音声研究　第6巻第2号
2002(平成14)年8月
66-74頁

Journal of the Phonetic Society
of Japan, Vol.6 No.2
August 2002, pp.66–74

The Possible Preferential Cues of Infants' Response toward Their Native Dialects Evidenced by a Behavioral Experiment and Acoustical Analysis

Ryoko MUGITANI*, Akiko HAYASHI** and Shigeru KIRITANI***

乳児の養育環境にある方言音声選好の手がかりとなる音響特性の検討
―乳児行動実験及び音響分析を用いて―

麦谷　綾子*・林　安紀子**・桐谷　滋***

要旨：音声知覚能力の発達は言語獲得を可能にする最も重要な要素のひとつである。乳児の音声知覚は生後1年間で飛躍的に発達し，母語の音声言語体系に適合した様式になる。故に，前言語期の乳児の音声知覚能力に関する知見を得ることは，人間の言語獲得過程をあきらかにする上で不可欠である。麦谷(2000)の報告において，乳児は生後8ヶ月で自分の養育環境にある関東方言を選好することが示された。この知見を踏まえ，本研究では乳児が方言音声への選好を示す際に手がかりとしている音声特徴を，乳児行動実験と音響分析を用いて検討することを目的とした。乳児行動実験から方言音声の韻律的特徴には乳児が選好を示すために十分な情報量が含まれていること，また音響分析からいくつかの音響特徴において方言間での違いが見出され，乳児の方言音声選好の手がかりとなっている可能性が示唆された。

Key words: Infants, Speech perception, Native dialect, Preferential cues, Development

1. Introduction

Language is a cognitive function that is unique to human beings. In order to acquire language, the ability to segment continuous speech into syllables, words and phonemes is essential. More over, the ability to discriminate and identify these units is also required. The acoustic characteristics of phonemes change by articulation co-ordination and assimilation. Therefore, in order to understand spoken language, not only phonemes but also prosodic cues such as accent, rhythm, and intonation must be used effectively in the processing of speech signals. Human vocal tracts differ depending on sex and age, with the acoustic characteristics of speech also depending on the speaker. In order to understand the language information contained in speech, the ability to standardize the various acoustic characteristics of different speakers is important.

The development of speech perception is so rapid that each of the abilities described above is acquired during the first year of an infant's life (Jusczyk, 1997). Therefore, studying the development of speech perception during the first year of life is indispensable to understanding the most primitive and essential processes of language acquisition.

Recent research of infant speech perception has revealed that infants are very sensitive to the speech stream surrounding them. The development of perception of speech surrounding infants has been studied by comparing the preferential response of the native language to a foreign language. Mehler et al. (1988) studied infants' preference to the native language and revealed that French 4-day-old infants prefer

* Doctor's Course, Graduate School of Arts and Sciences, University of Tokyo. Research Fellow of the Japan Society for the Promotion of Science（東京大学大学院総合文化研究科博士課程・日本学術振興会特別研究員）
** Associate Professor, Research Institute for the Education of Exceptional Children, Tokyo Gakugei University（東京学芸大学附属特殊教育研究施設助教授）
*** Professor, Department of Literature, Kobe Kaisei College（神戸海星女子学院大学文学部教授）

French to Russian. Moon et al. (1993) found that American 2-day-olds prefer English to Spanish. Hayashi et al. (1996) studied Japanese 4 to 11-month-old infants and reported that these infants showed preference to Japanese over English 200 days from birth.

Dialects are also a form of particular speech stimuli surrounding infants. A "dialect" is defined to be a local variant of a language (Sanada, 1988). Dialects can be considered as intra-language modifications of a single base language, and hence inter-language acoustic differences, which are usually assumed to be large and robust, could be smaller between dialects. Despite this slight difference, infants acquire their own local dialect. Thus, it is speculated that there should be a time when infants start discriminating amongst dialects, and show a preference to their own, just as infants prefer to listen to their native language and acquire it.

In order to verify this hypothesis, Mugitani et al. (2000) studied 67, 5 to 8 month old infants who had been born and raised in the environment of the Japanese Eastern dialect. This research demonstrated that those infants preferred to listen to the Eastern Dialect (native dialect) to the Western Dialect (foreign dialect) at 8 months of age.

