Title	A preliminary study on infants' learning of word-object relations : the role of social interaction
Sub Title	
Author	白野, 陽子(Hakuno, Yoko) 大森, 貴秀(Omori, Takahide) 山本, 淳一(Yamamoto, Junichi) 皆川, 泰代(Minagawa Kawai, Yasuyo)
Publisher	Centre for Advanced Research on Logic and Sensibility The Global Centers of Excellence Program, Keio University
Publication year	2012
Jtitle	CARLS series of advanced study of logic and sensibility Vol.5, (2011.) ,p.127- 136
JaLC DOI	
Abstract	
Notes	I. Study of Logic and Sensibility Part 2 : Genetics and Development
Genre	Research Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO12002001-20120224-0127

13 A Preliminary Study on Infants' Learning of Word-object Relations: The Role of Social Interaction

Yoko Hakuno[1], Takahide Omori[1,3], Jun-ichi Yamamoto[1,3], and Yasuyo Minagawa-Kawai[2,3]

[1] Department of Psychology, Keio University
[2] Graduate School of Human Relations, Keio University
[3] Centre for Advanced Research on Logic and Sensibility (CARLS), Keio University

I. Introduction

During the first year of life, infants begin to acquire spoken words. Previous research has revealed that English-learning infants are able to segment words from continuous speech at 7.5 months of age (Jusczyk & Aslin, 1995). As well as infants in English-speaking countries, Japanese infants seem to be capable of segmenting words at around 8 months of age (Sato et al., 2007). The age at which acquisition of word-object relations begins is controversial; however, recent research has documented that 6-month-olds are already able to learn word-object associations (Shukla et al., 2011). In order for young infants to learn words, the way of teaching is crucial. According to Kuhl et al. (2003), infants' foreign language phonetic learning is enhanced by social interaction. Although the significance of social interaction for language acquisition has been argued, there is not enough experimental data. In our study, we assessed the role of social interaction in infants' word learning.

We used a forced-choice preferential looking (FPL) paradigm (Teller, 1979; McCleery et al., 2007) to test infants' word learning. The FPL paradigm is based on the fact that infants look longer at a patterned stimulus on

one side of a display rather than nothing on the other side. On each trial, the stimulus appears on the left or right side of the display. An experimenter observes the infant's eye gaze behavior through a video camera monitor to judge where the stimulus locates. Correct judgments above chance level indicate that the infant is able to detect the stimulus. This FPL technique is often used to assess visual acuity in infancy, and therefore, only a visual stimulus is usually presented on each trial. In our study, we applied this paradigm to assess the time infants looked at the congruent and incongruent visual and auditory stimuli by presenting both the visual and auditory stimulus together.

In Experiment 1, the FPL paradigm was examined to determine if it can detect infants' preference for congruent visual and auditory stimuli. In Experiments 2-1 and 2-2, we tested whether 5- to 12-month-old infants can segment a word from continuous speech and acquire understanding of a word-object relation through learning with social interaction (Experiment 2-1) and without social interaction (Experiment 2-2).

II. Experiment 1

It is well known that infants develop perception of audiovisual temporal synchrony during the first few months of life. According to Dodd (1979), 2- to 4-month-old infants look at a human speaker longer when voice and lip movements are in synchrony rather than out of synchrony. In Experiment 1, we attempted to replicate the findings of Dodd (1979) using a different technique, the FPL paradigm, in order to establish that this paradigm can be used to detect infants' sensitivity to audiovisual correspondence.

1. Methods

1.1. Participants

Twelve Japanese infants (M = 7.0 months, range = 4 – 10 months; 8 males) served as participants. All participants were full-term infants with no history of medical problems. Informed consent was obtained from each parent before the infant's participation.

1.2. Stimuli

The stimulus was a 20-second movie clip of a female adult's speaking face and voice. There were two conditions: congruent (in synchrony) and incongruent (out of synchrony) condition. In the congruent condition, the speaker's lip movements and voice were temporally synchronized, whereas in the incongruent condition, the speaker's voice was delayed behind the lip movements by 5000 ms.

1.3. Procedure

The experiments were performed in a quiet testing room at Keio University. The infant sat on the parent's lap, facing a 19-inch display in the testing area. Auditory stimuli were presented from one of two speakers located at the right and left sides of the display. The visual stimulus was approximately 40 cm away from the infant's eyes. A video camera placed above the display recorded the infant's eye gaze. From an observation area in the testing room, the experimenter monitored the infant's eye gaze and controlled the presentation of stimuli using a computer. The observation area and the testing area were divided by a thick curtain. The parent and the experimenter were blind to the auditory stimuli presented.

