3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

DISCRIMINATION OF ENGLISH /r-l/ AND /w-y/ BY JAPANESE INFANTS AT 6-12 MONTHS: LANGUAGE-SPECIFIC DEVELOPMENTAL CHANGES IN SPEECH PERCEPTION ABILITIES

Teruaki Tsushima[1], Osamu Takizawa[2], Midori Sasaki, Satoshi Shiraki, Kanae Nishi, Morio Kohno[3], Paula Menyuk[1], Catherine Best[4]

1. Boston University, 718 Commonwealth Avenue, Boston, Massachusetts 02215, U.S.A. , 2. Kansai Advanced Research Center, Communications Research Laboratory, Ministry of Posts and Telecommunications. 3. Kobe City University of Foreign Studies, 4. Wesleyan University, and Haskins Laboratories.

ABSTRACT

In this study, we investigated language-specific developmental changes in Japanese infants' perceptual discrimination of English approximants, /r-l/ and /w-y/ at 6-12 months. It was found that both /r-l/ and /w-y/ were discriminated at 6-8 months. At 10-12 months, however, /w-y/ was discriminated, whereas /r-l/ was not. The present findings support the following conclusions; 1) a developmental decline in discrimination of non-native approximants occurs toward the end of the first year, consistent with earlier reports on non-native stop consonant contrasts; 2) phonemic factors are crucial in accounting for the observed decline, given that discrimination of a native approximant contrast remains good throughout the first year.

I. INTRODUCTION

The effect of speech input on speech perception has been a matter of great theoretical and applied interest. The language-specific effects on speech perception have been amply demonstrated for phonologically mature listeners. For example, a great deal of research has been conducted on Japanese listeners' perception of an (American) English contrast, /r, l/, which is not phonemic in Japanese. It has been found that Japanese adults show substantial difficulty in discriminating and identifying /r, l/ [1]. To date, no study has examined developmental changes in perception of /r, l/ in an early stage of perceptual development. In the present study, we attempted to examine how Japanese infants' perceptual discrimination of /r, l/ changes developmentally during the first year of life, and to examine whether phonological factors play a role in developmental changes in speech perception, using an English contrast, /w, y/, which is phonemic in Japanese.

The questions addressed in this study were the following. First, "Does Japanese infants' discrimination of English /r-l/ decline during the first year of life?" "If it does, when does the decline occur?" The English /r, l/ is classified phonetically as approximants, which are phonetically intermediate between stops and vowels in certain respects. Previous research has found language-specific perceptual effects by 10-12 months for non-native stop consonant contrasts, but by 6 months for non-native vowel contrasts. Yet, as stops and approximants share a primary cue (the onset frequency of formants, and their transition) for discrimination, the timing of the decline for approximants might coincide with that for stops, occurring around 10-12 months of age. Alternatively, it has been hypothesized that for a non-native contrast which has relatively low psychoacoustic saliency such as /r-l/, the developmental decline might occur earlier (e.g. 6 months) than that for stops [2].

The second question addressed in this study was, "How does Japanese infants' discrimination of /r-l/ compare to that of /w-y/?" Previous research suggests that phonemic factors are involved in the observed developmental decline. In contrast to /r-l/, English /w-y/ is phonemic in Japanese (although slightly different phonetically from Japanese counterparts). Therefore, comparing discrimination of both contrasts would enable us to examine whether phonemic factors are involved in the developmental decline in discrimination of non-native approximants. It was hypothesized that while discrimination of /r-l/ would decline, that for /w-y/ would be maintained. To test these hypotheses, we tested a group of infants 6-8 months of age, and 10-12 months of age, and compared their performance in discrimination of American English /r, l/ and /w, y/.

