# Geminates and singletons: On 'unstretchability' of single segments*

**Katsura Aoyama**

*University of Hawai'i at Mānoa*

## Abstract

When there are quantity distinctions in a language, short or single segments seem to be more restericted in their duration compared to their long or geminate counterparts. Trubetzkoy (1969) called this 'stretchability' and 'unstretchability' in a quantity opposition, and Jakobson and Halle (1971) also called short phonemes 'unstretchable', opposed to the long and sustained phonemes.

This paper provides empirical data on the notion of unstretchability from four experiments on quantity contrasts (n vs nn) in Finnish and Japanese. In production, it was found that the contrast (n vs nn) in Finnish and Japanese. In production, it was found that the contrast between [n] and [nn] was more distinctive in Finnish than in Japanese although it was found that quantity of the nasals was clearly differentiated in both languages. The more distincive contrasts in Finnish was achieved by making [n] short. In perception, identifying a word containing a short nasal was more difficult than identifying a word containing a geminate. Moreover, children's productions differed more from adults' in their productions of singletons than in their productions of geminates.

It appears that a key for making quantity contrasts is in the duration of the short segment; when there was a difference or difficulty in the distinction between singletons and geminates, it originated in making the single consonants short rather than making the geminates consonants long.

## 1. Quantity contrasts and unstretchability

Finnish and Japanese share a characteristic of phonemic quantity distinctions both in vowels and consonants. Phonetically there are infinite variations in duration of a segment; however, phonological quantity categories are limited to two (long or short) in languages such as Finnish and Japanese (Richardson 1998:23). The categorical boundary is considered to be rather flexible; there is no definite boundary in terms of absolute duration between the short and long segments (Richardson 1998:54, see also Jakobson et al. 1952:14).

Lehiste (1970) observed a much greater scatter of values in long syllable nuclei compared to the short syllable nuclei in Czech vowels, and suggested that

**Proceedings of LP'2000, edited by B. Palek and O. Fujimura**
**ISBN 80-246-0490-6**

**ISSN 0567-8269**
**The Karolinum Press 2001: 95-113**

this would support Trubetzkoy's notion of 'stretchability' and 'unstretchability' in a quantity opposition (Trubetzkoy 1969: 175-176). Jakobson and Halle also called short phonemes 'unstretchable', opposed to the long and sustained phonemes (1971:34). Lehiste stated that, in order for the feature of quantity to have a linguistic function, it must be assumed that the timing of articulatory movements can be controlled (1970:6). Since short or single segments are more restricted in their durations compared to their long or geminate counterparts, more control over articulatory movements would be required for short segments than for long segments.

This paper provides empirical data on the notion of unstretchability from four experiments on quantity contrasts (n vs. nn) in Finnish and Japanese. Because of the unstretchability, it seems that a key for making quantity contrasts is in the duration of the short segment; when there was a difference or difficulty in the distinction between singletons and geminates, it originated in making the single consonants *short* rather than making the geminate consonants *long*. In the experiment on production, it was found that the contrast between [n] and [nn] was more distinctive in Finnish than in Japanese although it was found that the quantity of the nasals was clearly differentiated in both languages. The more distinctive contrast in Finnish was achieved by making [n] short. It was also found that, in perception, identifying a word containing a short nasal was more difficult than identifying a word containing a geminate. Moreover, children's productions differed more from adults' in their productions of singletons than in their productions of geminates.

## 2. Experiment 1: Adult production in Finnish and Japanese

In the experiment on production, it was found that the contrast between [n] and [nn] was more distinctive in Finnish than in Japanese although the quantity of the nasals was clearly differentiated in both languages. The more distinctive contrast was due to the shorter (proportional) duration of [n] in Finnish than in Japanese.

### 2.1. Method

#### SUBJECTS

Subjects of experiment 1 were 10 native Finnish speakers (1 male, 9 females, mean age = 27;8) and 10 native Japanese speakers (2 males, 8 females, mean age = 30;5). All 10 Japanese subjects and 6 Finnish subjects were living in Honolulu, Hawai'i, and the recordings were done at either the subject's apartment (2 subjects) or at the University of Hawai'i at Mānoa (14 subjects). Data from other four Finnish subjects were collected in a quiet place in San Sebastian, Spain during an international conference.



#### STIMULI AND PROCEDURE

The target words are two names, *Hana* and *Hanna*, that are phonotactically well-formed in both languages, and that differ only with respect to the length of the medial nasal. These names were used in all four experiments reported in this paper. *Hanna* is a girl's name in Finnish, and *Hana* is a girl's name in Japanese.[1] Pictures of two girls were prepared; one with brown straight hair, which was identified as *Hana*, and another with yellow curly hair which was identified as *Hanna*. The names were not written on the pictures, but only presented to the subjects orally.