Which phonetic features made those infants identify and prefer their native dialect? In the above research, the perceptual experiments by adult subjects indicated that Japanese Eastern dialect speaking adults could discriminate the Eastern and Western dialects by using only the prosodic information of stimuli. In addition, infants are very sensitive to prosodic information from the beginning of their external life. Using only prosodic information as a discriminatory cue, it was reported that infants can differentiate between native and foreign languages (Mehler et al., 1988), mother's and unknown woman's voice (DeCasper et al, 1980), and infant directed and adult directed speech (Fernald, 1985). Also the preference that infants have to listening to familiar sounds such as native language (Mehler et al, 1988) and their mother's voice (DeCasper et al, 1980) can be determined by that sound's prosodic features. This sensitivity of infants and the evidence of adults' ability to discriminate stimulated the following research into the effectiveness of prosodic information of dialect as a cue of preferential response. We conducted the following behavioral experiment in order to ascertain whether infants still show the preferential response toward the native dialect when only prosodic information is available.

On the other hand, as former research indicates, at around 8 months of age when the preferential response toward the native dialect emerges, infants start to change their speech perception style from a global and universal manner to a more unique manner specified to the phonetic system of their native language. According to Mugitani's study, the preference of native dialect emerged at the same time as the specification of speech perception to the native language. Hence, it is speculated that development of the dialect preference is dependent on detailed acoustic information processing, which occurs after 8 months of age. In order to understand detailed acoustic features, which might be used by 8-month-old infants as the cues of preference, we execute acoustic analysis of the stimuli used by Mugitani et al. (2000) and identify the acoustic differences between the two dialects' stimuli.

2. Behavioral experiment

2.1. Creation of Stimuli

The voice recordings were made in a sound-proof booth. The speaker was a 22 years old woman fluent in both the Japanese Eastern and the Western dialects. She grew up in Tokyo until the age of 5 years, then moved and lived in Kyoto until the age of 19 years. At that time, she communicated in the Eastern dialect with her family members and in the Western dialect outside the household. She had lived in the United States for 1 year when she was 17 years old. She moved back to Tokyo, and has lived in Tokyo from the time when she was 19 years old up to the time of recording.

The speaker saw pictures in a book, and responded vocally as if she were explaining the

contents to an infant for about 3 min. Eight pictures were presented to the speaker and each picture was explained in both dialects. Thus, a total of 16 speech utterances (8 utterances in the Eastern dialect and 8 utterances in the Western dialect) were recorded on DAT tape (48KHz sampling rate). Consecutive speech samples lasting about 20 sec respectively were extracted from the 16 utterances. Thus, a total of 16 speech samples were made.

In the behavioral experiment, 2 samples of each dialect (a total of 4 samples) were presented to infants in the practice phase, and the remaining 6 samples of each dialect (a total of 12 samples) were presented in the experimental phase. The mean duration of the 6 samples of each dialect presented in the experimental phase was 21.29 sec in the Eastern dialect and 21.07 sec in the Western dialect.

All stimuli were low-pass filtered at 400Hz. In this way, segmental information of speech was removed though prosodic information was retained.

2.2. Methods

2.2.1 Participants in the experiment

46, 7 to 8-month-old infants (28 males and 18 females) from families that spoke only the Japanese Eastern dialect, participated in the experiment. The infant's average age was 238.0 days (SD=18.37). The definition of "speaking only the Eastern dialect" is; 1) The infant has had almost no opportunity to hear the Western dialect between birth and the time when the experiment was conducted. 2) All of the family members have spoken to the infant using only the Eastern dialect. 3) An interviewer identified the speech of the mother at the interview as typical Eastern dialect. In addition, if the mother came from an area where a different dialect from the Eastern dialect is spoken, 4) The mother had lived in Tokyo for more than 5 years.

16 infants were excluded for crying, not paying attention to the stimuli, and experimental errors. Therefore, 30 infants (19 males and 11 females) were subjects of final analysis. The average age of participants was 293.3 days (SD=17.44).

