

Strategy in short-term memory for pictures in childhood : A near-infrared spectroscopy study

Sanefuji, Masafumi

Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University

Takada, Yui

Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University

Kimura, Naoko

Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University

Torisu, Hiroyuki

Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University

他

<https://hdl.handle.net/2324/25460>

出版情報 : NeuroImage. 54 (3), pp.2394-2400, 2011-02-01. Elsevier

バージョン :

権利関係 : (C) 2010 Elsevier Inc.

**Strategy in Short-Term Memory for Pictures in Childhood: a Near-Infrared
Spectroscopy Study**

Masafumi Sanefuji, Yui Takada, Naoko Kimura, Hiroyuki Torisu, Ryutaro Kira, Yoshito
Ishizaki, and Toshiro Hara

Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University,
3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

Corresponding author: Masafumi Sanefuji, MD, PhD

Department of Pediatrics, Graduate School of Medical Sciences, Kyushu University,
3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan.

Tel: +81-92-642-5421, Fax: +81-92-642-5435; Email:

sane26@pediatr.med.kyushu-u.ac.jp

Abstract

In Baddeley's working memory model, verbalizable visual material such as pictures are recoded into a phonological form and then rehearsed, while auditory material is rehearsed directly. The recoding and rehearsal processes are mediated by articulatory control process in the left ventrolateral prefrontal cortex (VLPFC). Developmentally, the phonological strategy for serially-presented visual material emerges around 7 years of age, while that for auditory material is consistently present by 4 years of age. However, the strategy change may actually be correlated with memory ability as this usually increases with age. To investigate the relationship between the strategy for pictures and memory ability, we monitored the left VLPFC activation in 5 to 11 year old children during free recall of visually- or auditorily-presented familiar objects using event-related near-infrared spectroscopy. We hypothesized that the phonological strategy of rehearsal and recoding for visual material would provoke greater activation than only rehearsal for auditory material in the left VLPFC. Therefore, we presumed that the activation difference for visual material compared with auditory material in the left VLPFC may represent the tendency to use a phonological strategy. We found that the activation difference in the left VLPFC showed a significant positive correlation with memory ability but not with age, suggesting that children with high memory ability make more use of phonological strategy for pictures. The present study provides functional evidence that the strategy in short-term memory for pictures shifts gradually from

non-phonological to phonological as memory ability increases in childhood.

Keywords: child, short-term memory, recoding, rehearsal, near-infrared spectroscopy

1. Introduction*

Short-term memory plays fundamental roles in performing a wide range of cognitive skills including learning, particularly in childhood. The most influential account of short-term memory is provided by Baddeley's model of working memory, which proposes central executive and two slave systems (Baddeley, 1986). These slave systems represent a visuospatial sketchpad for visual images and a phonological loop for speech-based information. The phonological loop consists of a phonological store and an articulatory control process. The phonological store holds information in a phonological form, while the articulatory control process serves to rehearse the decaying information in the store. The articulatory control process is also implicated in recoding verbalizable pictures into a phonological form (Baddeley, 1986). Functional neuroimaging studies have demonstrated the association of the articulatory control process with the left ventrolateral prefrontal cortex (VLPFC) including Broca's area (Awh et al., 1996; Baddeley, 2003; Henson et al., 2000; Koelsch et al., 2009; Paulesu et al., 1993; Smith et al., 1998). Dysfunction of the articulatory control process in working memory is implicated in neurodevelopmental disorders such as autism and Down syndrome

***Abbreviations:** AI: auditory index, DLPFC: dorsolateral prefrontal cortex, NIRS: near-infrared spectroscopy, VI: visual index, VLPFC: ventrolateral prefrontal cortex, VAI: visual-auditory index

(Alloway and Gathercole, 2006).

Short-term memory shows developmental changes in strategy as well as efficiency during preschool through to early adolescence (Gathercole, 1999). A well-established strategy change is memory for serially-presented pictures of familiar objects which changes between 4 to 11 years of age. Children below 7 years of age are likely to rely on a visuospatial sketchpad, while children beyond this age tend to use the phonological loop by recoding pictures into a phonological form (Gathercole and Hitch, 1993; Halliday et al., 1990; Hitch and Halliday, 1983; Hitch et al., 1989). The strategy change was suggested to form a continuum from 5 to 8 years of age, rather than the simple dichotomous pattern (Palmer, 2000); the strategy progressed from no obvious strategy, to visual encoding, then to a dual (visual and phonological) code, and finally to phonological encoding. In that study there was also a close relationship between the phonological strategy and high recall ability within the same ages, or occasionally irrespective of age. Therefore, the strategy change may be correlated with children's memory ability rather than with age, as memory ability usually increases with age. In contrast to visual material such as pictures, auditory material has direct and obligatory access to the phonological store, and is rehearsed consistently after at least 4 years of age (Hulme et al., 1984). These findings are exclusively based on behavioral studies using word length effect and visual and phonological similarity effects. To the best of our knowledge, there are no functional neuroimaging studies addressing this issue. One

functional magnetic resonance imaging (MRI) study showed rehearsal-related activation in the left VLPFC with a visually-presented letter 2-back task in 9 to 11 year old children (Casey et al., 1995). However, that study only investigated the involvement of rehearsal process after recoding, and did not examine the emergence of recoding and rehearsal process.

