Intonation and Duration in English Questions of Japanese Speakers

Margaret Maeda

日本人英語学習者の発話における目立った特徴は主に２つあり，その１つは突出したWH疑問語と２つ目はやはり突出して発話される代名詞である。WHとyes/no疑問文の中でのピッチと長さのパターンを比較するため，先ず筆者は被験者を２つのグループに分けて，全員に対話文を発話してもらい，その発話をカセットテープに録音してもらった。グループ１は19名で全員が北米に住んだ経験がない。グループ２は15名で少なからず北米に住んだ経験をもっている。又，11人のアメリカ英語母語話者の協力を得た。次に筆者はKay Elemetrics CSLソフトウェアを使用し全ての録音済みテープを解析し，F0と長さを測りその結果を比較検討した。その結果は以下の通りである：グループ１　WH疑問語と代名詞のピッチが際立って高い。グループ２には１のような現象はないが，彼らは母語話者とはかけ離れているが，なるべくごまかすような方法を代名詞や助動詞の宿約のために使っている。グループ１はピッチの卓立を疑問文の最初に持っていく。その結果わかることは彼らは核音位置についての規則をまったく知らないということをF0と長さの測定結果と聴覚判定が証明している。グループ２は通常，核音がフレーズの最後の内容語上で落ちることを知っているようであるが，短い発話文章にもかかわらず，新情報と既知情報を区別するため正しく核音を使い分けることができたのは半数以下だった。

Key words: intonation, duration, WH words, yes/no questions, subject pronouns

Introduction

This study looks at intonation and durational patterns in the WH and yes/no questions of Japanese learners of English. Questions were chosen as the focus of the study because they include some typical features of Japanese learners' English. They pronounce WH words (*when, what, where, how,* etc.) with too much pitch prominence whereas native English speakers usually put the primary stress (the nucleus) later in the question and put a secondary or no stress on the WH word. In yes/no questions, such as "Do you like Japan?" or "Are you married?" elementary-level learners pronounce subject pronouns with pitch prominence. Native speakers do not normally stress subject pronouns (Allerton, 1978; Chafe, 1974). Learners pronounce the auxiliary verb and subject pronoun slowly and without contractions.

WH and yes/no questions were recorded and compared between 19 Japanese subjects who had not been to North America, 15 who had lived in North America for between one and seven years (see Appendix B), and 11 native speakers of American English. The learners were women students from two universities in Tokyo. All the subjects recorded a dialogue written to include questions which Japanese students often ask native speakers of English (see Appendix A). Measurements of fundamental frequency and duration were made of all the syllables in the recorded material using Kay Elemetrics speech analysis software.

The recordings of the Japanese subjects were judged by three native-speaking teachers of English for overall pronunciation level. They were not told that intonation and duration were the focus of the study. The recordings were judged for nucleus placement by the author.

1. Background

1.1 Structure of the intonation-group

This study is based on the unit of speech which has been referred to by writers on intonation using a variety of terms such as sense-groups, breath-groups, tone-units, intonational phrases and intonation-groups (Cruttenden, 1997). The term *intonation-group* is used in this study (Cruttenden, 1997). The questions in the dialogue were designed to consist of one intonation-group each.

The divisions of the intonation-group are those in the British tradition (Palmer, 1922; Kingdon, 1958; Crystal, 1969). An intonation-group always has a main stress, or *nucleus*. It may only have a nucleus, such as in the utterance, "Yes." The main pitch movement of the intonation-group is on the nucleus. (However for American yes/no questions, see Section 1.4.) Any syllables after the nucleus form the *tail*. If there are any secondarily stressed syllables before the nucleus, the first of them is the first syllable of the *head*. The head consists of all the syllables from the first stressed syllable up to, but not including, the nucleus. Any syllables before the first stress are the *pre-head*. The nucleus is the only obligatory part of the intonation-group. Figure 1 shows the structure of an intonation-group with all the parts.

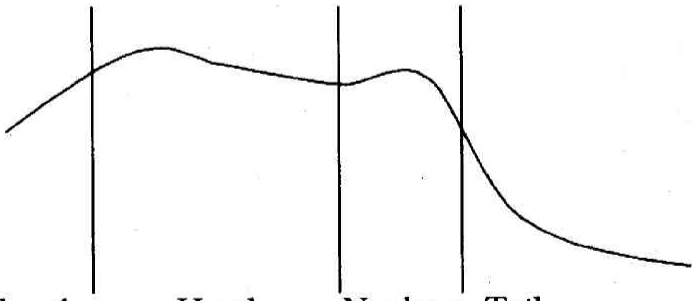


Figure 1 **The parts of the intonation-group**

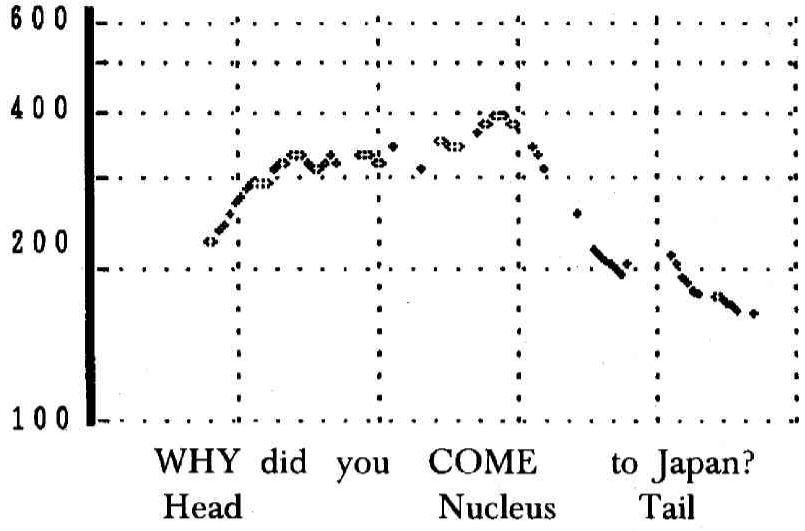


Figure 2 **The pitch line of a WH question (native speaker)**

1.2 Prosodic features

The three main prosodic features used in stressing syllables are pitch, length and loudness, usually considered to be important in that order. These terms are used for the features as they are perceived. The terms for the features as they are measured instrumentally are fundamental frequency, duration and intensity. Fundamental frequency will be referred to as F0. There is not an exact correspondence between the perceptual features and their acoustic correlates. For example a sound which is perceived to be twice as high as another may not have double the F0. Another example is the phenomenon of declination: when F0 peaks in an intonation-group decline gradually, they are perceived to be about the same height (Cooper & Sorensen, 1977). The uses of the terms for the acoustic and perceptual features are not kept strictly separate in this study.

