The language recognition ability of four-month-old infants in English and Japanese spoken by a bilingual and Japanese monolingual woman

Akane Watanabe
Waseda University

This study investigates whether four-month-old Japanese infants are able to distinguish English from Japanese spoken by a bilingual woman (level 9.5) (Experiment 1), and then, spoken by a Japanese monolingual woman (level 1) who has learned English fifty years ago and never been to English speaking countries (Experiment 2). According to Mehler et al. (1993), rhythmic cues play a role for infants when two languages are presented. The bilingual speaker has indicated English rhythmic patterns as well as Japanese, whereas the monolingual speaker has not.

The experimental procedure followed Bosch & Sebastian-Galles (1997 & 2001) that have investigated four-month-old infants' capacity to discern two rhythmically similar languages. In both researches, the scholars let infants sit between two loud speakers and recorded their reaction towards one of two speakers presenting a stimulus. Then, they have coded the infant's Reaction Time (RT) (1997) as well as listening time (2001) to the stimuli and analyzed them.

In our first experiment, ten Japanese monolingual-to-be infants were tested with English and Japanese stimuli narrated by a bilingual woman to assess the validity of the procedure. In the main experiment, mostly same infants were examined in above two languages spoken by a Japanese monolingual woman. The data from those two experiments were coded with special editing software to measure two different types of reaction, RT and listening time. The result of the first experiment indicated the validity of the procedure, i.e. that Japanese four-month-old monolingual-to-be infants distinguished English from Japanese when the bilingual woman spoke. In the second experiment, however, the infants did not discriminate those languages spoken by Japanese monolingual woman (level 1).

1. Introduction

1.1. Research questions

Reflecting the social demand on English, a large number of parents wish that their children acquire English as early as possible. Some of them use English by themselves with Japanese rhythmic property, inconsiderately. (To make the matter worth, some neglect the famous theory, "one-person one-language rule", suggested by the famous Jules Ronjat (1913) [1] and supported by many others.) The rhythmic, i.e. prosodic property is an indispensable element for young infants to recognize a language (Mehler et al., 1996; Nazzi et al. 1997; Bosch & Sebastian-Galles, 1997). Do infants distinguish, however, those two languages spoken by monolingual Japanese parents? Doesn't the rhythmic element used on both English and maternal Japanese make infants confuse?

This research was conducted to answere the following research questions:

1) Do infants differentiate English from Japanese spoken by a bilingual woman with RT and HPP methods carried by Bosch and Sebastian-Galles in 1997 and 2001?
2) Do infants differentiate English from Japanese spoken by a monolingual woman?

1.2. previous works

There are numerous tasks children face to acquire a language. For one thing, they "have to master the sound system of speech and attach the sounds to their reference." Then "there is a task of developing lexicon and a system of meanings. The syntactic system must also be acquired and finally, all of these must be used appropriately in social interaction." (Mc Laughlin, 1984)

The very first step of language acquisition preceding mastering the sound system will be speech perception ability. It was until around Leibermann's finding (1973) that many researchers believed newborns have no auditory perception due to the liquid in the tympanum. In realty, however, infants are interested in human voice from the very first day, in women's voice above all, but not to any noise nor tone (Largo, 1995). One-month-old infants are already capable to discriminate between consonant sounds, such as [ba] and [pa] (Eimas et. al., 1971). Those speech perceptions depend on interaction with the environment, since the children of deaf parents are not able to discover sound patterns merely through exposure to television or radio. The recurrence of sound patterns at time of significance to the child appears to mark off particular sounds for attention (Ervin-Tripp, 1970b).

The speech processing ability increases its importance particularly when children are exposed to two languages. In other words, separating languages will be a prerequisite for their language learning as Mehler & Christophe suggest (1995). In deed, there are many studies analyzed infant's early competence to distinguish between languages. In spite of some deficits such us the omission of subject's age or the use of connected speech, the following pioneer works are remarkable; A French newborn distinguished French from Russian (Mehler 1988), a five-month-old

[1] "There is nothing to be taught(children how to speak). It is sufficient simply to speak to him when the occasion to do so arises in one of the language he is to learn. But here is the crux: each language must be embodied in a different person. You, for instance, should always talk French to him and his mother German. Never switch roles. In this manner he will begin to speak two languages without noticing it and without having to put forth any special effort in learning them (Ronjat, 1913, p3)"

separated English and Spanish (Bahrick & Pickens 1988), two-days-olds distinguished English from Spanish and preferred mother tongue (Moon 1993).

In 1996, Mehler et al. proposed what kind of clue exactly infants use to differentiate two languages. Through their prior studies, they discovered that newborns (four-day-old and two-month-old) rely on robust acoustic cues, i.e. prosodic information, hence, duration (time) and amplitude (intensity) of vowels to discern two languages. The proposal was named TIGRE (Time and Intensity Grid REpresentation). In deed, neonates are already able to distinguish two non-maternal languages when they are prosodically distant; i.e. they can discriminate Russian from

French or Italian from English (Mehler et. al. 1988). It also agrees with other conclusions out of various approaches; that infants attune themselves to the maternal vowel structures (prosodic properties) up to six months of age (Kuhl et al.,1992; Werker & Ladonde., 1988), that infants' metaphysical recognition at phonological level develops from larger phonological unit to the smaller (Treiman, 1992), and so on.

In TIGERE proposal, they also predicted difficulty on the perception of two rhythmically similar languages (Mehler at.el,1996). Nazzi et al. (1997) offered evidence in this direction in that French newborns can distinguish English from Japanese but not from Dutch. In another study, neonates were neither able to distinguish Spanish from Italian nor English from Dutch (Nazzi, Bertoncini, & Mehler, 1998). However, Bosch & Sebastian-Galles (1997) presumed that the subjects were probably not enough mature to distinguish small difference in rhythmic cues between Italian and Spanish or English and Dutch. In fact, they showed that four-month-old infants were actually able to differentiate two other rhythmically close languages, Catalan from Spanish and Spanish from Catalan using RT method. In 2001, they provided the same results in a different method called HPP (See section 1.3).