2.2.2. Procedure and Apparatus

Headturn Preference Procedure (HPP) (Fernald, 1985; Kemler-Nelson et al, 1995) was employed. In HPP, two types of materials (ex. The Japanese Eastern and Western dialect) are compared. While infants watch the target (flashing red light), a speech sample is presented, and the watching time is measured as the length of time during which each sample is listened to. The average listening time is compared between the two types of materials. A significant difference between two types of materials indicates that infants discriminate between them, and also prefer to listen to the one with longer listening time.

The experiment was conducted in a three-sided test booth (Fig1) made of pegboard panels (120 × 180cm). Blue curtains were hung between the ceiling and the top of the three sides of the booth in order to block the infant's view of the rest of the room and also to make the booth symmetrical. Loudspeakers were mounted beyond the walls of the two side panels at a height roughly level with that of the infant's head. A red light was located on each side panel in the vicinity of the loudspeaker. The center panel, which the infant faced, had a green light mounted at the infant's eye level. Directly below the light there was a hole in the panel allowing for the insertion of the video camera lens. Lighting of the booth was so dim that the infant would be able to detect the flash of lights easily. An

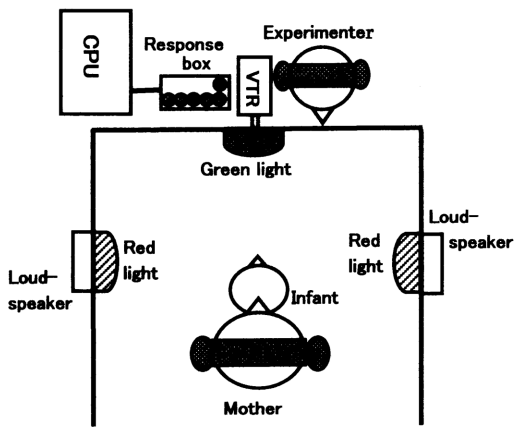

Fig. 1. Sketch of the test booth

experimenter hidden behind the center panel looked through a peephole of the pegboard and recorded the direction and duration of the infant's headturns using a response box connected to the PC.

The experiment was made up of a practice phase and a test phase. The practice phase consisted of four practice trials and the test phase consisted of 12 test trials. A total of 16 trials were executed consecutively. The practice phase aimed to accustom the infant to the flashing light and to train the infant to focus on it. Each infant was held on the lap of his/her mother who was seated on a chair in the center of the test booth. Each trial began by allowing the green light on the center panel to blink until the infant oriented in that direction. Then, the center light was extinguished and the red light above the loudspeaker on one of the side panels began to flash. When the infant turned and focused on the red light, the speech sample for that trial began to play. The speech sample continued until the infant had listened to it completely, or if the infant turned away from the target for more than 2 consecutive seconds, both the speech and blinking of the red light were terminated and the next trial started.

The order of presentation of the 12 test trials was randomized for each infant. The side of the loudspeaker from which the stimuli were presented was randomly varied from trial to trial.

In the test phase, mothers were instructed not to interfere with the infant's response by pointing or talking. As the experimenter and the infant's caregiver wore earphones and listened to masking music, they were unaware of the ordering of the test samples.

2.3. Results and Discussion

The shortest length of the 12 speech samples used in the test phase was 17.7 sec. Thus, even if the listening time of a sample exceeded 17.7 sec, that trial was treated as having a length of 17.7 sec even.

The participants were divided into a 7-month-old group and an 8-month-old group. The 7-month-old group included 15 infants (10 male and 5 females), and the average age was 224.6 days (SD=7.96). The 8-month-old group included 15 infants (9 males and 6 females), and the average age was 254.1days (SD=10.1).

A one-way ANOVA with the main within-subject effect of presentation order revealed a significant main effect(7-month-old: $F(11,154) = 6.56$, $p<.01$, 8-month-olds: $F(11,154) = 4.78$, $p<.01$). Tukey's HSD test for a multiple comparison found that the 7-month-old group listened to speech samples of the first trial significantly longer than those of all other trials, and the 8-month-old group listened to speech samples of the first trial significantly longer than those of the third to 12th trials. Thus, the listening time of the speech samples of the first and second trials were excluded from further analysis.