II. PROCEDURE

2-1. Experimental procedure

To maintain continuity with previous research, we used a "conditioned eye fixation procedure", which is most similar to the one employed in Best, McRoberts & Sithole (1988) [3]. In this procedure, an infant is presented with speech stimuli contingent on fixation to a visual stimulus. When the infant habituates to one phonetic category, speech stimuli from the contrasting category are introduced. A significant recovery of fixation-time after the sound-shift indicates discrimination. In this study, the habituation criterion was defined as a 50% or more looking-time decrement in two consecutive trials (termed "pre-shift trials": "pre-shift phase"), relative to the average of the two longest looking-time among the first three trials. Immediately after the habituation criterion was reached, two trials were held in which the same speech stimuli were presented (termed "control trials":"control phase"), to ensure that an increment in looking-time after the pre-shift phase was not due to a random variation in looking-behavior. Immediately following the control phase, three trials with a new set of speech stimuli were held (termed "post-shift trials":"post-shift phase").

2-2. Subjects

Sixty-two infants from Kobe, Japan, were tested at a lab at Kobe City University of Foreign Studies. They were divided into two age groups; the younger age group (N=28, Mean age:221 days; Range:186 to 254 days); the older age group (N=34, Mean age:330 days; Range:288 to 372 days). No subject had had a history of ear infection according to the caretaker's report. The subject was tested on the r/l and the w/y pair in one day with the order of the speech stimuli pair counterbalanced across subjects.

2-3. Speech stimuli and visual stimulus

Speech stimuli used in the experiment were a part of a 10-step continuum of synthesized /rak-lak/, and of /wak-yak/, generated on the OVE-IIIc synthesizer at Haskins Laboratories, which had been used in Best, & Strange (1992) [1]. An overall duration (330 ms), intonation contour and a spectral pattern of the final vocalic portion of the syllable (105 ms) were held constant across the stimuli. For the /rak-lak/ series, F_2 and F_3 onset frequencies, F_3 frequencies at the point of inflection were varied as well as duration of steady-state and transition of F_1. For the /wak-yak/ series, F_2 and F_3 onset frequencies were varied (See [1] for a detailed description of the stimuli). We used the first and the second tokens on a continuum for each phonetic category (e.g. [rak]) in order to provide for some within-category variability in the speech stimuli. When presented the two stimuli were pseudo-randomly ordered. The inter stimulus interval was 750 ms. The visual stimulus was a computer-generated geometric human face.

2-4. Experimental settings and apparatus

The experiment took place in an open-ended booth (approximately 1 x 1.2 m) surrounded by black walls and curtains in a sound-treated room. A caretaker sat on a chair facing the closed end of the black wall, and an infant sat on her or his lap, facing a video monitor (15 x15 cm) placed on the wall. A video camera was mounted on the video monitor, monitoring and filming the infant's face and the upper body through a square window (15x15) on the wall covered with a sheet of opaque black felt, which prevented the infant's viewing the camera. An observer monitored the subject's eye movement and pressed a button contingent on the subject's fixation to the visual stimulus. Presentation of the audio and visual stimuli was controlled by a specialized software program (IWATSU:ISEL) run on a PC (NEC:98) and associated hardware facilities (IWATSU:ISEL). The audio stimuli were presented at an intensity level of approximately 65 dB at the infant's ears. The looking-time of each trial was recorded on a PC file, and the infant's looking-behaviors were also video-recorded.

2-5. Exclusion of the data

A part of these data sets were excluded from the analyses for several reasons (e.g. the subject was restless and did not complete the session; the habituation criterion was not reached within 23 trials from the beginning of the session, etc.). Furthermore, in an attempt to determine objective exclusion criteria, we performed post-hoc analyses on possible external factors which might significantly affect looking-behaviors. Based on the analyses, we determined to exclude the data in which some behaviors that significantly affected looking-time were present in the control and the post-shift phase (e.g. reaching for the visual stimuli). Second, we also excluded data in which non-looking time was exceptionally long in the control or post-shift phase and in which the amount of time spent from the beginning to the sound-shift was exceptionally long. All together, 144 out of 201 data sets were excluded by these criteria.