1.3. Controversial prosodic classes

As has been mentioned, infants learn about prosodic properties of the maternal language around six month of age intensively. The acquired properties could play a role in fluent speech in the future (Culter 1990, Culter & Butterfield 1992). What the prosodic properties are, i.e. rhythmic categories are, controversial. The categories derived from speech perception studies of adult, especially from the famous research of Mehler et al. (1981). They suggested that French, Spanish, Catalan and Portuguese speakers segment the fluent speech into "syllable size units". On the other hand, Culter et al. (1986) reported that English and Dutch speakers analyze the speech in strong and weak syllables, i.e. in "stress". Those "syllable timed" vs. "stress timed" had been the only classes until the concept of "mora" appeared.

The word "mora"[1] appeared in Otake et al. (1993) for the first time. Their result was that Japanese speakers segment successive speech into mora, neither into syllable nor stress.

In spite this third prosodic class had been permeated, Otake himself suggested a new theory recently that Japanese children may be sensitive to syllables. According to his study, children identified two chunks (pan-da) rather than three (pa-n-da) (Otake and Yoneyama, 1999). Though it seemed just one of few exceptional cases, Otake declared children recognize mora at last in a school period. Before they learn to read and write letters, he says, children recognize syllables (and probably some mora) (Otake and Imai, 2001). Hence, it could be said that the rhythmic properties theory went back to the first position: syllable-timed and stress-timed. At any stage, English and Japanese have never been rhythmically similar, since Japanese have been categorized either into "syllable-timed" or "mora-timed" where English was always "stress-timed".

1.4. The study of Bosch & Sebastian-Galles (1997) and (2001)

The purposes of the first half of 1997's study were to measure the infants' ability on language discrimination in prosodically different and similar languages, and also to know if they rely on the prosodic properties to distinct them.

With 10 Catalan and 10 Spanish monolingual 4-monts-olds, recorded sentences, and RT (Reaction Time) method, they figured out following results; 1. Infants react faster to the familiar materials (maternal language, i.e. either Catalan or Spanish) than English. 2. They also discriminate above two intra-class languages (maternal Catalan from Spanish or maternal Spanish from Catalan). 3. They differentiate those languages through some kinds of distinct rhythmic properties. The method used for the last result was questioned by Lieberman in the following year (1998).

In 2001, they reinvestigated the first and second research questions from the above study applying another method. Number of subjects enhanced, recorded stimuli reordered, and apparatus slightly developed. In addition, they showed pictures between each stimulus. The new method was HPP (The Head-Turn Preference Procedure) which has been extensively used for infant' speech perception. Two results were the same as 1997. They did not though figure out how infants discriminate those intra-class languages.

1 "metrical" mora

Mora is one Japanese prosodic category suggested by Otake in 1993. Normal mora consists of either a vowel, a semivowel and a vowel, a consonant and a vowel, or a consonant, semivowel and a vowel. Special Mora is a long vowel, choked sound, syllable nasal.

2. Experiment 1

2.1. Purpose

In the first experiment, we assessed if the infants were able to differentiate English from Japanese spoken by a bilingual woman. At the same time, it tested the adequacy of the methods (RT and HPP) mostly following to Bosch and Sebastian-Galles (1997) (2001). In the RT method, we coded and analyzed both "the first" and "the second saccadic" movements to verify 1997's result on those movements.

2.2. Method

2.2.1. Subjects

Eleven four-month-old Japanese infants who were carefully assessed by means of interview and a questionnaire were chosen (See Appendix 1). They were between 136 to 152 days at the recording time, and did not have auditory problem. Their mother tongue was Japanese, for their parents spoke to them only in Japanese from the birth. They have hardly heard native English (probably shortly in a park or train) directly, and have heard English through audio-visual equipment only a little. (A mother listened to an English learning program for adults sometimes. Another mother had watched American movies near her awaken baby. Two mothers had presented their babies a world-language video called "Baby Einstein" up to the third months occasionally. Presentation of English in this video was not very long. Two mothers had listened to English songs for children at home or in the car).

Two additional subjects were also tested, but rejected due to fuzziness and failure of the camera.

2.2.2. Recording and Editing of Auditory Stimuli

A bilingual Japanese female adult lived in America from twelve to eighteen-year-old (A woman was selected since female voice catches infants' attention easier (Largo, 1995.)) was asked to read out sentences as if she narrated for children after she looked at a picture book of "The Snowman". She had to read each sentence in four seconds.

English sentences (n=20) were mostly quoted from the pioneer work (Bosch 1997), which have been prepared corresponding to drawings of "The Snowman" by Raymond Brigg. We, however, added coordinate clauses and reduced single clauses so that each type of clause appeared more or less equally. The sentences became, therefore, either single clause (n=6), coordinate clause (n=6), or a main clause with a subordinate (n=8) (See Appendix 2). In addition, they were all declarative as well as "phonological utterances", the top of seven prosodic hierarchies of Nespor and Vogel (1986). Repetition except for pronouns was avoided as much as possible.

Japanese sentences (n=20) had close sense to English ones, though they were not exact translation of the former. We rather paid attention so that it sounded naturally and become around 25 syllables. Each type of clause appeared almost equally; 6 single clauses, 7 coordinate clauses, and 7 main clauses with a subordinate (See Appendix 2). They were also declarative and "phonological utterances". Repetitions occurred very few.

A difference in the mean number of syllables between two languages were found (12.6 sylls. in Englihsh; 24.55 sylls in Japanese, probably due to the different (syllable vs. stress timed) properties of each language.

After recoding of those 21 sentences (16,000Hz) in a sound-attenuated room, the bilingual woman's English was introduced to an American man and estimated as native level (level 9.5) (See Appendix 3) in the matter of pronunciation and fluency. We asked him a further comment on her Pronunciation and Reading Fluency following to Phone Pass guideline (See Appendix 4). According to the man, "Her level of Pronunciation and Reading Fluency was very advanced", "She produced consonants, vowels, and stresses consistently in a native-manner", "She read with adequate rhythm, phrasing and pausing". Two German introductive sentences (n=2) were narrated by the author for familiarization phrases (See Appendix 2). The quality of pronunciation was guaranteed by a native German speaker.