Mean listening times to the two languages in the test phase were calculated for each infant. Across the 7-month-old group, the average listening times were 6.79 sec (SD=2.81) for the samples of the Eastern dialect, and 6.53 sec (SD=2.77) for those of the Western dialect. Across the 8-month-old infants, the average listening times were 7.17 sec (SD=2.36) for the samples of the Eastern dialect, and 5.94 sec (SD=2.15) for those of the Western dialect (Fig 2). A two-way ANOVA with the main within-subjects factor of dialect and main within-subject factor of direction of lights was conducted in each age group. As a result, the significant main effect was observed only in the main factor of

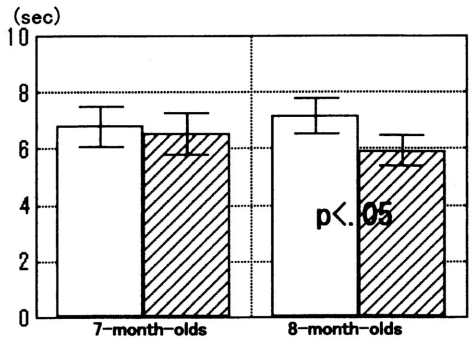

Fig. 2. Mean listening time of each dialect
The Japanese Eastern Dialect
The Japanese Western Dialect
Standard Error

dialect in the 8-month-old group ($F(1,14) = 5.86$, $p<.05$). There was no effect of the main factor of direction of lights ($F(1,14) = 0.97$, n.s.) and no interaction between the two factors ($F(1,14) = 1.00$, n.s.) in the 8-month-old group. Also, there was no effect of the main factor of dialect ($F(1,14) = 0.07$, n.s.) and direction of lights ($F(1,14) = 1.46$, n.s.) and no interaction between the two factors ($F(1,14) = 0.38$, n.s.) in the 7-month-old group.

This result indicated that the 8-month-old infants preferred to listen to the Japanese Eastern dialect as opposed to the Japanese Western dialect using prosodic information. Thus, even when most information issuing from other types of cues is stripped away by a low-pass filter, the remaining prosodic information contains sufficient cues to allow the infants to distinguish and prefer the native dialect. This suggestion fits well with previous observations regarding infants' preference for certain speech. For example, Fernald et al. (1985) have shown that prosodic factors are an important determinant in the preference that older infants display for infant-directed speech over adult-directed speech. Mehler et al. (1988) demonstrate that infants can discriminate their native language from foreign languages even when the speech was low-pass filtered. Similarly, it has been shown that certain discrimination results cannot be obtained under conditions where the natural flow of prosody is disrupted (Mehler et al. 1978).

3. Acoustic analysis

3.1. Procedure

Acoustic analysis was carried out in order to ascertain the characteristics of the 12 speech stimuli used in the experimental phase. Utterance duration, pause duration, speech rate, and pitch were analyzed as prosodic features. These parameters were the same as the previous research comparing the prosodic features of infant-directed speech in different languages (Fernald et al 1984). In addition to these four parameters, accent pattern was studied because the Eastern dialect and the Western dialect contrast sharply with each other in accent pattern (Kindaichi, 1975).

Vowel duration, consonant duration, duration of closure and devoiced vowels were analyzed as segmental features. It has been indicated that vowels are stronger and consonants are weaker in the Western dialect (Satoh, 1977; Hirayama, 1983). Therefore, it is assumed that vowel duration, consonant duration, and duration of closure may be different between two dialects. Since there is a notion that high vowels between voiceless consonants are often devoiced in the Eastern dialect but not in the Western dialect (Sugito, 1996), the proportion of devoiced vowels was also studied.

The speech samples were processed on a Roku-Bun-Ken speech analyzer. A pause was defined as a period of 300 msec or more of silence. An "utterance" was defined as a section of speech existing between pauses (Fernald et al. 1984). 40 utterances and 49 utterances were identified in the Eastern and the Western dialect samples respectively. Speech rate was measured by dividing the utterance duration by the morae number in the utterance. Pitch was extracted (temporal window = 5 msec) in 14164 observation points in the Eastern dialect, and in 13785 observation points in the Western dialect. 26 pairs of the same words were extracted from the 6 pairs of speech stimuli, and vowel duration, consonant duration, and closure duration before stops and affricates were measured. Excluding repeated pairs of words, 20 pairs of accent patterns were examined. The rate of devoiced high vowels was measured out of 24 vowels of the Eastern dialect and 16 vowels of the Western dialect under the condition that the target vowels were either between voiceless consonants, or in phrase final position with preceding voiceless consonant.