Each trial began with the presentation of a fixation point on the display and a fanfare sound from both speakers (Figure 1). As soon as the infant's gaze turned to the display, the movie clip of a female face appeared on the right or left side of the display. Shortly after the infant moved his eyes to the visual stimulus, the auditory stimulus was presented from the speaker on the same side as the visual stimulus. When the infant looked away from the visual stimulus for more than 2 seconds or continued to look at the stimulus for longer than 40 seconds, the trial ended and the experimenter stopped

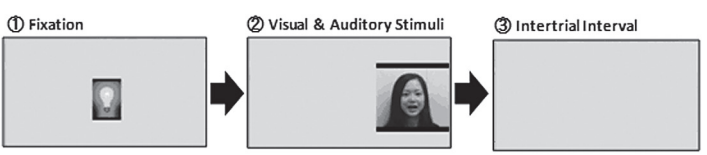

Figure 1. An example of stimuli in a trial using the forced-choice preferential looking paradigm.

presenting all stimuli. The intertrial interval was 1 second. Each congruent and incongruent condition was presented once on the right side and once on the left side of the display; thus, one session consisted of four trials. The presenting order of the two conditions was counterbalanced within one session and within participants. The time infants looked at each visual stimulus was recorded by the experimenter.

2. Results and Discussion

Twelve infants' mean looking time for the visual stimulus in the congruent condition was 19.7 seconds, which was significantly longer than the mean looking time in the incongruent condition of 12.9 seconds (t (11) = 3.12, p<.01) (Figure 2). These results indicate that, the infants preferred the stimuli when the speaker's lip movements and voice were temporally synchronized rather than when the voice was delayed behind the lip movements, and this accorded with the results of Dodd (1979). The results of Experiment 1 confirmed that the FPL paradigm could effectively detect infants' preference for audiovisual correspondence.

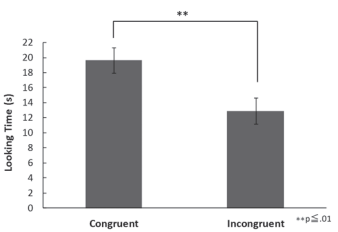

Figure 2. Mean looking time for congruent (in synchrony) and incongruent (out of synchrony) conditions in Experiment 1 (N = 12).

III. Experiment 2-1

The validity of the FPL paradigm using audiovisual stimuli was proved in Experiment 1. Here in Experiments 2-1 and 2-2, we used the same paradigm to test infants' word-object mapping ability. In Experiment 2-1, a live speaker sitting in front of the infant named and moved the object while making frequent eye contact. That is, the infants learned the word-object relation through training with social interaction.

1. Methods

1.1. Participants

Eight Japanese infants were tested as the experimental group (*M* = 6.5 months, range = 5–9 months; 4 males), and another eight infants comprised the control group (*M* = 6.3 months, range = 5–8 months; 6 males). They were all full-term infants with no history of medical problems. Informed consent was obtained from each parent before the infant's participation.

1.2. Stimuli

Two auditory words (*/yaQpuN/* and */doHmiN/*) acted as target words and one auditory word (*/naNkiH/*), as the control word. For all the words, Japanese pitch accent was assigned to the first mora, resulting in a High-Low-Low pitch pattern. We chose these nonsense words to ensure that the infants had never heard the three words before. The visual stimuli were three puppets corresponding to each auditory word (Figure 3). They were videotaped individually and made into three 40-second movie clips. The auditory stimuli were three short texts consisting of six sentences, with one target or control word in each text. These words were embedded in sentences (e.g., "*Yappun* lives in a forest. I first saw *Yappun* yesterday."). The texts were read by a female adult who was unfamiliar with the infants. In the test phase, there were four stimulus conditions (Table 1).

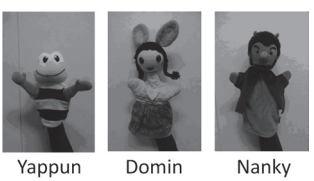

Figure 3. Puppets used as visual stimuli in Experiments 2-1 and 2-2. Yappun and Domin were treated as the target stimuli and Nanky was the control stimulus.

Table 1. Visual and auditory stimuli in four stimulus conditions in Experiments 2-1 and 2-2.