1.3 WH questions

Figure 2 has a typical pitch line for a WH question spoken by a native speaker. The WH word normally has a secondary stress or no stress, and the nucleus, with the main pitch movement, is usually on a word later in the question. Some questions have two-word WH phrases, for example "How long" and "How old." They are sometimes

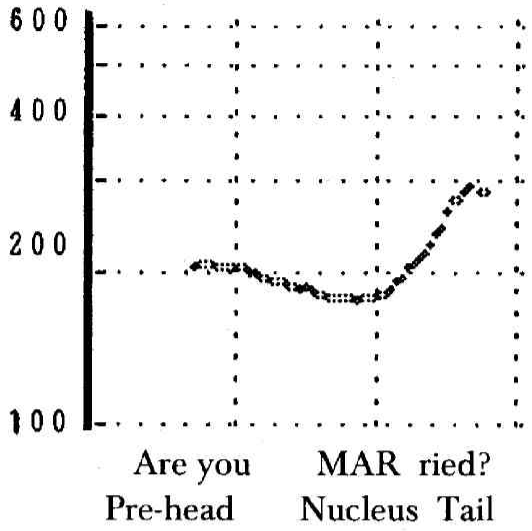


Figure 3 **The pitch line of an American English yes/no question**

pronounced with equal stress on both words, and sometimes with a stress on the second word.

1.4 Yes/no questions

Yes/no questions always have pre-heads because they start with unstressed auxiliary verbs, often followed by a subject pronoun. American English yes/no questions with the nucleus on the last or penultimate syllable typically have a low and level or gradually falling pitch line until the nucleus, and a sharp rise on the final syllable whether the final syllable is stressed or not (see Figure 3). When the nucleus is on the final or penultimate syllable, the final rise is typically the only major pitch movement in the question, and if there are secondary stresses, they are usually marked mainly by extra duration or intensity.

1.5 Duration

Syllables in the different parts of the intonation-group are spoken at different speeds. The unstressed syllables in pre-heads are pronounced more quickly than unstressed syllables in the other parts of intonation-groups (Jassem, 1952; Levelt, 1989). The greater speed of the pre-head is thought to contribute to indicating the beginning of a

new intonation-group (Cruttenden, 1997, Levelt, 1989). The next part, the head, is spoken at an intermediate speed, and the syllables in the nucleus and the tail are spoken the slowest. Final-syllable lengthening contributes to the perception of the end boundary of an intonation-group (Cruttenden, 1997). Stressed syllables are pronounced with greater duration than unstressed syllables within the same part of an intonation-group.

English seems to have wider variation in the duration of syllables than other languages. Delattre (1966) compared duration in English, German, French and Spanish. He found that the longest syllables in English were more than three times as long as the shortest, about 2.5 times as long in German and French, and about twice as long in Spanish. In English, there was more phrase-final lengthening than in German or Spanish. Unstressed final syllables in English were longer than stressed syllables in other parts of the English intonation-group. In German and Spanish, they were shorter. Cruttenden (1997) believes that final-syllable lengthening may be an intonational universal, but that there seems to be variation in the amount of final-syllable lengthening between languages, with particularly high lengthening ratio in English.

Phrase-final lengthening has been found for Japanese by Takeda et al. (1989). They compared durations of vowels in the syllables of isolated Japanese words with vowels in sentence-final syllables in real telephone conversations and in a conversation supposed to be read naturally by a professional speaker. In real conversation, the sentence-final vowels were longer than in isolated words, but in the read dialogue, sentence-final vowels were shorter than those in isolated words.

1.6 Nucleus placement

Teachers and textbooks often tell learners to put the nucleus on the most "important" word of a phrase or sentence, but it is not always easy to decide which is the most important word (Schmerling, 1974; Bresnan, 1971; Bolinger, 1972, 1989). The basic rule of nucleus

placement in English is that the last content word in an intonation-group is stressed (Cruttenden, 1997). When there are two content words near the end of an end of an intonation-group, a noun is often preferred over a verb, even if the verb is the last content word. This may be because nouns carry more semantic weight than verbs (Bolinger, 1989; Ladd, 1996). One function of nucleus placement is to highlight information which the speaker considers to be new to the listener. The dialogue used for this study has the question, "Do you like Japan?" (see Appendix A). Without a context, "Japan" would be stressed, since it is a noun, and it is at the end of the question. In this dialogue, the fact that the interviewee lives in Japan has been established, so Japan is *given* (old) information, and the new information is "like." Another function of nucleus placement is the marking of contrasts, for example, "I want the BLUE car, not the RED car." There are no examples of contrastive nucleus placement in this study.

The word which is logically the most important is often the one which takes the nucleus, and particularly if a phrase or sentence is short, it may not be difficult for learners to stress the right word. However, some common errors of nucleus placement by Japanese learners cannot easily be explained by claiming that they are not stressing the most important word.

Prominent WH words

Japanese learners tend to pronounce WH words with pitch prominence but there is no satisfactory semantic explanation as to why the nucleus is not normally on the WH word in English since it is logically the most important word (Ladd, 1996). In Japanese, the WH word or phrase is focused, and it generally has the main pitch prominence (Maekawa, 1991) and long duration. Native-speaking English teachers and speakers on teaching tapes tend to over-stress WH words to make sure that learners will catch what is semantically the most important word, but in English, and in many other European languages, the WH word does not normally take the nucleus. If the nucleus is placed on the WH word, this placement sounds marked

and, particularly in the case of "Why," may indicate some emotion such as exasperation.

Culicover & Rochemont (1983) give a syntactic rather than a semantic account of secondary stress on WH words. In languages which have WH movement, like English, the WH word moves from the site of the answer to the beginning of the question: "She lives IN TOKYO." "She lives WHERE?" "Where does she LIVE?" The moved WH word leaves behind a "trace," in the form of the nucleus, which stays toward the end of the question. (In English echo questions like, "You live WHERE?" the WH word is not moved, and the nucleus stays on the question word.) Japanese does not have WH movement. The question word is in the same position in the utterance as the word which answers the question: "DOKO ni sunde imasuka." "TOKYO ni sunde imasu," and the WH word and the word answering the question are both focused and pitch prominent.

Prominent subject pronouns

Japanese learners often stress subject pronouns: "Do YOU like Japan?" If learners are told to stress the important word, they may wonder why the person being addressed is not "important." The explanation given by Allerton (1978), Chafe (1974), and Halliday (1967) is that the pronoun is not stressed because the identity of the person addressed is obvious from the situation.