3.2. Result and Discussion

The data were analyzed and compared between dialects. Table 1. shows the means and SD of each parameter of each dialect. Significant differences were observed in several parameters.

1) Variance ($F(39,48) = 2.482$, $p<.01$) and the mean ($t(87) = 2.624$, $p<.05$) of utterance dura-

Table 1. Mean and SD of each parameter of each dialect

	Eastern Dialect			Western Dialect		
	n	Mean	SD	n	Mean	SD
Uttarance duration(msec)	40	2371.3	1226.2	49	1810.0	778.2
Pause duration(msec)	34	662.0	224.0	43	674.3	234.1
Speech rate(msec/morae)	40	156.5	30.1	49	157.8	27.2
Pitch(Hz)	14164	256.9	64.4	13785	256.9	76.4
Vowel duration(msec)	74	94.5	48.4	74	96.0	46.9
Consonant duration(msec)	85	64.2	33.9	85	58.5	26.6
Duration of closure(msec)	19	96.6	55.3	19	82.1	51.1

tion of the Eastern dialect were significantly greater.

2) Variance of pitch of the Western dialect was significantly greater ($F(13784,14163) = 1.407$ $p<.01$).
3) Variance of consonant duration in the Eastern dialect was significantly greater ($F(84,84) = 1.626$, $p<.01$).
4) The proportion of devoiced vowels in the Eastern dialect (67%) was higher than that in the Western dialect (50%).

There were no significant differences between dialects in variance and the mean of speech rate ($F(39,48) = 1.227$, n.s $t(87) = .0222$, n.s.), pause duration ($F(33,42) = 21.092$, n.s. $t(75) = .0233$, n.s.), vowel duration ($F(73,73) = 1.063$, n.s. $t(146) = 0.19$, n.s.), closure duration ($F(18,18) = 1.17$, n.s. $t(36) = 0.832$, n.s.), nor in the mean of pitch ($t(26913.73) = .057$, n.s.) and consonant ($t(168) = 1.213$, n.s.) duration. Pairs of different accent patterns between dialects were dominant. The accent pattern was the same in seven pairs (35%) and different in 13 pairs (65%).

This result indicated that there are differences in utterance duration, pitch and consonant duration between the Japanese eastern dialect and the western dialect. Accent patterns were different in most of the words that were analyzed. Also, vowels in the Eastern dialect were more often devoiced. Even though physical acoustic characteristics do not always match exactly the characteristics of human's auditory perception, these differences may be the detailed information that is used effectively by 8-month-olds to identify the native Eastern dialect and show preference towards it.

4. General Discussion

The results of the behavioral experiment indicate that the participating infants can discriminate between the two dialects and show preference toward the native one using only the prosodic information of the speech. Therefore, prosodic cues may be important in allowing 8-month-old infants to identify utterances as belonging to the native dialect and to show preference towards it. Acoustic analysis indicates that there are several differences in their prosodic and segmental features between the Japanese Eastern dialect and the Western dialect.

Combining these findings, it can be speculated as to which precise aspects of the prosodic information infants use for preferring the Japanese Eastern dialect samples. The results of acoustic analysis of prosodic features suggested that variance of pitch is greater in the Western dialect samples, and the mean of utterance duration length and variance were also different between dialects. Accent patterns were also different in most of the words that were analyzed. Though it should be mentioned that physical acoustic characteristics do not always exactly match characteristics of human's auditory perception, these prosodic cues may be important in order to discriminate between the speech of two dialects and show preference to the native one. Further testing which focuses on each prosodic difference will tell us which prosodic cue is reasonable for the preferential response.