After three experiments with two blocks each (one block had consisted of 7 sentences in each language), the stimuli were reedited in one block, since infants could not concentrate in the second block at all. Ten sentences holding more syllables in each type of clause were finally picked up out of each language, because the bilingual woman spoke faster than the monolingual and they were the ones being always around four seconds. Mean duration of utterances of those 20 sentences was similar in each language, 3680 ms. in English (SD=81), 3670 in Japanese (SD=66).

The selected 20 sentences, half English and half Japanese, were edited with a sound editing software carefully. Each sentence was attached to a silent interval of more or less four seconds, so that a phrase became just eight seconds approximately. Following to two German sentences through both speakers, those twenty sentences together with silence in each were randomized under the following restrictions: the stimuli can come out less than three times through the same side, and the sentences can be cited less than twice from the same language. Two images to get infants attention at the beginning (8×2=16s.), the familiarization(8×2=16s.), and twenty randomized sentences in Japanese and English (8×20=160s.) made a 192 seconds block for Experiment 1 (See Figure 1).

2.2.3. Apparatus

To record and edit the utterances, we used a microphone, computer, a sound editing software called “Sound Engine”. A digital video camera (Sony DSC-P2), two loudspeakers (Sony SRS-Z900), two computers, a curtain to hide the camera, and a lamp were settled for the test. A video editing software called “Video Editing Magic” with time code system even for a millisecond helped our coding.

The test booth was built in a quiet Japanese tatami-room. The control booth with a computer was settled behind a corner between four successive paper doors (1.8×0.9×4) in white. We made a chink between the doors for observation. A child chair in front of a mother’s seat with a MD player was set inside of the room at a distance of 75 cm facing a computer monitor in the center, camera above center, and two loudspeakers covered with woman photos at 35 degrees to the right and left on the linear of the monitor.

2.2.4. Procedure

The entire procedure took place in a single session, near 3.5 min. As we mentioned above, a trial began with two colorful images on the screen settled in the middle. As soon as the second image disappeared, the familiarization phase started: a German sentence from the left, then another from the right. Finally, the test phrases commenced. Bilingually spoken English (n=10) or Japanese (n=10) were presented through the left or right side of speakers randomly under the restrictions (See Figure 1).

We dared to take this procedure, which followed the Bosch and his colleague’s former work, both for RT measurement and HPP, though they inserted one image before each stimulus in 2001.

seconds	Stimuli			
01-08	Image A			
08-16	Image B			
16-24	Left speaker	German 1 (familiarization)	Right speaker	
24-32				German 2 (familiarization)
32-40				English 1
40-48		Japanese 1		
48-56				Japanese 2
56-64				English 2
64-72		Japanese 3		
⋮		⋮		⋮

Figure 1. The Stimuli

In the meantime, infants seated either separately on a child chair in front of the mother, or on her lap when he/she was fuzzy, facing to the computer monitor between two speakers covered by women’s photos. The mother was instructed not to move, and listen to relaxing music through a MD player. Commands were given through the computer behind the paper-made doors with a chink by the author. The whole session was recorded with the digital video camera.

The rational behind the procedure was that if the infants have listened to a familiar rhythmic representation only in Japanese sentences, they might show shorter reaction time in RT measurement and longer listening time in HPP than in English.

2.2.5. Coding procedure

We selected two kinds of method from the prior studies as have been mentioned. One was the RT (Reaction Time) measurement used by Bosch and Sebastian-Galles in 1997. RT was defined as the duration between the moment that a stimulus was given and that the infant looked at the sound source (or the other side mistakably). The other was HPP (The Head-Turn Preference Procedure) applied by the same researchers in 2001. HPP measures listening time between the RT and the end of infant's attention. Both durations were measured using special software scrolling mill-seconded time-code without the sound. The coding has been carried three times thoroughly for each method, until each came to the agreement.

At the beginning of 1997's paper, two possibilities to code RT, or "orientation latencies" in other words, were suggested. After full investigation, they selected the second "direct saccadic movements" as the "reaction" to one of two presented languages. The first movements, which seemed like the reaction to the sound, were finally rejected. We decided to code both the first and the second movements in experiment 1 to verify the prior study.

A trial became null 1) when the infant did not look at the display for the last two seconds (RT, HPP), 2) when she failed to orient either side (RT, HPP), and 3) when she did not be diverged after the next stimulus began (HPP). Wrong localizations were also included according to previous works that proved the insignificance on the localization.

2.3. Result and Analysis

2.3.1. Reaction Time

56% of valuable response to English and 60% of it to Japanese was seen including wrong localization (32%) that agreed to other developmental works on this method. Certain infants (n=2) often looked at the camera instead of two laud speakers. Those failures of the orientation were turned down (n=14 in English, 11 in Japanese) since it corresponded to the first restriction. Cases infants did not look at the display for the last two seconds have occurred always at the end of the test provably owing to the tiredness (n=10 in English, n=12 in Japanese). Other declined cases were that infants' older siblings opened the door and hindered the experiment, as well as that infants lost the interest and stopped reacting totally.

In fact, most of infants did not turn the head directly to the sound source in our study. In stead, they showed a kind of movement (the first movement) before they looked at one of two speakers (the second movement). What kind of movement appeared at first was due to the infant's attitude what he or she was doing just before the stimulus. When he was playing with his own tows, for instance, he provably looked up once, then started looking for the source. When his head had been sustained in the middle or rather upper (this was in fact very frequent case), he closed the eyes before turning the head. Other cases seemed like just a habit towards a preferred direction as Bosch explained.

After careful examinations, we defined the "direct saccadic movement" (the second movement) as the moment that the infant made the widest eyes to the sound source just after he/she turned the head (To reach the second restriction, he had to hold the posture for a while after the coding point.). The definition of RT (orientation latencies) became, thus, either from the moment stimulus was given to the first reaction or to that the infant opened his eyes at most to the sound source. That of the time for HPP (listening time) was length between the second movement and the deviation.