The two names were first introduced to the subjects, and they were asked to repeat each name three times. The subjects' production of these two names was then elicited three times each by asking them questions such as 'Who is this girl?', or 'Who is eating an apple?' in each language. Finnish subjects were asked to do the task together in pairs; thus the elicitation was done in Finnish except minimum interruptions by the investigator in English. Japanese subjects did not do the task in pairs, but the investigator interacted with all Japanese subjects in Japanese. Japanese subjects were asked to say the names without a common ending *chan* for easier comparison with the Finnish data.

A total of 6 tokens each of *hana* and *hanna* were elicited from each subject; thus, a total of 240 tokens (6 tokens of *hana*, and 6 tokens of *hanna,* from 10 subjects in Finnish and Japanese) were collected and analyzed.

#### ANALYSIS

The speakers' productions were analyzed using the programs Signalyze[2] and Speech Analyzer.[3] Wide-band spectrograms were produced for each word, and the durations of the vowel [a] and the medial nasal [n] were measured in milliseconds. The vowel [a] was identified by its formants. The medial nasal [n] was measured from the point at which an abrupt change in the formant structure from the preceding vowel occurred until the onset of the following vowel. The second [a] was measured up to the point where the voicing ends. I decided not to measure the duration of [h]; since the target words were uttered in isolation without carrier sentences, the beginnings of some of the [h] were difficult to determine. Thus, the whole-word duration discussed here actually corresponds to the duration of [ana] and [anna]. Durations of the whole word were measured first as the entire interval between the onset of the first [a] and the end of the last [a]. Second, the duration of each segment was measured then added. Both numbers were compared to ensure the accuracy of the measurements.

## 2.2 Results

Table 1 summarizes the data from experiment 1. The quantity of the nasals was clearly differentiated in production in both languages. First, the mean absolute durations of geminate nasals (178 ms. and 172 ms. in Finnish and Japanese respectively) were more than twice as long as those of single nasals (62 ms. in Finnish and 68 ms. in Japanese). There is virtually no overlap between [n] and [nn] in terms of absolute duration in either language; the longest [n] was 97 ms. (Finnish) and 95 ms. (Japanese) and the shortest [nn] was 105 ms. (Finnish) and 116 ms. (Japanese). The relative proportion of the medial nasals also shows a clear distinction between single and geminate nasals. In Finnish, the average proportion of single nasals in [ana] was 23.8% while that of geminates in [anna] was 49.3%. In Japanese, the average proportion of single nasals in [ana] was 32.7% while that of geminates in [anna] was 50.8%.

'Unstretchability' is observed in absolute duration in this experiment; the duration of [n] ranged from 33 to 97 ms. in Finnish and from 39 to 95 ms. in Japanese. On the other hand, the duration of [nn] ranged from 105 to 258 ms. in Finnish and 116 to 220 ms. in Japanese. This means that the value of [nn] spread across approximately 150 milliseconds in Finnish, and 100 milliseconds in Japanese, while the range of the value of [n] was much more restricted.

It is found that geminate nasals were similar between the two languages, both in terms of their mean durations (178 ms. in Finnish and 172 ms. in Japanese) and in their proportions in each word (49.3% in Finnish and 50.8% in Japanese on average). A paired t-test shows that the difference between the duration of [n] in the two languages was statistically significant ($t(59) = 3.015$, $p < 0.005$), whereas the difference between the duration of [nn] between the two languages was not ($t(59) = 1.338$, $p > 0.1$). In addition, the proportion of the nasal in the whole-word in [ana] was almost 10% larger in Japanese than in Finnish (32.7% vs. 23.8%).

| | Vowel 1 duration (ms.) | Vowel 1 proportion | Nasal duration (ms.) | Nasal proportion | Vowel 2 duration (ms.) | Vowel 2 proportion | Whole word (ms.) |
|---|---|---|---|---|---|---|---|
| F: [ana] | 90 | 34.5% | 62 | 23.8% | 109 | 41.8% | 261 |
| J: [ana] | 53 | 25.5% | 68 | 32.8% | 87 | 41.8% | 268 |
| F: [anna] | 102 | 28.3% | 178 | 49.3% | 81 | 22.4% | 361 |
| J: [anna] | 77 | 22.8% | 172 | 50.8% | 90 | 26.6% | 338 |

Table 1. The comparison of the vowels and the medial nasal in [ana] and [anna]



Table 2 and Figure 1 show the frequency distributions of *(h)ana* and *(h)anna* tokens by the proportion of the nasal. This analysis by the proportion of the segments also shows that the tokens of [anna], especially in Finnish, are more widely dispersed than the tokens of [ana].