In the teaching of question intonation to Japanese learners of English, the end of the question is usually the part which is given the most attention. Students are taught that WH questions have a fall at the end and yes/no questions have a rise. In fact both types of question commonly have both rises and falls. The most common non-native sounding patterns of pitch and durational prominence occur nearer the beginning of questions. These features were examined to see what could be learned about Japanese learner's intonation and durational patterns in order to provide information which could be useful in the teaching of prosodic features to Japanese learners of English.

2. Method

2.1. Subjects

The Japanese subjects of this study were 34 first- and second-year women students at two universities in Tokyo. The 19 students who had not been to North America are referred to as Group 1. The 15 students who had been to local schools in the United States or Canada for between one and seven years are called Group 2. Most of the students were English linguistics majors. The rest were in elective English courses. They volunteered to be the subjects of the study in return for feedback on their pronunciation. They did not receive payment. Although they knew that their pronunciation was being studied, they did not know that the prosodic features of pitch and length were the features of interest. The third group consisted of 11 women living in the United States, in their thirties and forties, who are speakers of General American. They are referred to as native speakers or Group 3.

2.2 Materials

The subjects recorded a dialogue (see Appendix A) in their own homes on their own machines which was supposed to be an interview of a native-speaking English teacher. The reason for recording in this way was to give the students as much time as they wanted to prepare and record free of pressure. Although the questions were the focus of the study, the subjects were asked to read both the questions and the answers because the answers influence the nucleus placement in subsequent questions. Question 8 was not included in the results because some subjects in each group, including the native speakers, stumbled over it, probably because it was too long.

2.3 Measurement

Measurements of F0, duration and intensity were obtained for all the syllables in the recordings using Kay Elemetrics Corp. Computerized Speech Lab Model 4300B Software Version 5.X. The

measurements of intensity were not used in this study because the recordings were not made in the sufficiently controlled conditions necessary for the measurement of intensity (Borden et al, 1994). All measurements were re-checked twice at intervals of several weeks by the author.

WH words: F0

F0 was measured at the highest points of WH words, and in the middle of the vowel of the next words (see Figure 4) with the aim of taking measurements which correspond approximately to the parts of the words which make the WH word sound prominent. The WH words were spoken with a pitch rise by most of the native speakers and the more proficient learners. The last third of a pitch rise is the part best perceived (Rossi, 1971). Less proficient learners had a low-high-low pitch pattern on the WH word. The words after the WH words were short, unstressed words. Short tones with small rises or falls are perceived as level tones rather than as pitch slopes (Rossi, 1971).

Yes/no questions: F0

F0 measurements were taken for all the syllables of the three yes/no questions in the study, and they were plotted to compare pitch patterns of yes/no questions between the three groups. The points for

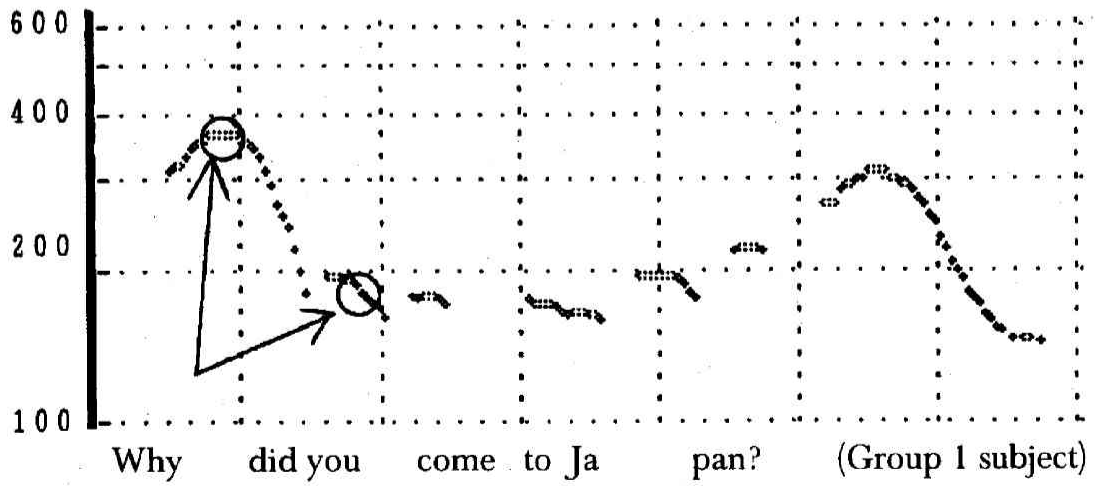


Figure 4 **Measurement points for F0 gap between "Why" and "did" (Question 7)**

taking F0 measurements were chosen, as for WH words, with the aim of using F0 values which correspond approximately to how the pitch contour is perceived. Prominent syllables were measured at the highest point if they were rising tones or low-high-low pitch curves, and short tones were measured in the middle of the vowel. If nuclei had a high-low-high curve or a downward slope, they were measured at the lowest point. Rossi (1978) has found that the last third of a pitch fall is the best perceived.

Duration of WH words and final syllable lengthening

Measurements of the durations of syllables were taken using waveforms, spectrograms and auditory judgment together. Durations of syllables can be difficult or impossible to measure accurately (Ladefoged, 2003). When measurements were being made for this study, there were two kinds of difficulties. In the questions with "What" followed by "do" or "does," the closure for /t/ is held through the closure for /d/, so an arbitrary cutoff point was chosen in the middle of the closure between the two words. Other places where an arbitrary cutoff point was chosen were between "your" and "wife" and between the two syllables of "married." There is a gradual change from vowel to approximant(s) to vowel in both cases. The lowest part of the second formant was taken as the cutoff point between the two syllables, following Ladefoged (2003). Durations of nuclei and tails were measured together as both are long, the nuclei because they are the main stresses and the tails because they are final syllables.

2.4 Judgment of nucleus placement

The author marked the nucleus placements of the subjects in the three groups where it was possible to do so (see Appendix A). Some of the native speakers and some of the subjects who had been abroad seemed to have two equal stresses in one question. In this case, both were noted. Most of the students who had never been abroad did not seem to have any idea of nucleus placement rules in English, so syllables which were prominent because of pitch movement and/or

long duration were noted and are referred to as *prominences* rather than as stresses or nuclei.

2.5 Judgment of overall pronunciation level

The recordings of the Japanese subjects were evaluated for overall pronunciation by three native-speaking university teachers of English. They were asked to make overall assessments of the subjects' pronunciation by dividing them into five groups graded "Very good," "Quite good," "Ordinary," "Rather weak," and "Very weak." These ratings are referred to as A, B, C, D and E (see Section 3.1). They did not know that intonation and duration were the features of interest or that some of the subjects had lived abroad and others had not.

