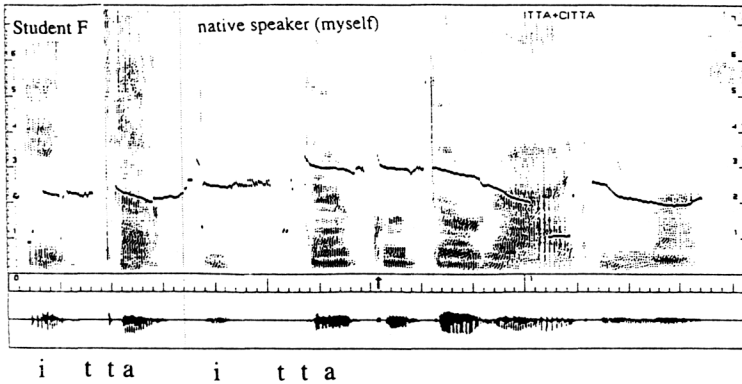
On the spectrogram it was observed that a clear pause exists after each lax vowel.

Figure 3: Spectrogram of "hito" by Student G and a native speaker (myself)



Therefore it can be stated that mistakes are due to mispronunciation. When students produce the flat pitch pattern correctly, they have difficulty keeping the appropriate length for the geminate consonant. In the word "kitto", pitch was moved from high to low and as a result it was relatively easy for speakers to produce the geminate consonant. On the contrary, in the word "itta" they did not keep the appropriate length because they pronounced the word in a flat pitch pattern.

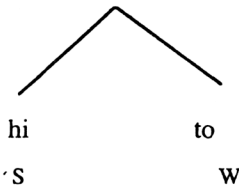
Figure 4: Spectrogram of "itta" by Student's F and a native speaker (myself)



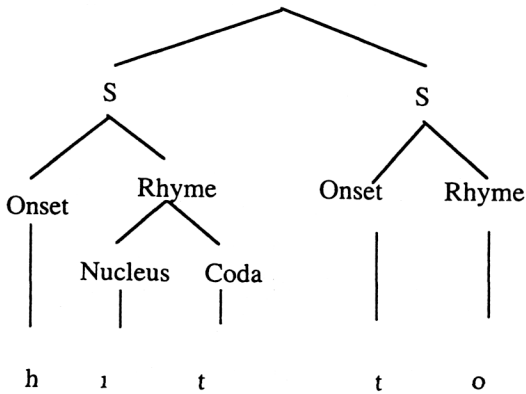
6. Discussion

Learners who are familiar with the rhythm of Japanese keep the duration correctly. Transfer of English prosody into Japanese words seems to be the reason why some learners produce durational mistakes. The mechanism of misproduction of the geminate is well explained from the phonological rule of English.

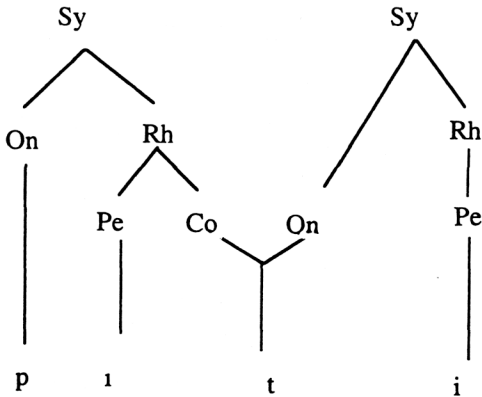
"hito" is a 2-syllable word.