As the former research revealed between English / Catalan and English / Spanish, there was no significant difference between the reaction time in English and Japanese when the first movements were concerned ($t(25.6)=-1.35$, $P>0.05$) . Nevertheless, the story was different when we concern about the second movements. The mean duration of the second movements in English was 1918 msec. (SD=133), where it was 1257 msec. (SD=119) in Japanese (See Figure 2). In contrast to the first movements, the paired t-test demonstrated significant difference in both orientation latencies ($t(10)= 5.43$, $P =0.0001$). (Both RT to English and to Japanese showed normal distribution (See Appendix 5.)

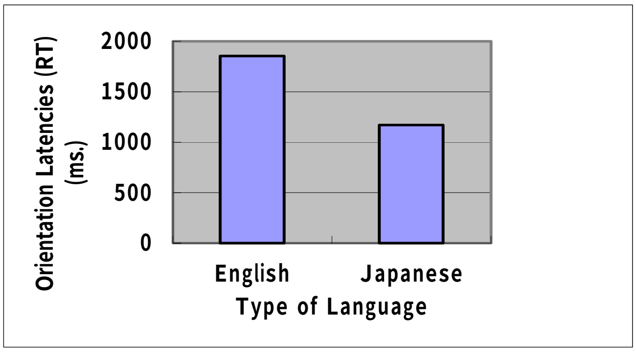

Fig. 2. Mean orientation latencies of Japanese infants to English and Japanese spoken by a bilingual woman (level 9.5)

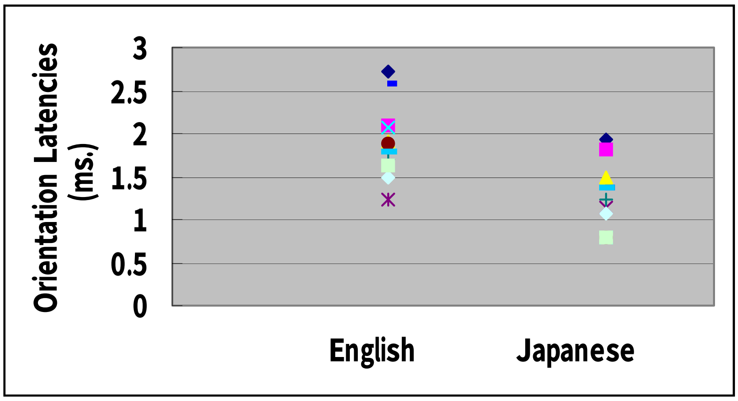

Fig. 3. Individual differences in RT to English and Japanese spoken by a bilingual woman (level 9.5)

2.3.2. Listening time

58% of valuable response to English and 58% of it to Japanese was seen including wrong localization (32%) as I have described.

Both RT to English and to Japanese showed normal distribution (See Appendix 5).The mean duration of each listening time between the second movements and moments of deviations in English was 2204 msec. (SD=251), where it was 3510 msec. (SD=290) in Japanese (See Figure 4). Paired t-test demonstrated significant difference in both orientation latencies (t(10)= -6.93, P <0.01).

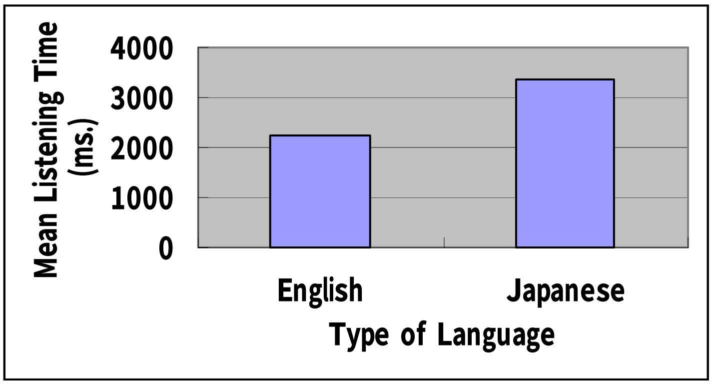

Fig. 4. Mean listening time of Japanese infants to English and Japanese spoken by a bilingual woman (level 9.5)

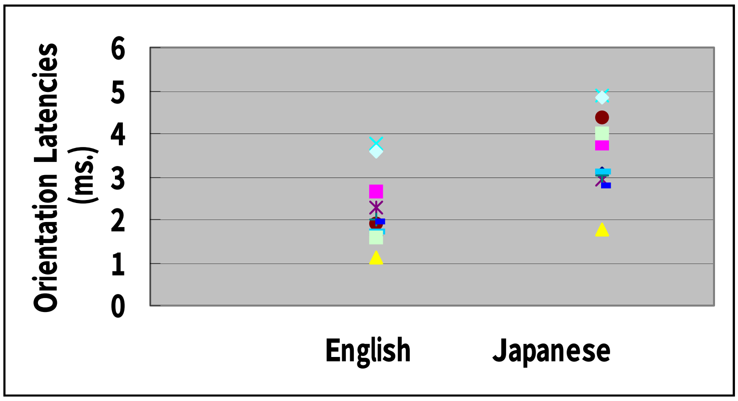

Fig. 5. Individual differences in RT to English and Japanese spoken by a bilingual woman (level 9.5)

2.4. Discussion

The first movements did not present any significance in Reaction Time whereas the second did. Though the first movements occurred much very often, the t-test verified that infants reacted at first just toward the sound, not to a language, as had been claimed in 1997's paper.

Due to the second movements, Japanese infants in monolingual environment showed a significant discrimination capacity between native English and Japanese, with slower reaction to the former and quicker to the latter, as well as shorter attention to the former and longer to the latter.

The Data indicate that Japanese four-month-olds were abele to distinguish English from Japanese, a pair of rhythmically different languages. This result corresponds to Nazzi et. al. (1998). Hence, the procedure can be used to test if Japanese infants discriminate English from Japanese spoken by a monolingual Japanese in level 1.

3. Experiment 2

3.1. Purpose

Based on the validity of both methods, we decided to run the second experiment that assesses the research question: if Japanese four-month-old monolingual infants can differentiate English from Japanese that spoken by a Japanese woman who has never been to English speaking countries.