3. Results

3.1 Judgment of overall pronunciation level

The three judges for the most part divided the subjects into those who had lived in North America and those who had not (see Table 1). Some exceptions were one Group 1 learner who was rated A by two judges and C by the other judge, and one Group 2 learner who was rated E by two judges, and C by the other judge. This Group 2 learner had been in the U.S.A. from age 14 to 18. There were two Group 2 learners who were rated differently by all three judges. Both were

Table 1 **Judgments of overall pronunciation level**

	Number of judgments	
Groups	1	2
Grade		
A	4(3)	21
B	8	12
C	9	8
D	11	3
E	24	2(1)

Note ( ) Number of subjects

rated A, C and D.

3.2 Nucleus placement:

Native speakers

Ideally, native speaking subjects' nucleus placement should be examined in recordings of natural conversation, but it is difficult to avoid using a scripted dialogue when the aim is to measure and compare exactly the same utterances spoken by a large number of subjects. Recording a scripted dialogue requires acting ability, but some native speakers seemed to think that speaking clearly was more important than speaking naturally, and this made some of their pronunciation sound over careful. Although most of the native speakers put the nucleus where the author expected, several native speakers put an additional stress on WH words which sounded as strong as their stress on the nucleus.

Group 2

Group 2's nucleus placement for Questions 2, 7 and 10 (see Appendix A) suggested they knew the basic rule that the nucleus commonly goes on the last *content* word of an utterance, not just on the last word, but most of them could not use nucleus placement to distinguish new and given information. In Questions 3 and 4, the last content word was not new information, and the native speakers placed the nucleus earlier in the questions. In Question 3, 7 of the 15 Group 2 subjects put a stress on the new information "like" compared with 10 of the 11 native speakers. In Question 4, only one Group 2 subject stressed the new information compared with 10 of the native speakers. She had lived in North America from the age of four to eleven and may have learned English more as a first than a second language. In Questions 3 and 4, the Group 2 subjects over-generalized the rule that the nucleus goes on the last content word.

Group 1

For Group 1, syllables were counted as prominent if they had large pitch movements, long duration or both. Most of the subjects in

Group 1 pronounced all the syllables evenly in the short Questions 3 and 6, so no prominences were counted for these questions. Some subjects had two or three equal-sounding prominences in one question. Each one was counted without indicating whether or not the subject had other prominences in the same question.

The WH words were prominent in about half the subjects' recordings, and "Why" was prominent in nearly all of them. About half the subjects put prominences on the last words of the questions whether they were content words or not. These results suggested that the subjects in Group 1 did not know the rules of nucleus placement in English.

3.3 F0: WH questions

The results for measurements of gaps in F0 between WH words and the next word are shown in Table 2. The native speakers, on average, said the WH words at a slightly lower F0 than the next word. Many of the subjects in Group 2 showed similar patterns to the native speakers. The mean F0 of their WH words was 16 Hz higher than the next word. A gap of about 16 Hz is perceivable but does not make a WH word sound very prominent, and because of declination (see Section 1.2), it may hardly sound prominent at all. The F0 gaps for Group 1 were large at an average of 66 Hz. A gap of this size makes a word sound clearly pitch prominent.

Table 2 **Mean F0 Height of WH Words Compared with Next Words**

	Group		
Question	1	2	3
1	66.42	28.47	0.73
4	89.37	14.00	−15.55
5	40.37	3.73	−8.10
7	67.53	17.80	1.18
Mean	65.92	16.00	−5.44

Note − (minus) The mean F0 for the WH word was lower than the value for the next word

Table 3 **F0 Gap Between the Peak of "How long" and "have" in Question 2**

	Group		
	1	2	3
Mean	89.37	82.53	17.27
Median	95	56	12

There were two questions with two-word WH phrases, "How long" and "How old," Questions 2 and 10. F0 gaps were measured for Question 2 between the highest point of the two words "How long" and the third word "have" ("How long have you lived there?"). With two words and a longer pitch line, there was more variation between subjects, both native and non-native, so the results for "How long" are listed separately (see Table 3). Two native speakers placed the nucleus on "How long." They may have treated the rest of the question as given (old) information, but this seems an unlikely stress pattern in natural conversation. Four native speakers had equal stresses on "How long" and "lived." This pronunciation sounded over careful. Four subjects in Group 1 put a pitch prominence on "How" only—an impossible placement in the context.

To check that the prominent WH words of Group 1 were not just part of a wider pitch range throughout the questions, the pitch ranges after the second word (the third word in Question 2) were compared among the three groups (see Table 4). The results show that Group 1 had the narrowest pitch ranges, Group 2 had pitch ranges between Groups 1 and 3, and the native speakers had the widest pitch ranges. It was concluded that the larger pitch movements on the WH words of Groups 1 and 2 were not part of larger pitch ranges generally.

In Question 10, "How old are they?" the WH phrase is normally focused. The nucleus is on the phrase "How old," and "are they" is the tail, so putting pitch prominence on the WH phrase "How old" is the normal placement. F0 gaps between the highest point of "How old" and "are" gave similar average values for Groups 1, 2 and 3. These values were 85, 89 and 97 Hz respectively. Group 1's pronunciation of

Table 4 **Mean Pitch Ranges of WH Questions After WH Words**

	Group		
Question	1	2	3
1	68	87	115
2	93	111	122
4	98	135	143
5	82	99	108
7	118	153	149
M of means	92	117	127

Note M = mean

Table 5 **Mean Total Durations of the Nine Questions in Seconds and as Percentages**

	Group		
	1	2	3
Seconds	12.265	11.022	9.053
Percentages	135.48	121.75	100.00

this question nevertheless sounded different from the native speakers' pronunciation because of different durational patterns and use of glottal stops before vowels. These patterns are described in Section 3.5.

3.4 Speech rates of the three groups

The mean durations of Questions 1 to 10, excluding 8, were added together in each Group. The results are listed in Table 5 in seconds and in percentages with the native speakers' durations set at 100%.

The articulation rates of Groups 1, 2 and 3 were respectively 205, 229 and 278 syllables per minute. These results show that the speed of speaking of Group 2 subjects was nearer to Group 1 than Group 3.