Moreover, there are differences in the segmental aspects of the stimuli. The acoustic analysis indicated that variance of consonant duration is larger in the Eastern dialect. More

devocalization took place in the Eastern dialect than in the Western dialect. Phrase-ending expressions are also often different between the dialects. For example, the phrase-ending expression "-dane" in the Eastern dialect is usually replaced by "-yana" in the Western dialect. These facts lead to a possibility that segmental differences including those found in the acoustic analysis might also become sufficient cues of preference. Behavioral experiments with synthesized stimuli, from which prosodic information is removed and segmental information is retained could offer further insights into this question. If preference to the native dialect were again elicited in this experimental condition, it would indicate that the segmental features can also be cues of infant's preference of a native dialect. On such an occasion, the segmental aspects uncovered by the acoustic analysis should be tested individually.

In the meantime, the present finding demonstrates that several acoustic features are suspected to be used as preferential cues by 8 month olds, and prosodic factors must be given serious consideration in accounting for the way in which infants prefer their native dialect.

5. Conclusion

This study aimed to investigate the possible cues for preferential listening toward their native dialect in 8 month old infants. Behavioral experiments demonstrated that prosodic information of the stimulus dialects contained sufficient information to elicit the preferential listening of the infants. Also, acoustical analysis identified several segmental and prosodic features that were different between the two stimulus dialects. There is a possibility that infants utilize these differences as cues of preferential response.

References

DeCasper, A. and W. P. Fifer (1980) "On human bonding: newborns prefer their mother's voices," *Science* 208, 1174–1176.

Fernald, A. and T. Simon (1984) "Expanded intonation contours in mother's speech to newborns," *Developmental psychology* 20(1), 104–113.

Fernald, A. (1985) "Four-month-old infants prefer to listen to motherese," *Infant Behavior and Development* 8, 181–195.

Hayashi, A., Deguchi, T. and S. Kiritani (1996) "Response Patterns to Speech Stimuli in the Headturn Preference Procedure for 4-to 11-month-old Infants," *The Japan journal of logopedics and phoniatrics* 37, 317–323.

Hirayama, T. (1983) *Kenbetsu Hougen no Tokusyoku.* Tokyo: Kadokawa Syoten.

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

Kemler-Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A. and L. A. Gerken (1995) "The headturn procedure for testing auditory perception," *Infant Behavior and Development* 18, 111–116.

Kindaichi, H. (1975) *Nihon no Hogen: Akusento no Hensen to Sono Jissou.* Tokyo: Kyoiku Syuppan.

Mehler, J., Bertoncini, J., Barriere, M. and D. Jassik-Gershenfeld (1978) "infant perception of mother's voice," *Perception* 7, 491–497.

Mehler, J., Jusczyk, P. W. and G. Lamberts (1988) "A precursor of language acquisition in young infants," *Cognition* 29, 143–178.

Moon, C., Cooper, R. P. and W. P. Fifer (1993) "Two-day-olds Prefer Their Native Language," *Infant Behavior and Development* 16, 495–500.

Mugitani, R., Hayashi, A. and S. Kiritani (2000) "Developmental change of 5 to 8-month-old Infants' Preferential Listening Response for The Native Dialect," *Jouranal of Phonetic Society of Japan* 4(2), 62–71.

Satoh, K. (1977) *Kokugogaku Kenkyu Jiten.* Tokyo: Meiji Shoin.

Sanada, Shinji (1988) "Hougen to Kyotsugo." In Kindaichi, H., Hayashi, O. and T. Shibata (eds.) *Nihongo Hyakka Daijiten.* (pp. 913–924). Tokyo: Taisyukan Shoten.

Sugitoh, Miyoko (1996) *Nihongo no Oto.* Tokyo: Izumi Shoin.

APPENDIX – Kana Transcripts of 6 samples of each dialect used for the experimental phase in infant behavioral experiment and acoustic analysis. Underlines represent moraes examined for vowel devoicing and vertical bars represent pauses with 300 msec or more of silence. Parenthetic digits represent the duration of the sample.