3.2. Method

3.2.1. Subject:

Ten four-month-old Japanese infants from experiment 1 were tested. One infant fell in sleep between the experiments and did not been assessed. The subjects were between 136 to 152 days at the recording time, and did not have auditory problem. Their mother tongue was Japanese, for their parents spoke to them only in Japanese from the birth. They have hardly heard native English (probably shortly in a park or train) in front of them, and have heard English though audio visual aids not very much. (A mother had watched American movies near her awaken baby. Two mothers presented their babies a world-language video called “Baby Einstein” up to the third months occasionally. Presentation of English in this video was not very long. Two mothers had listened to English songs for children at home or in the car).

Two additional subjects were also tested, but rejected due to fuzziness and failure of the camera.

3.2.2. Auditory Stimuli

A Japanese woman who has never been to an English speaking country was asked to read out English (n=21) and Japanese (n=21) sentences used in experiment 1. Before reading, she looked at a picture book of “The Snowman” and practiced more than three times approximately. She had to read each sentence in four seconds. We used a very quiet room this time.

After recoding, her English narrating were introduced to an American man and estimated as level 1 in ALC categorization (See Appendix 3) in the matter of pronunciation and fluency. We asked him a further comment on her Pronunciation and Reading Fluency following to Phone Pass guideline (See Appendix 4). According to the man, she “ pronounces in a non-native manner and reads not fluently”, “produces consonants, vowels, and stresses in a non-native manner”, “reads without English rhythm, phrasing and pausing”.

After three experiments with two blocks each (one block consisted of 7 sentences in each language), the stimuli were reedited in one block, since infants could not concentrate in the second block at all (They got tired even sooner than experiment 1). Ten sentences holding less syllables in each type of clause were finally picked up out of each language, because the monolingual woman spoke slower than the bilingual and they were the ones being always around four seconds. Mean duration of utterances in those selected sentences became therefore similar in each language, 3880 ms. in English (SD=86), 3600 ms. in Japanese (SD=100). One-factor ANOVA showed non-significant difference among four kinds of stimuli used in Experiment 1 and being used in Experiment 2 ($F(3/25)=1.31$, $P=0.29$).

Those 20 sentences, half level-1 English and the other Japanese, were edited with a sound editing software carefully. Each sentence was attached with silence of more or less four seconds, so that a phrase became just eight seconds approximately. Following to two German sentences, those twenty sentences together with silence in each were randomized under the same restrictions as Experiment 1. The image, German familiarization, and twenty randomized sentences in Japanese and English made a block for the experiment 2 (See Figure 1).

3.2.3. Apparatus

The apparatus was the same as Experiment 1.

3.2.4. Procedure

After the experiment 1, infants and mothers took either a walk or lunch outside, to avoid any influence and tiredness to the test.

The whole procedure took place in a single session, near 3.5 min, comprised familiarization phase, ten monolingually spoken English and ten Japanese sentences randomly edited. The design was same as Experiment 1.

The rational behind the procedure was that if the infants have listened to a familiar rhythmic representation only in Japanese sentences, they might show shorter reaction time as well as longer listening time than towards English.

3.2.5. Coding procedure

The procedure was same as Experiment 1

3.3. Result and Analysis

3.3.1 Reaction Time

76 % of valuable response to English and also 82 % of it to Japanese was seen including wrong localization (34 %) that agreed to other developmental works on this method.

Since the invalidity of the first movements were proved in Experiment 1, we neglected those in this time. Both RT to English and to Japanese showed normal distribution (See Appendix 5). The mean duration of each reaction time (the second direct saccadic movement) in English was 1715 msec. ($SD=198$), and in Japanese was 1495 msec ($SD=124$) (See Figure 6). Paired t -test revealed the insignificant difference in reaction time to both languages. ($t(9)=-0.9$, $P=0.39$). The analysis of variance (One-factor ANOVA) indicated a significant interaction between reaction time towards bilingual Japanese, monolingual Japanese, and monolingual English ($F(2/62)=2.29$, $P>0.05$).

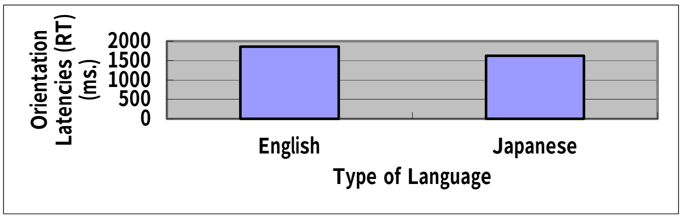

Fig. 6. Mean orientation latencies of Japanese infants to English and Japanese spoken by a monolingual woman (level 1)

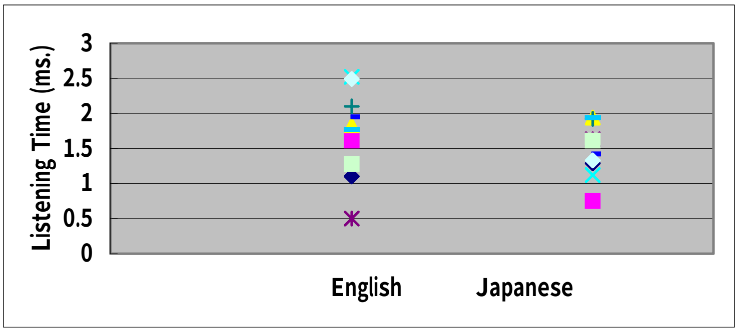

Figure 7. Individual differences in RT to English and Japanese spoken by a monolingual woman (level 1)

3.3.2 Listening time

Only 44 % of valuable response to English and also 46 % of it to Japanese was seen including wrong localization (34 %). Hence, among the data which were valuable for Reaction Time, about 58% of responses to English and 56% of those to Japanese were valid also for listening time. The low validity was caused by fussiness as well as that infants gazed at one of the speakers even short after the next stimulus came out. The latter case had to be rejected since we could not combine the listening duration to two different stimuli (restriction 3)).

Both RT to English and to Japanese showed normal distribution (See Appendix 5). The mean duration of each listening time between the second movements and the moments of deviation in English was 3356 msec. (SD=389), where it was 3644 msec. (SD=254) in Japanese (See Figure 8). Paired t-test demonstrated insignificant difference in both orientation latencies ($t(9)= 0.98$, $P =0.35$).