3.5 Durations of parts of WH questions

The durations of the WH words in Questions 1, 2, 4, 5 and 7 were added together and calculated as proportions of whole questions. The same was done for the nuclei of Questions 1, 5 and 7, and the

Table 6 **Durations of WH Questions and Final Syllables as Percentages of Whole Questions**

	Group		
	1	2	3
WH words	23.02	21.16	20.33
Final syllables	41.36	44.48	46.95

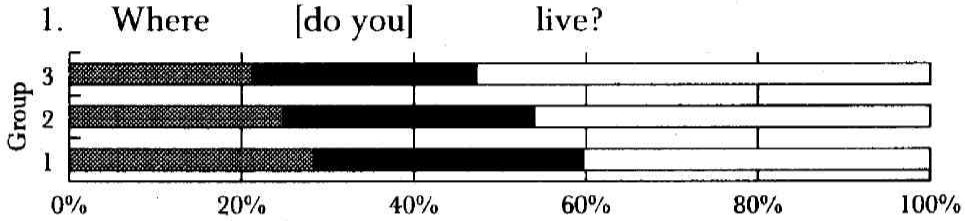
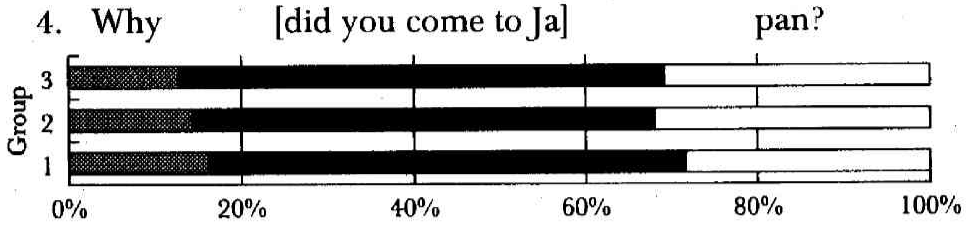


Figure 5 **Proportional durations in Question 1**

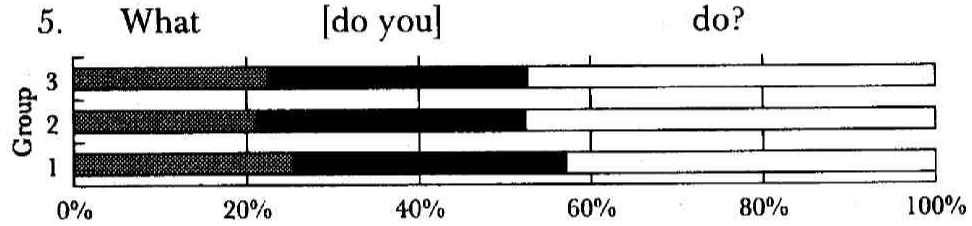
nuclei and tails of Questions 1 and 4 (see Table 6). These results showed that overall, WH words were the longest in the speech of Group 1, and nuclei and tails were the shortest. The reverse was true for the native speakers, and Group 2 had values between the other two groups. The durations in Question 1 were typical of this pattern (see Figure 5).

Questions 4 and 5 had different patterns from the general trend shown in Table 6 and Figure 5. In Question 5, "What do you do?" the duration of "What" for Group 2 was proportionately shorter, and slightly shorter in actual duration than for the native speakers (see Figure 6). They had about the same final syllable lengthening, proportionately, as the native speakers. All the learners said Question 5 more confidently and quickly than the other questions, probably because it is short and does not have any difficult sounds for Japanese learners. The Group 2 subjects had a native pattern of final syllable lengthening in Question 5, but they may have hypercorrected their durations for "What," or the native speakers may have pronounced WH words too carefully.

In Group 2's Question 4, the last syllable "pan" of "Japan" was both proportionately and actually longer than the native speakers'



Group	Why	[did you come to]	pan?
1	254	875	446
2	206	762	458*
3	155	677	369



Group	What	[do you]	do?
1	219	275	373
2	165*	245	374
3	167	221	351

Note *Values which were different from the general trend

***Figure 6* Proportional and actual durations (in ms) of Questions 4 and 5**

"pan" (see Figure 6) because of wrong nucleus placement. Most of the native speakers put the nucleus on "come," because "Japan" was given information, but most of Group 2 put it on "pan" (see Section 3.2). The native speakers' mean duration of "come" was 246 ms, and Group 2's was 229 ms. There was a similar pattern in Question 3 (see Appendix C), "Do you like Japan?" where "Japan" is normally treated as given information. Most of the native speakers put the nucleus on "like." Seven of Group 2 put it on "like," but two of the seven put an additional equal stress on "pan." The native speakers' mean duration of "like" was 279 ms and Group 2's was 240 ms. In Group 2, the last syllable in Questions 3 and 4, "pan," was long because of wrong

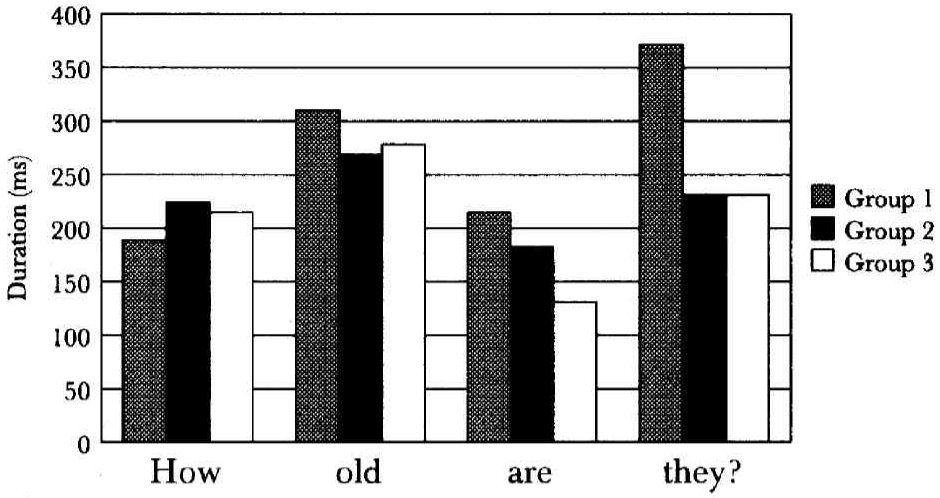


Figure 7 **Durations in "How old are they?"**

nucleus placement and final syllable lengthening. The native speakers put the nucleus, and therefore extra duration, on the middle sections of the questions while Group 2 put a secondary stress rather than the main stress on the verbs "come" and "like." The results of measuring durations support the conclusion reached from judging prominence placement by ear that most of Group 2 could not use nucleus placement to distinguish given and new information and instead over-generalized nucleus placement on the last content word.

3.6 Durations and glottal stops in Question 10, "How old are they?"

In Question 10, Group 1 pronounced "they" with very long durations (see Figure 7). This made the question sound as if it had two stresses on "old" and "they." They may have done this because beginners and elementary learners tend to pronounce too many stresses generally (Watanabe, 1988), or because they were stressing the subject pronoun.