There is a very small overlap between *(h)ana* and *(h)anna* in Finnish; only one token of *(h)ana* had a nasal-to-word proportion of 38.7%, which is typically in *(h)anna* range. The two peaks are well separated from each other; 26 tokens of *(h)ana* had about a nasal-to-word proportion of 20-24%, and 19 tokens of *(h)anna* had a nasal-to-word proportion of 50-54%. The two peaks were both less sharp and closer to each other in Japanese than in Finnish. There is also a much larger overlap in Japanese; 5 tokens of *(h)ana* and 9 tokens of *(h)anna* had about 40-44% of the nasal portion.

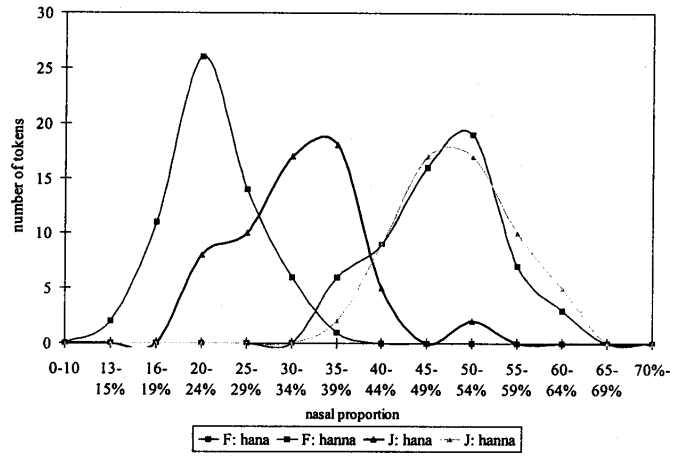


Fig 1. The distributions of tokens by the nasal proportion

Finnish

| % of the nasal | 11-15 | 16-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65- |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (h)ana | 2 | 11 | 26 | 14 | 6 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| (h)anna | 0 | 0 | 0 | 0 | 0 | 6 | 9 | 16 | 19 | 7 | 3 | 0 |

Japanese

| % of the nasal | 11-15 | 16-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65- |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (h)ana | 0 | 0 | 8 | 10 | 17 | 18 | 5 | 0 | 2 | 0 | 0 | 0 |
| (h)anna | 0 | 0 | 0 | 0 | 0 | 2 | 9 | 17 | 17 | 10 | 5 | 0 |

Table 2. The distribution of tokens by the nasal proportion

The difference seem to come from the distributions of *(h)ana*; in both languages, The peak of *(h)anna* was about the same (around 50%). There were 13 tokens of *(h)ana* whose medial nasal consisted of less than 20% of the whole word in Finnish, while there was no token of *(h)ana* which had that small a nasal proportion in Japanese. In contrast, 35 out of 60 Japanese *(h)ana* tokens had the nasal proportion around 30-39%, while only one token had such a large nasal proportion in *(h)ana* in the Finnish data.

In summary, it was found that the contrast between single and geminate nasals was clearly distinguished in both languages, although the distinction between the two categories appears to be clearer in Finnish than in Japanese. It was found that Finnish and Japanese were similar in their productions of [nn]. The mean absolute duration of geminate nasals was 178 ms. in Finnish and it



was 172 ms. in Japanese, and the difference between the two was not statistically significant ($t(59) = 1.338$, $p > 0.1$). Proportionately, [nn] in [anna] was also very similar: 49.3% in Finnish and 50.8% in Japanese on average. The difference between the two languages was in durations of [n]; the mean duration of single nasals was significantly shorter in Finnish (62 ms.) than in Japanese (68 ms.) ($t(59) = 3.015$, $p < 0.005$). The proportion of [n] in [ana] was also much larger in Japanese (32.8% in Japanese vs. 23.8% in Finnish on average). Therefore, the clearer distinction between single and geminate nasals in Finnish as compared to Japanese was due to the more strict 'unstretchability' of [n] in Finnish.

## 3. Experiment 2: Adult perception (Japanese)[4]

Experiment 2 was designed to examine the perceptual boundary between the two categories in Japanese. It was found that incremental changes in duration were perceived categorically in Japanese. It appears that identifying a word containing a short nasal is more difficult than identifying a word containing a geminate.

### 3.1. Method

**SUBJECTS**

A total of 25 native speakers of Japanese (6 males, 19 females, mean age = 27;6) participated in experiment 2. The data were mainly collected in Urbana-Champaign, Illinois, and Honolulu, Hawai'i.