	Target Object		Control Object	
	Congruent	Incongruent	Congruent	Incongruent
Visual Stimulus:	Yappun or Domin puppet	Yappun or Domin puppet	Nanky puppet	Nanky puppet
Auditory Stimulus:	*/yaQpuN/* or */doHmiN/*	*/naNkiH/*	*/naNkiH/*	*/yaQpuN/* or */doHmiN/*

1.3. Procedure

The experiments were performed in the same testing room with the same apparatus used in Experiment 1.

(1) Training phase: The infants in the experimental group learned a target stimulus (e.g., "yappun" and the Yappun puppet), and the infants in the control group learned a non-target stimulus (e.g., "domin" and the Domin puppet). Infants were exposed to the word inserted in short sentences 30 times by a live female speaker with a moving puppet for about 1.5 minutes in duration. The speaker used infant-directed speech and made frequent eye contact and at least one physical contact with the infants.

(2) Test phase: The FPL paradigm utilized in Experiment 1 was used to test infants' word learning. Throughout the phase, the parent and the experimenter were blind to the auditory stimuli presented. The target stimulus ("yappun" and the Yappun puppet or "domin" and the Domin puppet) and the control stimulus ("nanky" and the Nanky puppet) were presented to both groups. During congruent trials, the puppet on the display and the repeated spoken word embedded in sentences corresponded (e.g., "yappun" and the Yappun puppet). During incongruent trials, the object and the word did not correspond (e.g., "nanky" and the Yappun puppet). Four stimulus conditions were presented once on the right side and once on the left side of the display, thus one session consisted of eight trials. The presenting order of the two conditions was counterbalanced within one session and within participants. The time infants looked at the visual stimulus in each condition was recorded by the experimenter.

2. Results and Discussion

Mean looking times were submitted to a mixed ANOVA with group (2) as a between-participants factor and visual stimulus (2) and condition (2) as within-participants factors. A marginally significant interaction was found among the group, visual stimulus, and trial types ($F(1, 14) = 4.53$, $p = .0516$) (Figure 4). The post-hoc tests revealed that the experimental group looked at the target word marginally longer than the control group ($p < .1$). This suggests that 6.5-month-olds are already capable of segmenting words from continuous speech through training with social interaction. However, the infants in the experimental group looked at the stimuli with the target word

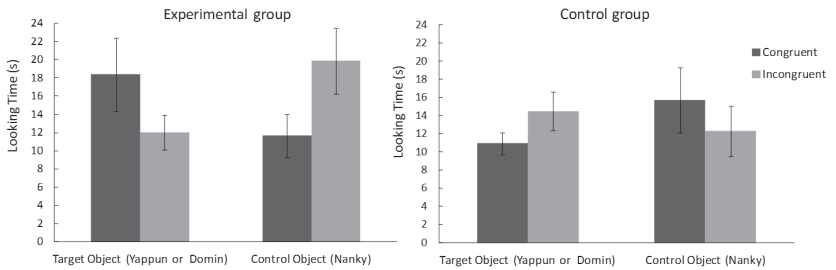

Figure 4. The experimental group and control group infants' mean looking time in each stimulus condition in Experiment 2-1.

equally long regardless of the visual stimulus. This result indicates that they failed to learn the word-object association.

IV. Experiment 2-2

In Experiment 2-1, the infants were exposed to the auditory word and the object presented by a live speaker, and they succeeded in segmenting the word from continuous speech. The question remains, would it be difficult for infants to learn a word in the absence of social interaction? To assess the role of social interaction in infants' word learning, the same speaker presented the word and object via television monitor in Experiment 2-2.

1. Methods

1.1. Participants

Eight Japanese infants participated as the experimental group (M = 8.0 months, range = 5–11 months; 3 males) and another eight infants participated as the control group (M = 7.8 months, range = 5–12 months; 4 males). They were all full-term infants with no history of medical problems. Each parent gave informed consent before his or her infant's participation.

1.2. Stimuli

As the training stimuli, we created two movie clips of the speaker's face and

the moving puppet. One was for "yappun" with the Yappun puppet and the other for "domin" with the Domin puppet. The speaker was the same female adult as in Experiment 2-1. As in Experiment 2-1, the speaker repeated the target word that was embedded in short sentences using infant-directed speech. The speaker also moved the puppet as she did in Experiment 2-1. The test stimuli were the same as in Experiment 2-1.