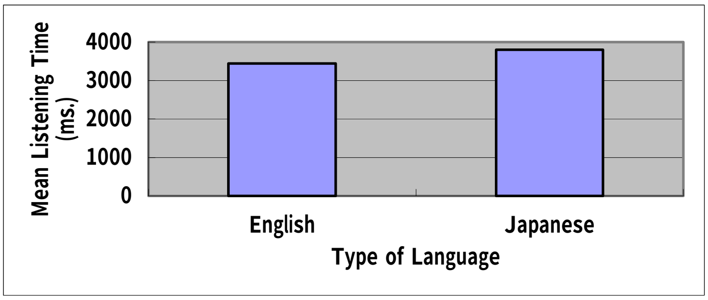

Fig. 8. Mean listening time of Japanese infants to English and Japanese spoken by a monolingual woman (level 1)

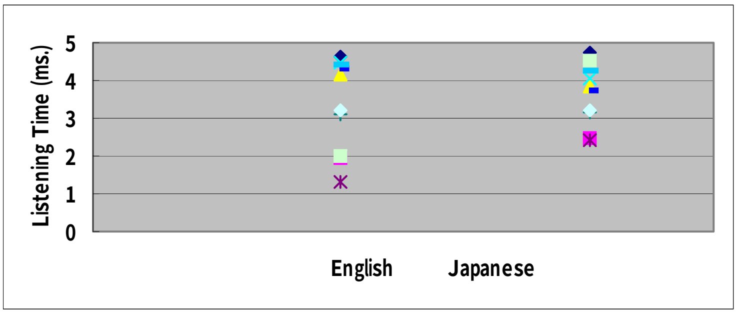

Fig. 9. Individual differences in RT to English and Japanese spoken by a monolingual woman (level 1)

3.4. Discussion

In the first experiment, infants significantly discriminated English from Japanese spoken by bilingual woman (level 9.5). On the other hand, this couldn't be the case in the second experiment that they could not significantly recognize English spoken by a monolingual female Japanese (level 1).

Both Reaction Time and listening time did not show any significant difference when the stimuli were presented. Those results indicate that four-month-old infants do not recognize English spoken by a Japanese monolingual woman in the first level. We have to warn parents (and teachers, besides) reckless use of language without training for proper pronunciation at home (and school).

In the other hand, many data of very long listening time both to English and to Japanese were rejected according to the restriction 3). If we had shown images among stimuli in order to take infants attention back, these endless listening attitudes could be made shorter, be adopted, make the mean duration of listening time longer, and lead another kind of result. We could see here clearly the necessity of presenting images among auditory stimuli as Bosch and his colleague did in 2001.

4. Conclusion

The study mostly followed the method of two consequent works conducted by Bosch & Sebastian-Galles (1997&2001). First of all, it has been shown that four-month-old Japanese infants could discriminate between natively spoken English and Japanese. Second, the same infants could not differentiate English from Japanese spoken by a Japanese monolingual woman.

This result, however, is nothing to reject the usage of two languages in a family, as long as they are pronounced in a higher level and one person uses only a language. When two distinguishable languages are presented by two different persons, either SUP or CUP (Cummins, 1981) might be constructed in one's brain. The bilingual-to-be infants under those environment will utter a socially, linguistically, psychologically preferable language to transmit the wish as they starts speaking (Dodson, 1985). As one can presume from the stimulative title of Grosjean (1989, "Neurolinguists, Beware! The bilingual is not two monolinguals in one person."), "the historical interpretation against bilingualism has been influenced by the relationship between social environment and bilingualism" in each era (Hoffmann, 1991). As Hoffmann suggests, a number of researchers are not against bilingualism for children nowadays, even though "smaller vocabulary and easier construction in sentences are remarkable until the beginning of school age." (Largo, 1995)

Though we have gained clear results from two kinds of measurements, we do not know exactly why the infants could not discriminate monolingual's English from Japanese. As I described, Bosch & Sebastian-Galles assumed additional vowels in Catalan language helped infants to distinguish it from Spanish. But our case does not come under theirs. We also have not discovered which level of successive utterance can be recognized as English by Japanese infants. Do they discriminate English spoken by Japanese English teachers? If not, don't we need to reconsider the educational system for English education? Both the reason of failure and the recognizable Level(s) should be investigated in the future.

Appendix 1: Questionnaire

1. あなたのお子さんの生年月日をお書きください。
2. あなたのお子さんは、音や声に、反応しますか（顔をそちらに向けるなど）？
3. あなたのお子さんの母国語をお書きください。
4. 母国語以外にご両親が話し掛ける言葉があればお書きください。またそれは、どの位の頻度ですか？
5. あなたのお子さんが、生の英語に触れる機会は今までありましたか？それはどの位の時間ですか？
6. あなたのお子さんが、テレビや CD の英語に触れる機会は今までありましたか？それはどの位の時間ですか？

Appendix 2:

1. Stimuli narrated by a bilingual woman

(S= single sentence, C=coordinate clause, MS= main clause + sub coordinate clause)

1.1. English sentences		
-To his surprise, the snowman greeted him po/lite/ly.	12	S
-Next morning, the boy saw only the scarf in the garden.	14	S
-The two went upstairs into the litt/le boy's parents' bedroom	15	S
-As he brushed his teeth, he kept his eyes on the snowman.	13	MS
-Since he couldn't fall in sleep, he looked out of the window.	13	MS
-Though the snowman wanted, he was afraid of the sunrise.	14	MS
-He stayed in there until he became cool enough to come out	15	MS
-He was frightened by a lamp as well as by a ra/di/a/tor.	16	C
-He went downstairs, put on a pair of boots, then went into the garden	17	C
-Because the container was full of ice, the snowman got fully happy	18	C
1.2. Japanese sentences		
-でんきを　つけたり　けしたりするのが　おもしろくて　たまりません	28	S
-おどろいたことに　じょそうのあと　ふたりは　そらを　とんでいました	28	S
-ゆきのなかを　とんだって　ふしぎと　すこしも　さむくありません	27	MS
-おとこのこが　にわに　みたものは　みどりの　えりまき　だけでした	27	MS
-はを　みがいている　あいだじゅう　きになって　しかた　ありません	27	MS
-ひのでを　みたいと　いっていたのに　なぜかとても　こわがりました	28	MS
-テレビも　みるきに　ならないし　たべるきだって　しないのです	26	C
-おとこのこは　ながぐつに　あしをつっこみ　にわへ　とびだしました	28	C
-だいすきな　れいとうこに　きがついて　あわててくるまを　おりました	29	C
-じゅうぶん　つめたい　からだになって　どっこいしょと　でてきました	27	C