**STIMULI AND PROCEDURE**

The words used were the same as in experiment 1 (*hana* and *hanna*). In order to determine the categorical boundary between single and geminate nasals, ten incrementally different audio-stimuli were prepared. The original stimulus was taken from the speech of a Japanese female subject producing the word *hanna* in experiment 1.[5]

| Stimulus | *hanna1* | 2 | *3* | 4 | *5* | *6* | *7* | *8* | *9* | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nasal (ms.) | 59 | 71 | 83 | 95 | 107 | 119 | 131 | 143 | 155 | 167 |
| Whole (ms.) | 242 | 254 | 266 | 278 | 290 | 302 | 314 | 326 | 338 | 350 |
| % nasal | 24.4 | 28.0 | 31.2 | 34.2 | 36.9 | 39.4 | 41.7 | 43.9 | 45.9 | 47.7 |

Table 3. The stimulus continuum in experiment 2

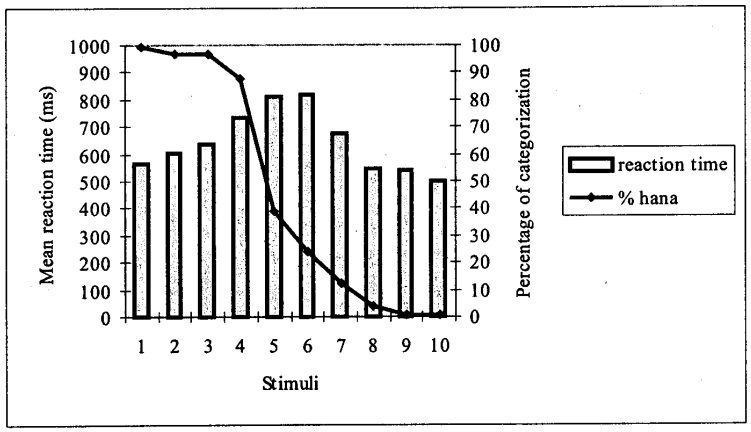


Fig 2. Mean reaction time and categorization

The original stimulus was digitized using SoundEdit 16,[6] using a sampling frequency of 44.1 kHz and a 16-bit quantization. The original *hanna* was labeled as stimulus 10 (henceforth *hanna10*), and nine other stimuli were prepared by incrementally deleting the nasal portion. To minimize the unnaturalness of artificially modified stimuli, stimuli 1 to 9 were all made from the same original. In other words, each stimulus was modified only once; *hanna9* was prepared by deleting 12 milliseconds of the nasal portion from *hanna10*, *hanna8* was prepared by deleting 24 milliseconds from *hanna10*, and so on. The total duration of stimulus *hanna10* was 500 milliseconds; approximately 70 milliseconds were left before and after the target word so that it did not start and end too abruptly. It was expected that the impression of the perceived stimulus would shift from *hana* (*hanna1*) to *hanna* (the original, *hanna10*) (see Table 3).



The prepared audio-stimuli were presented to the subjects using the program Psyscope (Cohen et al. 1993). The subjects were asked to work individually with a Macintosh G3 laptop computer with a headset in a quiet place. Subjects were shown pictures of *Hana* and *Hanna* and told the names of the girls in the pictures before the experiment. They were instructed, in Japanese, to press the appropriate buttons as they heard the names. The buttons were marked with colored stickers (a red one for *Hana*, a blue one for *Hanna*, names were not written on the stickers). A short instruction for the experiment was then presented in English on the computer screen in the beginning and ten stimuli were presented randomly. This procedure was repeated ten times; a total of 100 stimuli were presented to each subject.

There was a 1.5-second interval after the button was pressed for one stimulus before the onset of the next one. It took approximately 5 minutes for each subject to complete the task. Both the button pressed and the reaction time in milliseconds were recorded. Reaction time was measured from the offset of the final vowel of the stimulus until a key was pressed.

ANALYSIS

In experiment 2, the 10 stimuli were presented 10 times each, and they were identified by 25 subjects. The first block of 10 responses was excluded from the analysis as a practice. Thus, a total of 2250 identifications (90 identifications by 25 subjects) was analyzed. Outlying responses (those faster than 50 ms. or slower than 3000 ms.) were treated as errors (16 responses, 0.7% of the data). Reaction times were also compared to see whether it took longer for the subjects to identify the stimuli close to the categorical boundary.

## 3.2. Results

Table 4 and Figure 2 show the mean percentage of categorization as *hana* and the mean reaction time for each stimulus in experiment 2. Incremental changes in duration were perceived categorically by Japanese speakers; stimuli *hanna1* to *hanna3* were almost always categorized as *hana*, and stimuli *hanna8* to *hanna10* were almost always categorized as *hanna*. *Hanna5* seems to be the most ambiguous stimulus for the Japanese speakers; it was identified as *hana* 39.1% of the time and as *hanna* 60.9% of the time.

| Stimulus | *hanna1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| % *hana* | 99.1% | 96.9% | 96.9% | 87.8% | 39.1% | 24.3% | 12.2% | 3.6% | 0.9% | 0.9% |
| Reaction time (ms.) | 567 | 605 | 635 | 731 | 810 | 820 | 678 | 544 | 538 | 501 |

Table 4. Percentage of identification (as *hana*) and reaction time for each stimulus

Overall, the stimuli that were identified less categorically (*hanna4* to *hanna7*) took longer than the others (ranging 678 ms. to 820 ms.). It took longest to respond to *hanna5* and *hanna6* (810 and 820 ms. on average respectively). The other stimuli generally took shorter to respond; *hanna3* averaged 635 ms., while the reaction times for *hanna1* and *hanna2* and *hanna8* to *hanna10* ranged between 501 ms. and 605 ms.