III. RESULTS

3-1. The final data

The final data were grouped into the following four groups according to age (young or old) and speech stimuli (RL, or WY): 1) Young:RL, 2) Young:WY, 3) Old:RL, 4) Old:WY. The data consisted of 44 data sets from 30 subjects (out of 62 subjects tested). The younger age group consisted of 14 subjects (Mean age=220 days; Range=186 to 248 days), and the older age group consisted of 16 subjects (Mean age=327 days; Range=291 to 370 days). The number of data sets in each group was, 1) YOUNG:RL (10), 2) YOUNG:WY (10), 3) Old:RL (11), 4) Old:WY (13).

Since we had to exclude a part of the data sets based on the post-hoc analyses of the data, the final data turned out not to be totally balanced in terms of some experimental conditions such as the session of experiment (1st session or 2nd session), the order of experiment (RL->WY or WY->RL). However, statistical analyses revealed no significant effects of these factors on the looking-time.

3-2. Reliability of an observer's measurement of looking-time

Reliability of observers' measurement of looking-time was assessed by calculating a reliability coefficient (Cronbach's alpha), using looking-time measured by the observers and its corresponding looking-time measured independently by researchers on the video. Looking-time of all the trials in the final data was included in the analyses. The reliability coefficient was 0.987 (N=477).

3-3. Analyses of the amount of habituation

Statistics on the amount of habituation (i.e. a sum of looking-time until the habituation criterion was reached) among four groups, using a MANOVA (AGE (YOUNG, OLD) X LANGUAGE (RL, WY), revealed no significant effects. Therefore, we can be assured that the subjects tested in each condition received a reasonably equal amount of habituation.

3-4. Analyses of looking-time increments: evidence for discrimination

In the procedure used in this study, evidence for discrimination would be obtained when there was no significant looking-time increment between the pre-shift and control phase, and at the same time, there was a significant increment in looking-time between the control and post-shift phase. The average looking-time increment between the pre-shift and control phase, and between the control and post-shift phase in (a) Young group and (b) Old group is presented below in Figure 1.

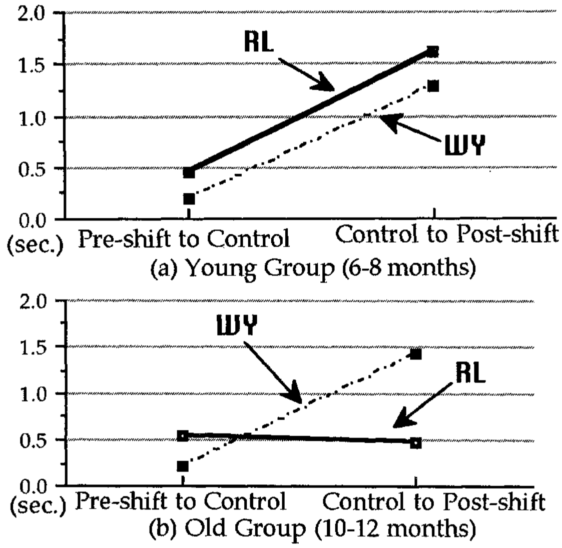

Fig. 1- Looking-time changes between the pre-shift and the control phase, and its recovery in the post-shift phase.

A MANOVA repeated measures design was run using PHASE (PRE-SHIFT, CONTROL, POST-SHIFT) as a within-subject factor, and AGE (YOUNG, OLD) X LANGUAGE (RL, WY) as between-subject factors. The main effect of PHASE was significant (F (2,29)=15.54: p=.000), indicating that overall, looking-time increased significantly from the pre-shift, control to the post-shift phase. However, none of the interaction terms was significant.