2. Stimuli narrated by a Japanese monolingual woman (level 1)

2.1. English sentences		
-He drew a big smile on the Snowman's face.	10	S
-He enjoyed switching on and off the light	10	S
-After an approach run, they started fly/ing.	11	S
-They had dinner before they washed the dishes	10	MS
-As he woke up, he saw it was snowing.	10	MS
-When they found a freezer, he went out of the car.	12	MS
-The snowman was so big that it took him all day long.	13	MS
-They flew long in the snow, but it was/n't cold at all.	13	C
-But he could/n't watch TV and he could/n't eat.	12	C
-The boy shook his hand and asked him to come into the house.	14	C
2.2. Japanese sentences		
-あったかい　らんぷや　だんぼうを　とても　こわがります	23	S
-なんと　ゆきだるまが　ゆっくり　こちらを　ふりむいたのです	25	S
-こおりを　たべる　ゆきだるまは　ほんとうに　うれしそうです	25	S
-うえの　しんしつで　ぱぱと　ままが　ぐっすり　ねむっています	25	S
-あさおきると　まどのそとは　まっしろなゆきでした	22	MS
-こおりで　いっぱいだったので　おおよろこびの　ゆきだるま	25	MS
-とても　おおきく　したので　つくるのに　いちにち　かかりました	26	MS
-ゆうしょくをたべ　そして　ふたりで　おさらをあらいます	23	C
-てをとって　あくしゅをし　いえのなかに　あんないしました	24	C

-はなをつけ　おおきな　くちを　ゆびでかくと　さあ　できあがり	25	C
3. German stimuli used as familialization		
-Wilkomen Kind. Ich wuensche dir eine shoene Zeit	13	
-Und ich hoffe, dass du mir ein bischen zuhoeren kannst.	14	

Appendix 3-1:
Excerpt of assessment criteria of SST (Standard Speaking Test) by ALC (Level 1-10)

Level	Pronunciation	Fluency
9	Candidate speak almost in native manner.	Fluency is like native speakers
8	Pronunciation is quite native-like, but still influenced by Japanese slightly. Candidates have difficulty to produce certain consonants such as l, r, th, v.	Speech is almost fluent and not stagnated.
7 6	Pronunciation is still influenced by Japanese, but produce some native-like words at the same time.	Speech become sometimes slowly, but in other time not.
5 4	Pronunciation is much influenced by Japanese. It sometimes hinders the listener from understanding, but mostly understandable.	Candidates speak sometimes without stagnation though the speech is slow..
3	Pronunciation is like Japanese. It sometimes hinders the listener from understanding.	Candidates repeat the words and make long pauses sometimes.
2	Pronunciation is strongly influenced by Japanese. It hinders listener from understanding quite often.	Candidates repeat the words and make long pauses quite often.
1	Pronunciation is strongly influenced by Japanese. It often hinders listener from understanding.	Candidates repeat the words and make long pauses often.

A 3-2 http://www.alc.co.jp/edusys/ssteng2.htm

Appendix 3-2: The correlation between ACTFL OPI and SST.

Levels	SST Levels
Superior	Level 9
Advanced High	
Advanced Mid	
Advanced Low	
Intermediate High	Level 8
Intermediate Mid	Level 7
	Level 6
Intermediate Low	Level 5
	Level 4
Novice High	Level 3
Novice Mid	Level 2
Novice Low	Level 1

Appendix 4: Excerpt of assessment criteria of Phone Pass
-Pronunciation
Pronunciation is the ability to produce consonants, vowels and stress in a native-like manner in sentence context; candidate performance depends on knowledge of the phonological structure of everyday words.
-Reading Fluency
Reading Fluency is the ability to read aloud using appropriate rhythm, phrasing and timing.
ppass@odinate.com

Appendix 5: Normal distribution

	RT to bilingually spoken English	RT to bilingually Spoken Japanese	Listening time to bilingually spoken English	Listening time to bilingually spoken Japanese	RT to monolingually spoken English	RT to monolingually spoken Japanese	Listening time to monolingually spoken English	Listening time to monolingually spoken Japanese
X2	2.52	1.24	2.52	0.92	0.48	2.16	2.16	0.9
P	0.11	0.26	0.11	0.34	0.49	0.14	0.14	0.34

Appendix 6: Individual data

RT to bilingually spoken English	RT to bilingually Spoken Japanese	Listening time to bilingually spoken English	Listening time to bilingually spoken Japanese	RT to monolingually spoken English	RT to monolingually spoken Japanese	Listening time to monolingually spoken English	Listening time to monolingually spoken Japanese
2.72	1.93	1.63	3.09	1.1	1.29	4.62	4.74
2.09	1.82	2.66	3.75	1.61	0.75	1.93	2.48
1.9	1.49	1.13	1.79	1.83	1.94	4.15	3.85
2.07	0.78	3.78	4.9	2.52	1.12	4.48	4.04
1.24	1.15	2.3	2.91	0.5	1.63	1.32	2.44
1.89	0.8	1.92	4.39				
1.69	1.24	1.98	3.04	2.1	1.92	3.13	3.16
2.59	1.38	1.96	2.78	1.95	1.42	4.3	3.75
1.8	1.38	1.72	3.1	1.77	1.94	4.43	4.25
1.49	1.07	3.59	4.86	2.49	1.33	3.19	3.19
1.62	0.79	1.58	4	1.28	1.61	2.01	4.54

Reference

Bahrick, L., E., Pickens, J.N., 1988. **Classification of bimodal English and Spanish language passages by infants.** Infant Behavior and Development 11, 277-296.