Among the stimuli that were categorized as either *hana* or *hanna* more than 95% of the time, there was a tendency for the stimuli categorized as *hana* to take longer than the stimuli categorized as *hanna*. The overall mean reaction time for *hanna1* to *hanna3* (categorized as *hana* more than 95% of the time) was 602 ms. while it was 527 ms. for *hanna8* to *hanna10*. A paired t-test on each subjects' mean reaction time for these stimuli shows that this difference was statistically significant ($t(74) = 4.375$, $p < 0.001$).

In summary, it was found that a gradual change in duration was perceived categorically by Japanese subjects, and that the stimuli close to the categorical boundary took longer for the subjects to respond to than did the other stimuli. In addition, it was also found that identifying a word containing a single nasal took longer to respond than identifying a word containing a geminate.

## 4. Experiment 3: Adult perception: A crosslinguistic study

In experiment 3, it was examined whether Finnish speakers identify Japanese single and geminate nasals as such. Considering the differences in production found in experiment 1, it was hypothesized that Finnish speakers may not always identify single and geminate nasals as intended by Japanese speakers. In order to test this hypothesis, Finnish speakers were asked to categorize the Japanese speakers' productions of single and geminate nasals.



### 4.1. Method

**SUBJECTS**

A total of 20 Finnish adults (8 males, 12 females, mean age = 24;6) participated in experiment 3. One subject was living in Honolulu, Hawai'i at the time when she participated in this study. All other 19 subjects were living in Jyväskylä, Finland.

**STIMULI AND PROCEDURE**

The same target words (*hana* and *hanna*) as in the previous experiments were used in experiment 3. The stimuli were taken from productions of *hana* and *hanna* by three female native speakers of Japanese in experiment 1. The reason for taking stimuli from three speakers was to balance out individual variations in production. The reason for using the female speakers' productions was simply because the majority of the participants were female in experiment 1. An effort was made to select speakers whose productions were relatively close to the overall average in Japanese, but attention was also paid to the conditions of the recordings.

Eighteen tokens (3 tokens each of *hana* and *hanna*, from 3 speakers) were prepared by using SoundEdit 16. When preparing the stimuli, a sampling frequency of 44.1 kHz and a 16-bit quantization were used. Approximately 70 milliseconds were left before and after the target word so that it did not start or end too abruptly. In order to assure that the stimuli were unambiguously identifiable as *hana* and *hanna* in Japanese, three native Japanese speakers were asked to listen to the prepared stimuli before the actual data collection from the native Finnish speakers.

The prepared audio-stimuli were presented to the subjects using Psyscope. Stimuli were grouped together into three blocks according to speaker, but each speaker's utterances were randomized within blocks.Then the three blocks were presented in random order (one cycle). This cycle was repeated 5 times. Thus, each stimulus was presented 5 times with a total of 90 stimuli presented to each subject.

The subjects were asked to work individually with a Macintosh G3 laptop computer with a headset in a quiet place. They were shown the pictures of *Hana* and *Hanna* and given short oral instructions in English. They were told the names of the girls in the pictures, *Hana* and *Hanna*, and to press the appropriate buttons as they hear the names. The keys were marked with colored stickers (a red one for *Hana*, a blue one for *Hanna*, the names were not written

on the stickers). A short instruction for the experiment was also given in English on the computer screen at the beginning.

There was a 1.5-second interval after a button was pressed for one stimulus and before the onset of the next stimulus. It took approximately 5 minutes for each subject to complete the task. Both the button pressed and reaction time in milliseconds were recorded. Reaction times were measured from the offset of the final vowel in the stimuli until a key was pressed.

ANALYSIS

A total of 18 stimuli (3 *hana* and 3 *hanna* from 3 subjects) were presented 5 times each, and they were identified by 20 subjects. The responses for the first block (6 responses) were excluded from the analysis as a practice. Thus, a total of 1680 identifications (84 identifications by 20 subjects) were analyzed. Outlying responses (those faster than 50 ms. or slower than 3000 ms.) were treated as errors (3 responses, 0.2% of the data).