Picture with red Dragonflies

(Eastern) あっ，あかいとんぼだねえ，いっぱいとんでるよ。| このとんぼなんていうかしってるかなあ。| あかとんぼっていうんだよ。| あかーいとんぼだから，| あかとんぼっていうんだね。| めがついてるね。おめめがちゃんとふたつついてて，| はねがついてて，| そらをとべるんだよね，とんぼさんは。(24.37sec)

(Western) あ，これあかとんぼやなあ。| おそらにいっぱいあかとんぼとんでんなあ。| いつやろなあ，きせつ。| あかとんぼとんでんのは，| あきやで。| あきになったらあかとんぼいっぱいとぶなあ，おそらに。| いいなあ，きもちよさそうにとんでんなあ。(21.03sec)

Picture with a white dog sleeping on a sofa

(Eastern) しろいわんわんが，| みどりいろのソファーのうえでねているよ。| なんのゆめをみているのかなあ。| しろいわんわんきもちよさそうに，ぐうぐうってねているねえ。| しっぽまでしろいよ。| きっと，| おいしいごはんのゆめをみているのかもしれないね。かわいいねえ。よしよしってなでてあげてね。(22.87sec)

(Western) みてみ。| ソファのうえになんかねてんな。なんやろ。| わんわんやなあ。| しろいわんわんやで。| きもちよさそうにねたはんなあ。| なんのゆめみたはるんやろ。| きっと，たべもののゆめみたはるんかもしれんなあ。| ぐうぐうってねてんで。(21.93sec)

Picture with a monkey handstanding on a tricycle

(Eastern) おさるさんがねえ，さんりんしゃにのってるよ。| おさるさん，さんりんしゃのうえでさかだちしているよ。| すごいねえ。| おさるさん，よくさんりんしゃのうえでさかだちできるねえ。| おさるさんは，なあんか，サーカスのおさるさんみたいだね。| ピンクいろのおぼうしと，ピンクいろのおようふくをきているよ。(24.03sec)

(Western) おさるさんがなあ，| じてんしゃにのったはるわ。| これさんりんしゃやなあ，じてんしゃじゃなくて。| なあ，ほんであかいさんりんしゃのうえでなにしたはる？| さかだちしたはんで。| すごいなあ，このおさるさんさかだちしたはるわ。| これサーカスのおさるさんかなあ。(20.97sec)

Picture with a couple of clouds in a sky

(Eastern) まっさおなおそらのうえに，くもがふたあつポカポカってうかんでるねえ。| くもっていろーんなかたちにかわるんだよ。| このくもは，| なあんか，| ソフトクリームみたいなかたちしてるねえ。| おいしそうだねえ。たべられるかな？(19.37sec)

(Western) そらにくもがうかんでんで。| なにいろやろ。| しろいしろいくもやなあ。| このくもなんか，| かたち，にてへん？| なににてるかわかる？| ソフトクリームそっくりやなあ。| むっちゃおいしそうなくもやなあ。ちょっとたべてみたくなんなあ。(19.70sec)

Picture with a mother and a baby elephant

(Eastern) ぞうさんのおやこは，とってもなかがいいのしってる？｜ぞうさんのおかあさんはねえ。｜あかちゃんぞうさんにおっぱいあげるんだよ。｜それでねえ，ねるときはね，じぶんのからだのしたでね，｜あったかーいところにあかちゃんぞうさんをねせるの。｜すごくなかがいいんだよ。(20.37sec)

(Western) ぞうのあやこやなあ。｜ぞうのおやこ，むっちゃなかいいねんで。｜おかあさんぞうはなあ，｜こぞうのことをようかわいがらはんねん。｜ほんでなあ，｜ねるときはな，｜いっしょにねんねんで。｜おかあさんぞうはなあ，｜こぞうをな，｜よしよしって，｜なでてあげはんねんで。(22.70sec)

Picture with snow falling on a forest

(Eastern) あ，ゆきがふってるねえ。｜はやしのうえにゆきがふってるよ。｜ゆきって，つめたくってねえ。｜なあんかしろくて，｜わたみたいにふわふわしているんだよ。で，｜おててとかにのっかるときえちゃうのがゆきなんだよ。(17.70sec)

(Western) あっ，おそらからなんかふってきてんな。｜しろいしろい，｜わたみたいなこれなんやろ。｜これはなあ，｜ゆきやで。｜ゆきはなあ，｜ふゆになると，｜おそらからいっぱいふってくんねん。｜つめたいけど，｜むっちゃきれいやねんで。(21.03sec)

Received 18 December, 2001;
Revision received 21 March, 2002;
Accepted 22 May, 2002.