A feature which was found for most of the Group 1 subjects and some of the Group 2 subjects, but only one of the native speakers, was glottal stops before "old" and "are" (see Table 7). Instead of linking "How_old_are they?" they separated "How old," and "old are" with glottal stops. Three of the glottal stops made by the Group 1 subjects

Table 7 **Glottal Stops Before Word-Initial Vowels in "How old are you?"**

Group	*old*	*are*	*old* and *are*	Total
1	6	0	9	24
2	3	1	1	6
3	1	0	0	1

were as long as a short syllable (55–88 milliseconds). Glottal stops before word-initial vowels is a common feature of Japanese learners' English. Native speakers sometimes put glottal stops before stressed words beginning with vowels, but too many and too long glottal stops in short utterances reduces the fluency of Japanese learners.

3.7 Prominent subject pronouns

The questions in the study with subject pronouns in the pre-head were:

3. Do you like Japan?
6. Are you married?
9. Do you have any children?

Other questions pronounced with prominent subject pronouns were Questions 2 and 4. The prominent subject pronouns found by aural judgment in the data were as follows: 17 in the recordings of ten of the Group 1 subjects, only one prominent subject pronoun in Group 2 and none among the native speakers. The pronouns were judged to be prominent if they had pitch rises or pitch peaks of 50 to 120 Hz and if they also sounded too prominent. Some pronouns with rises of over 50 Hz did not sound prominent because they had short durations or adjacent sounds were louder. These were not counted as prominent. Native speakers sometimes have raised pitch on subject pronouns without giving an impression of stress.

3.8 Intonation of yes/no questions

The results of plotting mean F0 measurements for Questions 6 and 9 showed that both groups of learners had relatively high F0 on

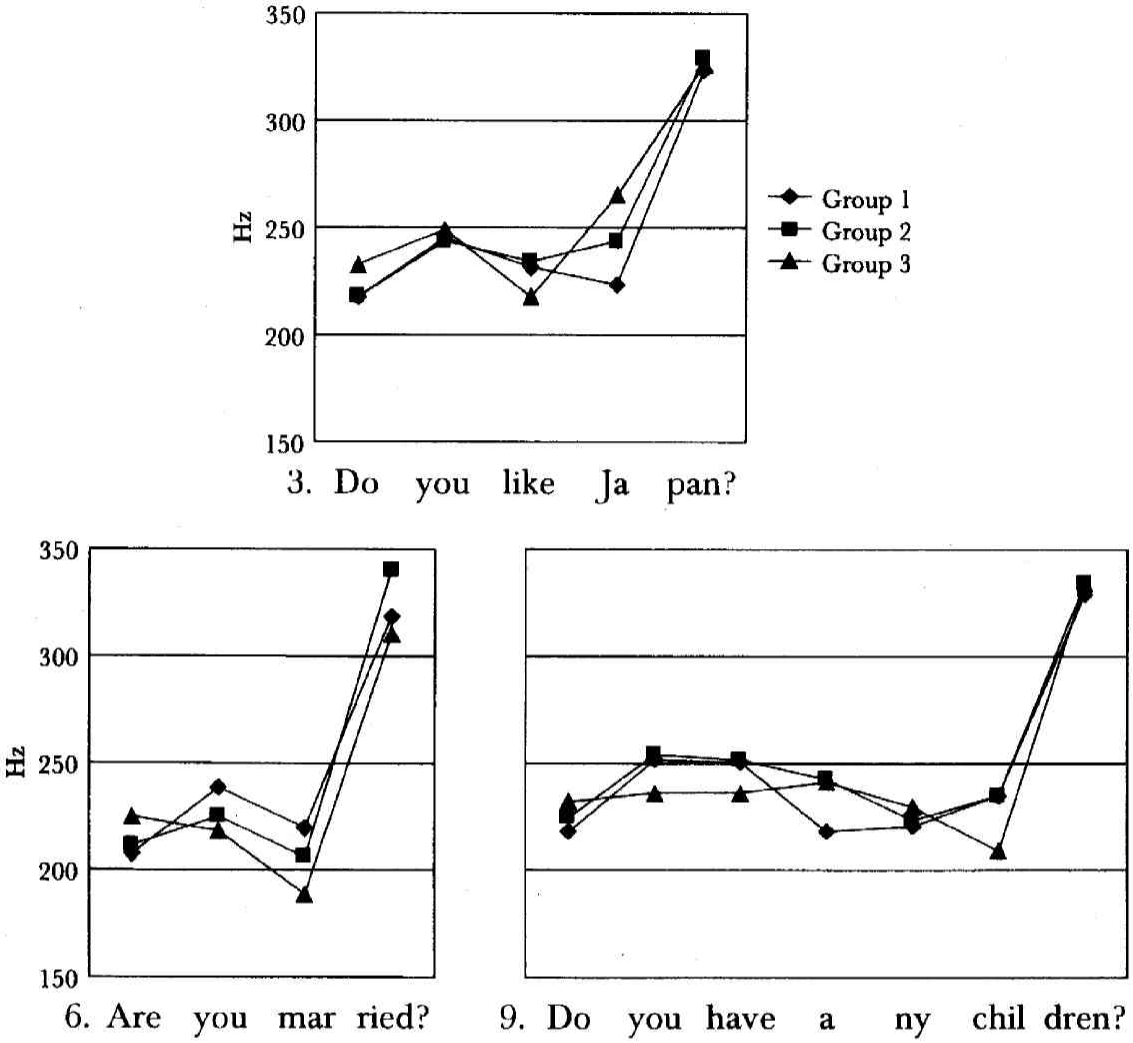


Figure 8 **Mean F0 of syllables in yes/no questions**

"you" and "have" (see Figure 8). In Questions 6 and 9, on average the native speakers had almost straight pitch lines before the final rise. In Question 9, "have" has a secondary stress but the native speakers mostly did not mark this stress with a change of pitch direction but with greater duration, intensity, or both. The lowest F0 of the native speakers was on the nuclei in all the three questions. Group 2 also had the lowest F0 on the nucleus in Question 6. This is the easiest question for the choice of nucleus. In Question 9, some native speakers had a raised first syllable on "children" and Group 2's pattern did not sound non-native.