Simple effects analyses at each level of AGE (YOUNG, OLD) X LANGUAGE (RL, WY) revealed that the effect of PHASE was significant at Young:RL (F(2,39)=6.35: p<.01), Young:WY (F(2,39)=3.40: p=.04), and Old:WY (F (2,39)=5.28: p<.01), but not significant at Old: RL (F(2,39)=2.13: p=.13), indicating that the looking-time significantly increased from the pre-shift to post-shift phase except for the Old:RL group.

T-tests were further conducted to test whether for each group the looking-time significantly increased between the pre-shift and control phase and between the average of the first two phases and the post-shift phase. The results showed that none of the looking-time increment between the pre-shift and control phase was significant, while the increment between the average of the first two phases and the post-shift phase was significant for Young:RL (p<01), Young:WY (p=.01) and Old:WY (p=.01), but not for Old:RL (p=.15).

Therefore, the overall results indicate that the infants in the younger group discriminated both English /r-l/ and /w-y/, while those in the older group discriminated /w-y/, but failed to discriminate /r-l/. Therefore, the data provided evidence of a developmental decline in discrimination of /r-l/ by 10-12 months, as well as evidence that discrimination for /w-y/ was maintained throughout the latter half of the first year.

IV. DISCUSSION

First, it was found that Japanese infants' discrimination of /r-l/ declined between 6-8 and 10-12 months. This finding supports the conclusion that a developmental decline in discrimination of non-native approximants occurs by 10-12 months. This finding is consistent with earlier findings on a developmental decline in some non-native stops. On the other hand, the finding does not support the hypothesis that a developmental decline for acoustically less salient non-native contrasts including approximants occurs earlier than acoustically more salient contrasts such as place-of-articulation differences in stops [2]. In the light of recently reported evidence that discrimination of non-native vowels further declines from 8-10 months to 10-12 months [4], the overall findings suggest profound language-specific effects on speech perception at around 10-12 month.

Second, it was found that while discrimination of /r-l/ declined between 6-8 to 10-12 months, that for the phonemic contrast, /w-y/, was maintained. This finding supports the conclusion that phonemic factors play a crucial role in accounting for the observed developmental decline in discrimination of non-native contrasts at 10-12 months of age. In this regard, the finding is compatible with the hypothesis that the developmental decline is associated with a shift from perceiving language-general information about phonetic (acoustic and articulatory) properties of speech to perceiving information about phonologically relevant properties in the speech signal [5].

However, little is known about the mechanisms which bring about the language-specific effects on speech perception. For example, it is not clear why the language-specific perceptual effects begin to appear earlier for vowels than stops and approximants. Therefore, a further direction for research would be to attempt to identify the factors (linguistic, cognitive, memory, etc.) which might underlie the developmental changes in speech perception abilities.

References:

[1] Best, C. T. and W. Strange. "Effects of phonological and phonetic factors on cross-language perception of approximants." Journal of Phonetics 20 (1992): 305-330.

[2] Burnham, D. K. "Developmental loss of speech perception: Exposure to and experience with a first language." Applied Psycholinguistics 7 (1986): 207-240.

[3] Best, C. T., G. W. McRoberts, and N. M. Sithole. "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants." Journal of Experimental Psychology: Human Perception and Performance 14 (1988): 345-60.

[4] Polka, L. and J. F. Werker. "Developmental changes in perception of non-native vowel contrasts." Journal of Experimental Psychology: Human Perception and Performance (in press).

[5] Jusczyk, P. W. "From general to language-specific capacities: the WRAPSA Model of how speech perception develops." Journal of Phonetics 21 (1993): 3-28.

Acknowledgment:

We would like to thank the following people for their invaluable help in conducting this research; Janet Werker and Toshisada Deguchi for helpful advice on various aspects of the experiment; Kenneth Stevens for his instruction on stimulus development in the early stage of the study. This study was supported, in part, by a grant, "Research and Development in the Frontier of Electronic Communication", from Japanese Ministry of Posts and Telecommunications to Morio Kohno, and also by a grant, "Development of Cognition and Language", from Japanese Ministry of Education to Morio Kohno.