### 4.2. Results

Table 5 shows the numbers and percentages of misidentified tokens, and the mean reaction times for the *hana* and *hanna* stimuli. In experiment 1, the difference in production between Finnish and Japanese was found to lie in the duration of [n] rather than the duration of [nn]; thus I expected that, if misidentifications occur, it would be in the identification of *hana* rather than in *hanna*. *Hana* was indeed misidentified more often than *hanna* (27 vs. 5 times), but a paired t-test on the number of each subject's misidentifications showed that this difference was not statistically significant ($t(19) = 1.448$, $p > 0.1$). Most subjects had no or very few misidentifications out of their total of 84 responses. It is apparent that in general Finnish speakers did not have difficulty in identifying Japanese speakers' productions of *hana* and *hanna*. In another word, Japanese speakers' intention of producing *hana* or *hanna* matched with Finnish speakers' perception of them.

| | Total number of identification | Misindentified | % of misidentification | Mean reaction time (ms.) |
|---|---|---|---|---|
| Target *hana* | 839 | 27 | 3.2% | 503 |
| Target *hanna* | 838 | 5 | 0.6% | 401 |

Table 5. Percentages of misidentification and the mean reaction time (ms)



However, even though the Finnish subjects correctly identified the *hana* and *hanna* stimuli, it appears that it took longer to identify a word containing a short nasal than to identify a word containing a geminate. The subjects' mean reaction time was significantly longer to identify a stimulus as *hana* (503 ms.) than to identify a stimulus as *hanna* (401 ms.) ($t(19) = 4.372$, $p < 0.001$). In addition, there were three subjects whose percentage of misidentification was higher than for most of other subjects, and for these three, percentages of misidentification were higher *only* for *hana* targets.

In summary, it was found that a word containing a short nasal took longer to identify than a word containing a geminate even when they were both correctly identified.

## 5. Evidence from acquisition

When children have a difficulty in making a contrast in quantity, or there is a difference in their production of quantity contrast between adults and children, it appears that their difficulty or the difference is likely to be originated in their productions of the short segment of the contrast. Here I report experimental data on the acquisition of the quantity contrasts in Finnish, and other studies.

### 5.1. Experiment 4. The acquisition of quantity contrasts in Finnish

#### 5.1.1. Method

SUBJECTS

A total of 24 Finnish children participated in experiment 4. There were three groups; 10 children in the 3-year-old group (mean age 3;5), 8 in the 4-year-old group (mean age 4;7), and 8 in the 5-year-old group (mean age 5;3). Production data from 10 Finnish adults in experiment 1 were used as an adult control group[7] (mean age 27;8). The experiment was administered in Jyväskylä, Finland.[8] A female native speaker of Finnish (referred to as the Experimenter) interacted with the children; thus all questions were asked in Finnish.

PROCEDURE

The same target words as in previous experiments were used: *hana* and *hanna*. This is the only pair of names used in the experiment because there was not enough time for each child to be familiarized with many unknown characters. *Hana* is an existing word for 'water faucet' in Finnish, but the children seemed to be comfortable accepting *Hana* as a name for the girl in the picture.

There were three tasks in the experiment; imitation, production and perception. The two names were first introduced to the children, and they were asked

to repeat each name three times after the Experimenter (imitation task). The children's production of these two names was then elicited three times each by asking them questions such as *Kuka tämä on?* 'Who is it?' (production task). Three non-imitative utterances each of *hana* and *hanna* were elicited from each child in the production task. The last task was a perception task. Children were asked to point to the appropriate picture when asked *Kumpi on Hana/Hanna?* 'Which one is *Hana/Hanna*?'. Different types of questions were asked randomly to prevent children from answering based on the order and patterns of the questions. Four or more questions were asked for each subject in the perception task. The perception task was given at the end of the session, because it is easier to detect children's uncertainty in the production task than in a picture-pointing task such as this one.

In addition to the investigator's observation notes, each session was audio-recorded on a DAT tape and videotaped. It took approximately 5 to 20 minutes for each child to complete the tasks.

**ANALYSIS**

Productions from the imitation and production tasks were combined, and durational measurements in milliseconds were carried out using Signalyze. For each word, the durations of each vowel, and the medial nasal were measured. The duration of the [h] was not measured for the same reason as in experiment 1. A total of 158 tokens of *(h)ana* and 169 tokens of *(h)anna* from 24 children were measured. A total of 60 adult tokens each of *(h)ana* and *(h)anna* from experiment 1 were compared with the children's productions. The perception task was analyzed in terms of the number of matching responses to the questions.

### 5.1.2. Results

**PRODUCTION**

The data from the imitation task and the production task were combined and will be reported as the production data (Table 6), as opposed to the perception data (Table 7). Adult productions were remarkably shorter than children's productions in general, due to their faster rate of speech compared to young children. The mean duration of [nn] was approximately twice as long as that of [n] in every group if we examine the contrast between the two words in each group. A paired t-test conducted on the 52 3-year-olds' tokens each of [ana] and [anna] showed that the duration of the geminate nasal was significantly longer than that of the singleton ($t(51) = 9.451$, $p < 0.001$).