In Question 3, the nucleus is before the penultimate syllable, so it

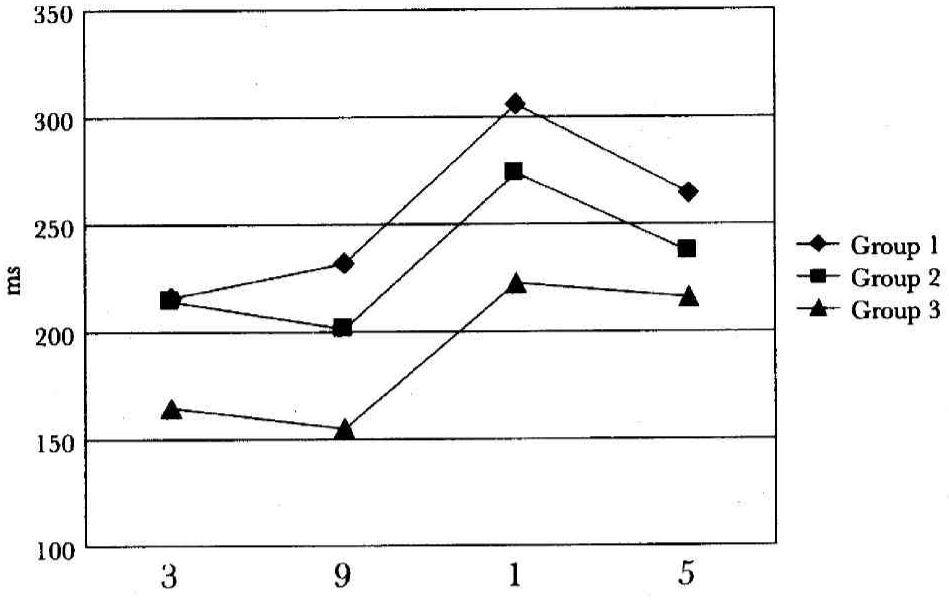
has its own pitch movement. The native speakers marked "like" with a big pitch movement, mostly a fall. The native speakers' "you" was high, but their rise on "Do you" had half the pitch range of the learners' rise, and the duration was shorter, so "you" did not sound prominent. The native speakers reached the lowest F0 on "like." Group 2 had a smaller drop in pitch on "like," probably because half the Group chose "like" as the nucleus and half chose "Japan." In Question 3, Group 1 had almost the correct pitch pattern for this question when "Japan" is the new information. However on listening to their Question 3, all the syllables seemed to be equally stressed despite the correct pitch pattern for a neutral context. Group 1 learners seemed to have learnt the pitch shape of yes/no questions in neutral contexts, but they did not choose any syllable for nucleus placement.

In Question 6, most of the native speakers started the final rise in the middle of the /r/ of "married." Some Group 1 and 2 subjects had a gradual rise from the beginning of the word "married" to the end of it, and some had a gradual rise from the beginning to the end of the whole of Questions 3 and 6. These patterns suggested that the learners did not know where to start the final rise.

3.9 Duration and contractions in pre-heads

It has been noted that the parts of the intonation-group are spoken at different relative speeds, with the pre-head being the fastest. The durations of "Do you" in pre-heads (Questions 3 and 9) are shown in Figure 9 together with the durations of "do you" in heads for comparison (Questions 1 and 5). All three Groups had shorter durations in pre-heads than heads. Durations for "Are you" (see the table in Figure 9) are longer because it is not contracted, but in the speech of the native speakers, "Are you" is still shorter than "do you" in heads.

Groups 1 and 2 had almost exactly the same durations for "Do you" in "Do you like Japan?" probably because Group 1 found this question easy compared with Question 9, "Do you have any children?"



	Questions		Groups 1	2	3
Pre-heads	3	Do you	222	219	164
	9	Do you	239	205	154
	6	Are you	292	252	198
Heads	1	do you	322	286	228
	5	do you	275	245	221

Figure 9 **Durations (in ms) of "Do you" in pre-heads and "do you" in heads**

Seven learners, one in Group 1 and six in Group 2, used non-native strategies for shortening pre-heads. These strategies were quite successful because the author listened to the questions many times before noticing them. They used an elongated "Do" and skipped "you," pronouncing "Do you like" and "Do you have" as "Do-oo like," and "Do-oo have." Some subjects pronounced the two words with a very weak [j] between them, and in two cases used a weak [w] instead of [j] resulting in something like "Dowoo" [duwu]. The [w] was probably easier to articulate in the context than [j]. One Group 2 learner used another strategy in Questions 3 and 6. She pronounced "you" as [ɪ] resulting in [duɪlaik] and [ɑːrɪmærɪd] for "Do you like?" and "Are you married?"

Summary and Recommendations

Group 2's WH words were similar in pitch prominence to those of the native speakers, and they had few prominent subject pronouns. Their durations in the first and last parts of the intonation-group were between those of Group 1 and 3. Some subjects in Group 2 appeared to be aiming at native durations for pre-heads but used non-native strategies for contracting auxiliary verbs and subject pronouns. Group 2 knew the rule of nucleus placement on the last content word, but they could only use nucleus placement to distinguish new and given information in a short utterance. Prominent WH words had almost disappeared, even in the recordings of learners who had only lived in North America for a year. This was probably a result of living in an English-speaking environment rather than having been taught to make WH words less prominent. It may be reasonable to suppose that unstressed WH words is not a difficult feature to acquire.

Teachers of English in Japan need to provide pronunciation teaching which will help learners to improve their pronunciation even without the benefit of a stay in an English-speaking country. The author believes it is necessary to draw learners' attention to their prominent WH words and contrast them with secondary or no stress on WH words in English. The learners' pitch patterns in WH questions can be contrasted with the native pattern by pronouncing both the learners' pattern and the native pattern, and by drawing pitch lines of learners' and native WH questions on the board. The teacher can explain that WH words are prominent in Japanese but not normally prominent in English. It is important to point out that prominent WH words may make questions sound aggressive, particularly questions with "Why?" Students should be exposed as little as possible to "Teacher Talk" questions with prominent WH words. They are often modeled on teaching tapes with large prominences on WH words.

Pronunciation of yes/no questions, particularly pre-heads, can be taught using the same technique as for WH questions. The learners'

patterns, for example prominent subject pronouns, can be contrasted with the target pattern by pronouncing both patterns and by drawing pitch lines for both patterns on the board. Teachers should explain that the first part of the question is flat, low and fast, and that learners should aim their intonation downwards to the nucleus, and then up on the final rise. This is not the only pattern that native speakers use, but it is frequent and will give a native-like pitch pattern to their yes/no questions.

Many Japanese learners think that native English speakers are simply skilled at articulating very fast, and they do not realize that native speakers use shortcuts, for example, reducing two syllables to one. This could be seen in Group 2's strategies for pronouncing "Do you." Although many textbooks draw attention to contractions, they are still new to a surprising number of learners. It is important to teach "D'you" or "Ju" and other contractions, such as the dropping of /h/ in "Is he," and the pronunciation of schwa as in "Is there a" /ɪzðərə/. If learners use these contractions, they will avoid the low-high-low pattern that is so often heard in the classroom in phrases such as "Do you like...?" or "Is there a...?" Contractions will also reduce the duration of pre-heads relative to the other parts of the utterance.