1.3. Procedure

(1) Training phase: The word-object exposure was similar to Experiment 2-1 except that the infants were exposed to the word and object via a television monitor. The same speaker as in Experiment 2-1 appeared only in the monitor; hence, no actual eye contact or physical contact was made during the training phase.
(2) Test phase: The procedure in the test phase was identical to that reported in Experiment 2-1.

2. Results and Discussion

As in Experiment 2-1, mean looking times were submitted to a mixed ANOVA with group (2) as a between-participants factor and visual stimulus (2) and condition (2) as within-participants factors. No significant main effect or interaction was found (Figure 5). Thus, it seems clear that infants trained without social interaction failed in both the word-object association and word segmentation from continuous speech.

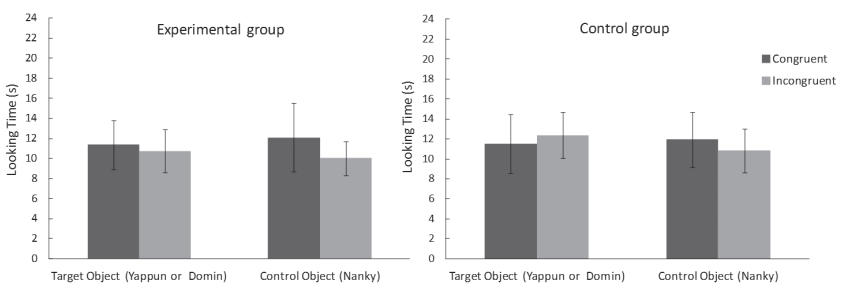

Figure 5. The experimental group and control group infants' mean looking time in each stimulus condition in Experiment 2-2.

V. General Discussion

The present study examined (i) the validity of the FPL paradigm to detect infants' sensitivity to audiovisual correspondence using both auditory and visual stimuli, and (ii) the effect of social interaction on word-object mapping in infancy. Experiment 1 asked whether, using the FPL paradigm, infants look longer at stimuli when the speaker's lip movements and voice are temporally synchronized than when lip movements and voice are out of synchrony. Experiments 2-1 and 2-2 tested two methods of teaching word-object relation in infants. In Experiment 2-1, a live speaker presented the auditory word and the object while interacting socially with the infants. In Experiment 2-2, the infants were exposed to the auditory word and the object by the speaker appearing on a television monitor.

The results from Experiment 1 demonstrated that the FPL paradigm effectively revealed infants' preference for audiovisual congruency. In Experiment 2-1, the 6.5-month-old infants seemed to be able to segment a word from spoken sentences, although they did not succeed in acquiring the word-object relation through training with social interaction. By contrast, in Experiment 2-2, 8-month-olds were unable to learn the word-object association or segment the word from sentences in the absence of social interaction. These results documented that social interaction possibly enhances learning of auditory words in infancy. In addition, the results indicate that, given training with a live person, Japanese infants are capable of segmenting a word from continuous speech at around 6 months of age, which is 1 to 3 months earlier than previously reported. However, the present study could not prove that 6.5-month-old infants can learn word-object relations. This was perhaps due to the difficulty of the tasks we set in our study. Therefore, it is necessary to either modify the level of the tasks or raise the age of the participants to, for instance, 9 months. Although the present preliminary study revealed a crucial role of social interaction in word learning for 5- to 9-month-old infants, further study evaluating detailed age differences is needed to elucidate developmental changes in the role of social interaction for word learning.

Acknowledgement

Our study was supported by the Centre for Advanced Research on Logic and Sensibility (CARLS), global Center of Excellence (gCOE), Keio University. We also appreciate the assistance offered by Sayaka Ishii, Noriko Morisawa, and Akiko Hakuno.

References

Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, **11**, 478–484.

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

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences,* **100(15)**, 9096–9101.

McCleery, J. P., Allman, E., Carver, L. J., and Dobkins, K. R. (2007). Abnormal magnocellular pathway visual processing in infants at risk for autism. *Biological Psychiatry*, **62**, 1007–1014.

Sato, K., Kajikawa, S., Sakamoto, K., and Matsumoto, H. (2007). The role of accent patterns and phonotactics in Japanese infant's word segmentation from fluent speech. *Journal of the Phonetic Society of Japan*, **11(3)**, 38–47.

Shukla, M., White, K. S., and Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, **108**, 6038–6043.

Teller, D. Y. (1979). The forced-choice preferential looking procedure: A psychophysical technique for use with human infants. *Infant Behavior and Development,* **2**, 135–153.