| | nasal (ms.) | whole word (ms.) | Proportion of the nasal |
|---|---|---|---|
| Adults: [ana] | 62 | 261 | 23.9% |
| Adults: [anna] | 178 | 361 | 49.3% |
| 5-year-olds: [ana] | 122 | 389 | 31.4% |
| 5-year-olds: [anna] | 225 | 457 | 49.2% |
| 4-year-olds: [ana] | 100 | 365 | 27.4% |
| 4-year-olds: [anna] | 224 | 445 | 50.3% |
| 3-year-olds: [ana] | 150 | 461 | 32.5% |
| 3-year-olds: [anna] | 305 | 588 | 51.9% |

Table 6. The mean durations of the nasal and the whole word

The proportion of the whole word taken up by the nasal was also examined in each target word; [ana] and [anna] were clearly distinguished from each other even among the 3-year-olds, although the distinction was clearest among the adults. The proportion of the nasal in [anna] was similar across all groups, around 50%. However, the proportion of the nasal in [ana] was larger in children's productions as compared with in adults' productions (e.g., 32.5% of [ana] among 3-year-olds vs. 23.8% among adults).

The results of the perception task are reported by the percentages of responses matching to the designated pictures (Table 7). Adults' responses were always matching to the designated pictures. The percentage of correct responses was around 90% among children, and the responses were better than chance in all groups (e.g., among 3 year-olds, $t(7) = 5.612$, $p < 0.001$).

| group | Number of subjects | matching | not matching |
|---|---|---|---|
| 3-year-olds | 8 | 90.6% | 9.4% |
| 4-year-olds | 8 | 93.0% | 7.0% |
| 5-year-olds | 8 | 88.4% | 11.6% |
| Adults | 10 | 100% | 0% |

Table 7. The percentages of matching responses in the perception task

In summary, it was apparent that Finnish children made a clear distinction between [n] and [nn] both in production and perception. However, when we compare Finnish children with adults, the distinction between [n] and [nn] was still clearer in Finnish adults' productions as compared to Finnish children's. The proportion of the nasal in [anna] was similar among all groups (around 50%). The proportion of [n] in [ana] was larger in children's productions (around 30%) when compared to adults' productions (23.8%). Thus the difference between children and adults was in production of [n] not in [nn].

## 5.2. Other observations

The same kind of difficulty in producing the short segments were also found in my observations of two children who were both acquiring Japanese as their primary language[9] (Aoyama 1999). Two pictures that were named *nene* and *nenne* were introduced to a 2-year-old (2;9). She correctly identified the two pictures, but she clearly avoided saying *nene* with a singleton. She referred to *nene* by saying 'not *nenne*':

(1) C (2;9)

C: kore *nenne* chan chiyau. 'This is not *nenne*.'
K: chigau, chigau nee. 'No, it is not. You're right.'
H: *nene* chan tte wakatteru. 'C knows that it is *nene*.'
C: (loud) kore *nenne* chan chiyau. 'This is not *nenne*.'
H: *nene* chan, *nenne* chan. '(This one is) *nene*, (that one is) *nenne*.'
K: un, soo soo. 'Yes, right, right.'

C: kore *nenne* chan chiyau. (crying) 'This is not *nenne*.'
K: chigau nee. 'No, it is not.'



(H: C's 6-year-old sister, K: investigator)

A 3-year-old (3;5) claimed that he thought *nene* and *nenne* were the same in the first session. A week later, in his second recording session, he said that he knew the difference between the two, although his productions of both *nene* and *nenne* still sounded *nenne* to adults. In (2), he was describing the difference between the two names to his father:

(2) S (3;5)

S:(pointed to *nenne*)
*kocchi ga* [nɛnnɛ], *deshoo*. 'This one is [nɛnnɛ], okay?'
F: *hai*. 'Yes.'
K: *un soo soo soo. sooda yo*. 'Oh yeah, yeah. That's right.'
S: (pointed to *nene*) *kocchi ga* [nɛnnɛ]. 'This one is [nɛnnɛ]'
F: *a*. 'Oh.'
S: [nɛnnɛ]. [nɛnnɛ].
(F: S's father, K: investigator)

This tendency was also reported in Estonian. Vihman (1997) reported that young children tend to lengthen the medial consonant, while geminate consonants are generally produced accurately. For example, *ema* 'mother' was produced as [am:a][10] (1;7), and *nuga* 'knife' was produced as [nug:a] (1;10) (Vihman 1997:216-217).

The difference was also found in productions of the pseudowords *ata* and *atta* between 18-month-old children with high genetic risk for dyslexia (GR+) and their controls (GR-) (Richardson 1998). In general, it appears to be the case that GR+ children had difficulties in using the durational cues of the Finnish quantity distinctions. This difference was in the production of *ata*, but not in *atta*; the difference in mean absolute duration of [tt] in *atta* was not statistically significant between the two groups (302 ms. among GR- children, 267 ms. among GR+ children). The mean duration of [t] was significantly longer among GR+ children (223 ms.) than GR- children (128 ms.) (Richardson 1998:169-170).