When the foot contains more than one syllable, the first syllable is more prominent than the others. Besides a stressed syllable has to have a branching rhyme. Therefore the syllable structure is rearranged as follows:

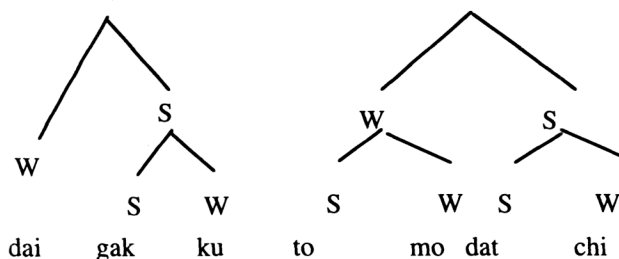


As a result, /hito/ is pronounced as /hitto/. However, there is still a question that needs to be answered: In the English word "pity"[piti], [pi] is stressed and still not followed by geminate [tt]. In this type of syllable structure, [t] is ambisyllabic as is shown below.



Therefore even if English speakers place a stress on [hi], they should be able to pronounce [hito]. [hitto] is an interesting developmental stage in second language phonology.

The stress assignment on the penultimate syllable "da" in "tomodachi" and "ga" in "daigaku", and the resultant geminate insertion are also analysed by English phonological rule as follows:



According to the Foot Assignment Rule in Nouns, a bisyllabic foot is assigned to the penultimate syllable if it is heavy.

To be able to state that all the durational mistakes of Japanese words by English speakers are derived from the application of English prosody, a further examination of mistakes in other special morae /CjV/, /R/ and /N/ will be necessary.

Incorrect duration does occur concurrently with wrong pitch assignment. However, wrong pitch assignment does not always trigger incorrect duration. Whether it is caused by individual differences or as a result of different stages of development is not known. Through interviews it was confirmed that the learners who got both duration and pitch wrong, pronounce the same way in connected speech as well. On the contrary, it was observed that in connected speech the learners who pronounced the isolated words correctly still made mistakes in pitch or duration.

It seems to be easy to correct mistakes in individual words, but acquisition of the entire rhythm of one language requires more.

Whether these durational mistakes can be remedied by teaching correct pitch assignment needs further pedagogical examination.

References

- Bolinger, D. (1958) A theory of pitch accent and sentence rhythm *Word* 14, 109-149
- Fry, D.B.(1955) Duration and Intensity as physical correlates of linguistic stress
Journal of the Acoustical Society of America 27 765-768
- Giegerich, H.J. (1992) *English Phonology* Cambridge University Press
- Jones, D. (1969) *The pronunciation of English* Cambridge
- Kenstowicz, M (1994) *Phonology in Generative Grammar* Blackwell Publisher
- Kindaichi, H. (1986) *Nihongo no tokushitsu* NHK
- Kubozono, H (1989) The mora and syllable structure in Japanese: evidence from speech errors. *Language and Speech* 32, (3) 249-278
- Kurisu, K (1994) Further evidence for bi-moraic foot in Japanese 368-370 *ICSLP*
- Ladefoged, P (1982) *A Course in Phonetics* Harcourt Brace Jovanovich Publisher
Lawrence Erlbaum Associates Inc.
- Meser, R.Armin (1990) " Patterns of Truncation" *Linguistic Inquiry*, 21/3, 478-485
- Nagano-Madsen, Y. (1992) *Mora and Prosodic Coordination : A phonetic study of Japanese, Eskimo and Yoruba* Lund University Press

Otake, T (1992) Segmentation unit in perception of Japanese: syllable and mora
Research report by Ministry of Education Group E10 38-47

Otake, T. et al. (1993) Mora or syllable? Speech segmentation in Japanese. In *Journal of Memory and Language* 32, 258-278

Pike, K.L. (1945) *The Intonation of American English* Ann Arbour: University of Michigan Press

Poser ,W.J. (1990) Evidence for foot structure in Japanese *Language* 66/1, 78-105

Sato,H (1993) Gairaigo wo zairyo to shita akusento no kentoo Gaikokujin wo taishoo to suru nihongokyouiku ni okeru onseekyooiku no hoosaku ni kansuru kenkyuu 5-18

Takagi & Mann (1994) A perceptual basis for the systematic phonological correspondences between Japanese loan words and their English source words *Journal od Phonetics* 22, 343-356

Trubetzkoy, N.S. (1969) *Principles of phonology* Berkeley:University of California Press

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