Bosch, L., Sebastian-Galles, N., 1997. **Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments.** Cognition 65, 33-69.

Bosch, L., Sebastian-Galles, N., 2001. **Evidence of Early language Discrimination Abiliies in Inafants from bilingual environments.** Infancy 2(1), 29-49.

Culter, A., Mehler, J., Norris, D., Segui, J., 1986. **The syllable's differing role in the segmentation of French and English.** Journal of Memory and Language, 25, 385-400

Culter, A., 1990. **Exploiting prosodic probabilities in speech segmentation.** In G.T.M. Altman (Eds.), Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives (pp. 105-121). Cambridge: MIT Press

Culter, A., Butterfield, S., 1992. **Rhythmic cues to speech segmentation: Evidence from juncture misperception.** Journal of Memory and Language, 31, 218-236

Cummins, J. 1981. **The role of primary language development in promoting educational success for language minority students.** In Office of Bilingual Bicultural Education (Eds.), Schooling and Language minority education: a theoretical framework. Sacramento (pp. 3-49). State Department of Education

Dodson 1985. **Second Language Acquisition and Bilingual Development: A Theoretical Framework.** Multilingual & Multicultural Development, 325, 326-327

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J., 1971. **Speech perception in infants.** Science, 171, 303-306

Ervin-Tripp, S., 1970 b. **Structure and process in language acquisition.** Monograph Series on Language and Linguistics. 23. 313-344

Grosjean, F., 1989. **Neurolinguists, Beware! The bilingual is not two monolinguals in one person.** Brain and Language, 36, 3-15

Hoffmann, C., 1991. Sociocultural Aspects of bilingualism: an introduction to bilingualism.(pp.136-154). Longman, London and New York

Ronjat, J., 1913. Le développement du langage observé chez un enfant bilingue: Librairie ancienne H. Champion. Paris.

Jusczyk, P. W., Aslin, N. R., 1995. **Infants' detection of the sound patterns of words in fluent speech.** Cognitive Psychology 29, 1-23

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., Lindblom, B., 1992. **Linguistic experiences alter phonetic perception in infants by 6 months of age**. Science, 255, 606-608

Largo, H. R., 1995. Babyjahre: Die fruehkindliche Entwicklung aus biologischer Sicht. (pp. 351) Piper Muenchen Zuerich

Leibermann, A., Sohmer, H., Szabo, G., 1973. **Cochlear audiometry (electrocochleography) during the neonatal period.** Developmental Medicine and Child Neurology 8-13

Lieberman, P., 1996. **Some biological constrains on the analysis of prosody.** In J.L.Morgan & K. Demuth (Eds.), Signal to syntax: Bootstrapping form speech to grammar in early acquisition (pp.55-65). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Mc Laughlin, B., 1984. **Second-Language Acquisition in Childhood: Volume1. preschool Children**. Lawrence Erlbaum Associates, Publishers (pp. 24) London

Mehler, J., Dommergues, J,Y., Frauenfelder, U., Segui, J., 1981. **The syllable's role in speech segmentation.** Journal of Vebal Learning and Bebal Bejavior, 20, 298-305

Mehler, J., Jusczyk. P. W., Lambertz, G., Halsted, N., Bertoncini, J., Amiel-Tison, C., 1988. **A precursor of language acquisition in young infants.** Cognition, 29, 144-178

Mehler, J., Bertoncini, J., Dupoux, E., and Pallier, C., 1995. **The role of suprasegmentals in speech perception and acquisition**. Dokkyo International Review, 7, 343-376. Tokyo

Mehler, J., Sebastian-Gallés, N., Altmann, G., Dupoux, E., Christophe, A., Pallier, C. 1993. **Understanding compressed sentences: the role of rhythm and meaning**. In A. M. G. Paula Tallal, Rodolfo R. Llinas, Curt von Euler (Ed.), Temporal information processing in the nervous system: Special reference to dyslexia and dysphasia. (pp. 272-282) Annals of the New York Academy of Sciences.

Mehler, J., Dupoux, E., Nazzi, T.,Dehaene-Lambertz, G., 1996. **Coping with linguistic diversity: the infant's viewpoint.** In:Morgan, J.L., Demuth, K., Signal to Syntax. L.E.A. Publ., Mahwah, (pp.101-116). New Jersey

Miwa, J., 1999, **Mora and Syllable.** www@sp.cis.iwate-u.ac.jp?subject=Mora

Moon, C., Cooper, R., Fifer, W., 1993. **2-day-old prefere their native language.** Infant Behavior and Development, 16, 495-500.

Nazzi, T., Bertoncini, J., (1997) **Language discrimination by newborns: towards an undestanding of the role of rhythm.** Journal of Experimental Psychology: Human Perception and Performance (in press).

Nazzi, T., Bertoncini, J., & Mehler, J., 1998. **Language discrimination by newborns**: **Toward an understanding of the role of rhythm.** Journal of Experimental Psychology: Human Perception and Performance, 24, 756-766

Nespor, M.,Vogel, I., 1986. Prosodic Phonology**.** Foris Publ., Dordrecht.

Otake, T., Hatano, G., Culter, A.,Mehler, J.,1993. **Mora or Syllable? Speech segmentation in Japanese.** Journal of Memory and Language, 32, 358-378.

Otake, T., Yoneyama, K., 1999. **Listeners' representations of within-word structure: Japanese preschool children.** Proceedings of the 13th International Congress of PhoneticSciences, (pp. 2193-2196). San Francisco, 3,

Otake, T., Yamamoto, K., 2001. **Phonological awareness by monolingual and bilingual speaking of Japanese and English.** Journal of the Phonetic Society of Japan, 5, 107-116

Traiman, R. 1992. **The role of intrasyllabic units in learning to read and spell.** In P.B.Gough, L.Ehri, and Treiman (Eds.)Reading Acquisition, Hillside, NJ:Erlbaum

Weker, J. F., and Lalonde , C. E., 1988. **Cross-language speech perception: Initial capabilities and developmental change.** Developmental Psychology, 24, 672-683