I originally expected that CV.CV structure would be easier for children to articulate than CVC.CV structure (see Demuth 1997). However, this does not seem to be the case as far as geminates are concerned. Children's difficulty or difference in making quantity contrasts originates in 'unstretchability' of singletons, in another word, difficulty in making the single consonant short, rather than making the geminate consonant long.

## 6. Summary

It was pointed out that, when there is a quantity or length contrast, a much greater scatter of values is allowed for long segments compared to their short counterparts (Jakobson and Halle 1971:34, Lehiste 1970:36, Trubetzkoy 1969). The empirical data presented in this paper seem to support the notion of 'unstretchability' of single segments; the difficulty or difference can surface in production and perception of short segments rather than of long segments. In adult production (experiment 1), the more distinctive contrast between [n] and [nn] in Finnish than in Japanese was achieved by making [n] short. In adult perception (experiments 2 and 3), identifying a word containing a short nasal seems to be more difficult than identifying a word containing a geminate. In acquisition, even though Finnish children (aged 3-5) made a clear distinction between short and long segments, the proportional duration of short segments was longer in children's production compared to adults', who made the clearest distinction between the two categories.

Lehiste noted that control over the timing of articulatory movements has to be involved for a quantity contrast (1970:6). In a pair of given short and long segments, more control over articulatory movements would be required for short segments rather than for long segments because of the 'unstretchablity' of short segments. Due to the relative nature of quantity contrasts, together with the necessary control over the timing of articulatory movement, it appears that the key in making a quantity contrast is in the duration of the short segments rather than the long segments.

## Notes

* This is a revised version of chapter 7 of my dissertation (Aoyama 2000). I thank Matti Leiwo and researchers of Early Language Development and Dyslexia project (ELDD) at the University of Jyväskylä for their help in collecting the Finnish data. Financial support for this research was provided from Center for International Mobility (CIMO) of Finland.

1. In Finnish, *hana* is a real word which means 'water faucet'.

2. Signalyze is a product of the InfoSignal company and is a speech analysis program for the Macintosh with spectral analysis tools.

3. Speech Analyzer is distributed by the Summer Institute of Linguistics, and is a speech analysis program for the IBM-compatible computers with the Windows operating system.

4. All the experiments were conducted in both Finnish and Japanese, but I only report the ones that are relevant to the discussions of unstretchability in this paper. For the results of other experiments, and cross-linguistic comparisons, see Aoyama 2001.



5. This token was chosen because it was relatively close to the overall average of *(h)anna* in the Japanese data.

6. SoundEdit 16 is an audio editing program for the Macintosh. It is a product of the Macromedia company.

7. The procedure in the adults' experiment (experiment 1) was kept consistent with that in experiment 4 for the purpose of this comparison.

8. I am grateful to Ulla Richardson, Pirjo Turunen of the University of Jyväskylä and Päivi Innilä of Nenäinniemitalo for their cooperation in the data collection.

9. The observations were done in Honolulu, Hawai'i, and the children naturally had an exposure to English.

10. I followed Vihman for transcribing long consonants here.

## References

**Aoyama, K.**: 1999, Nene or nenne?: Acquisition of the mora nasal. Poster presented at the VIIIth International congress, IASCL-International Association for the Study of Child Language, July 13, 1999. Donostia-San Sebastian, Basque Country/Spain

**Aoyama, K.**: 2001, *A psycholinguistic perspective on Finnish and Japanese prosody: Perception, production and child acquisition of consonantal quantity distinctions*, University of Hawai'i at Manoa Ph.D. dissertation; Boston, MA: Kluwer Academic Publishers

**Demuth, K.**: 1997, 'Stages in the development of prosodic words', *Proceedings from the 27th Child Language Research Forum* (ed. by E. V. Clark), 39-48. CSLI, Stanford, CA

**Jakobson, R., Fant, C. G. M.; Halle, M.**: 1952, *Preliminaries to speech analysis: The distinctive features and their correlates,* The MIT Press, Cambridge, MA

**Jakobson, R., and Halle, M.**: 1971, *Fundamentals of language*, Mouton, The Hague

**Lehiste, I.**: 1970, *Suprasegmentals*, The MIT Press, Cambridge, MA

**Richardson, U.**: 1998, Familial dyslexia and sound duration in the quantity distinctions of Finnish infants and adults, University of Jyväskylä PhD dissertation.

**Trubetzkoy, N. S.**: 1969, *Principles of phonology*, University of California Press, Berkeley and Los Angeles

**Vihman, M. M.**: 1997, 'The acquisition of accent: A framework for the study of Estonian prosodic development', *Estonian prosody: Papers from a symposium* (ed. by I. Lehiste and J. Ross), 203-221. Institute of Estonian language and authors, Tallinn