Speech contests are popular in Japanese schools and universities, and teachers who prepare students can indicate where nuclei should be placed on their scripts and train students to put major pitch movements and longer durations later in their sentences where the nucleus usually is. Students tend to start sentences high and loud, and fade out on the last part. Since this part usually has the important information, the meaning is often lost. If students put longer durations and bigger pitch movements on the nuclei, they will help the listeners to process the important information and make their speeches easier to understand.

In conversation, learners have to decide for themselves which word in an intonation-group takes the nucleus. Some rules of nucleus placement are easier to learn than others. Group 2 seem to have learned the rule for placement on the last content word, but only one

subject used nucleus placement for distinguishing new and given information to the same extent as the native speakers. In using nucleus placement to separate new and given information, it is thought that the speaker unconsciously estimates what is already in the listener's consciousness and uses nucleus placement to highlight information considered to be new to the listener (Chafe, 1974). This use of nucleus placement could be introduced to learners through the marking of nuclei in dialogues, but judging from the lack of this nucleus placement in the recordings of Group 2, particularly in the longer Question 4, and possibly because of complicated mental processes involved in deciding what is new and given information, it may be that this is a difficult use of nucleus placement to acquire.

Note

The printouts of pitch lines in Figures 2, 3 and 4 were obtained using *Onsei Roku-Bun-Ken* [sound recording-hearing-viewing] speech analysis software (Imagawa & Kiritani, 1989).

References

Allerton, D. J. (1978). The notion of 'givenness' and its relation to presupposition and to theme. *Lingua* 44, 133–168.

Bolinger, D. (1989). *Intonation and its uses.* Palo Alto, CA: Stanford University Press.

Borden, G. J., Harris K. S. & Raphael L. J. (1994). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech* (3rd ed.). Baltimore: Lippincot Williams & Wilkins.

Bresnan, J. (1971). Sentence stress and syntactic transformations. *Language* 47, 257–282.

Cooper, W. E. & Sorensen, J. M. (1977). Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America* 62, 683–692.

Chafe, W. L., (1974). Language and consciousness. *Language* 50, 111–133.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge University Press.

Crystal, D. (1969). *Prosodic systems and intonation in English.* Cambridge

University Press.
Culicover, P. & Rochemont, M. (1983). Stress and focus in English. *Language* 59, 123–65.
Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* 4(3), 183–198.
Imagawa, H. & Kiritani, S. (1989). Extraction of pitch and formants at real time using DSP, and its application to pronunciation training. *Technical Report of the Society of Electronic Information Communication* SP 89–36, 17–24.
Jassem, W. (1952). *The intonation of conversational English.* Warsaw: La Société des Sciences et des Lettres.
Kingdon, R. (1958). *The groundwork of English intonation.* Longman: London.
Ladd, D. R. (1996). *Intonational Phonology.* Cambridge: Cambridge University Press.
Ladefoged, P. (2003). *Phonetic Data Analysis.* Oxford: Blackwell.
Levelt, W. J. M. (1989) *Speaking: From Intention to Articulation.* Cambridge, MA MIT Press.
Maekawa, K. (1991). Perception of intonatonal characteristics of WH and non-WH questions in Tokyo Japanese. *Proceedings of the 12th Congress of Phonetic Sciences* 4, 202–205.
Palmer, H. E. (1924). *English intonation with systematic exercises.* Cambridge: Heffer.
Rossi, M. (1971). Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 23, 1–33.
Rossi, M. (1978). La perception des glissando descendants dans les contours prosodiques. *Phonetica* 35, 11–40.
Takeda, T, Sagisaka, Y. & Kuwabara, H. (1989). On sentence-level factors governing segmental duration in Japanese. *Journal of the Acoustical Society of America* 86(6), 2081–2087.
Watanabe, K. (1988). Sentence stress perception by Japanese students. *Journal of Phonetics* 16, 181–186.

Appendix A
Test Dialogue with Expected and Actual Nucleus or Prominence Placement

Native Speakers

1. Where [3] do you **live**(11)?	In Tokyo.
2. How long (1)[4] have you **lived**(10) there?	Five years.
3. Do you **like**(10) Japan(1)[2]?	Yes, very much.
4. Why(1) did you **come**(10) to Japan[2]?	I wanted to work abroad.
5. What(1) [2] do you **do**(10)?	I'm an English teacher.
6. Are you **mar**(11)ried?	Yes, I am.
7. What[2] does your **wife**(11) do?	She's an English teacher, too.
8. How does **she** like living in Japan?	She likes it a lot.
9. Do you have(1)[3] any **chil**(10)dren?	Yes, two.
10. **How old**(9) are they(1)?	Eleven and fourteen.

Group 2

1. Where(4) [3] do you **live**(8)?
2. How long (2)[3] have you **lived**(12) there(2)?
3. Do you **like**(7) Japan(3)[2]?
4. Why(1)[2] did you **come**(1) to Japan(11)?
5. What(3) do you **do**(12)?
6. Are[1] you **mar**(12)ried?
7. What[3] does your **wife**(14) do?
9. Do you have(1)[2] any(1) **chil**(14)dren?
10. **How old**(9) are(1) they(1)?

Notes

1. Syllables in bold print are those the author expected to take the nucleus.
2. ( ) the number of subjects who put the nucleus on the preceding

syllable.

3. [ ] the number of subjects who also had a nucleus on the syllable in bold print.
4. Where nuclei could not be detected, the total number of nuclei is smaller than the number of subjects.
5. "How long" and "How old" were treated as one word for nucleus placement.
6. Question 8, and all the answers, were not used in the study (see Section 2.2).

Group 1

1. Where(10) do you live (9)?
2. How long(8) have you(1) lived there(9)?
3. Do you(3) like(1) Japan(1)?
4. Why(17) did you come(2) to Japan(7)?
5. What(8) do you do(12)?
6. Are(2) you(1) mar(4)ried?
7. What(7) does your wife(8) do(11)?
9. Do you have(1) a(2)ny chil(6)dren?
10. How old(10) are(1) they(10)?

Note

( ) the number of subjects who put a prominence on the preceding syllable.

Appendix B

Ages and Lengths of Stay in North America of Japanese Subjects

Age in years	Length of stay (years)	Age in years	Length of stay (years)
14–18	4	14–18	4
18–19	1	18–19	1
4–12	7.5	20–21	1
7–11	4	10–16	5.5
16–20	4	16–20	4
10–16	5.5	10–13	3
7–11	4	14–19	5
21–23	2		

Appendix C

Duration of Beginning and End Syllables