Near-infrared spectroscopy (NIRS) is a noninvasive neuroimaging method for monitoring cerebral hemodynamics by measuring changes in the attenuation of near-infrared light passing through tissue. NIRS exhibits some advantages over functional MRI or positron emission tomography (PET) in that it is noiseless and is less restrictive to body movements, allowing even preschool children to perform tasks under natural conditions equally with auditory and visual presentation (Minagawa-Kawai et al., 2007; Moriguchi and Hiraki, 2009). As NIRS has a sufficient temporal resolution and a good signal-to-noise ratio, it is suitable for event-related task paradigms with relative few experimental trials (Schroeter et al., 2002). However, limited regions of interest need to be determined before experimentation as each NIRS channel monitors only a small area in the lateral cortical surface with relative low spatial resolution (Villringer et al., 1993).

To investigate the relationship between the strategy in short-term memory for serially-presented pictures and memory ability in children, we monitored the left VLPFC activation during free recall of familiar objects with visual and auditory presentation

using event-related NIRS. We hypothesized the following. (1) The left VLPFC would exhibit activation for auditory material as children rehearse auditory material consistently. (2) The left VLPFC would exhibit greater activation for visual material when a child uses a phonological strategy of recoding and rehearsal, and as such, the left VLPFC would show greater activation for visual than for auditory material. (3) The left VLPFC would exhibit minimal activation for visual material when a child relies on a non-phonological strategy without recoding and rehearsal, and as such, the left VLPFC would show reduced activation for visual than for auditory material. Based on these hypotheses, we presumed that the difference in activation for visual material compared with auditory material in the left VLPFC would represent a tendency to use a phonological strategy in short-term memory for pictures. As this tendency should be more pronounced in children with high memory ability (Palmer et al., 2000), we used correlation analyses to examine the relationship between the activation difference that could represent strategy and memory ability.

2. Methods

2.1. Subjects

Thirty-one right-handed healthy Japanese children (15 girls, age range: 4.3 to 11.0 years; mean [SD]: 7.3 [1.8]) were recruited for subjects. The handedness was confirmed by the Edinburgh Handedness Inventory (Oldfield, 1971). Twelve children were excluded from the analysis due to poor task performance below 0.75 ($n=5$) and a large number of trials

with motion artifacts in the NIRS signals ($n=7$). The remaining 19 children (10 girls, age range: 5.4 to 11.0 years; mean [SD]: 7.5 [1.6]) were analyzed. No subjects had preexisting neurological or developmental disorders. The present study was approved by the Ethics Committee of the Graduate School of Medical Sciences, Kyushu University. Written informed consent was obtained from parents before the investigation.

2.2. Instruments

A multichannel NIRS system (OMM-2001; Shimadzu, Kyoto, Japan) was used to monitor temporal changes in oxy-, deoxy-, and total-Hb concentrations in four channels. For each channel, the absorption of near-infrared light at 780, 805, and 830 nm wavelength was measured at a scanning rate of 70 msec. A visual-auditory stimuli presentation system (IS-703 AV Tachistoscope; Iwatsu, Tokyo, Japan) controlled stimuli presentation with on-line connection to the NIRS system.

2.3. Channel positions

Four channels were set on F3, F4, F7, and F8 according to the international 10-20 system for electroencephalography (Jasper, 1958). These approximate positions correspond to the left and right dorsolateral prefrontal cortex (DLPFC) and the left and right VLPFC, respectively, based on an anatomical craniocerebral correlation study (Okamoto et al., 2004a). Each channel was constructed by a pair of an emitter probe and a detector probe at a distance of 3 cm from each other. Bolt-combined probes were screwed one by one into four separate head-holders arranged to cover the exact channel positions

(Supplementary Fig. 1). To visualize the anatomical positions of the channels on the cerebral cortex, we obtained structural head MRI of a 6 year-old child. The positions of the four channels were labeled with MRI-detectable markers (lipid-containing capsules), and superimposed on the MRI of the three-dimensionally reconstructed cerebral cortex using software Fusion© (Shimadzu). The superimposed channel positions conformed well to the results of the anatomical craniocerebral correlation study (Fig. 3).

2.4. Stimuli

Stimuli consisted of children-familiar items (e.g., car, rabbit, grapes). These items have two or three morae (mora is a phonological unit of sound in Japanese, analogous to syllable in English) when pronounced. In the visual condition, each item was presented as a picture at the center of a display, and was subtended approximately 3 degrees of visual angle. In the auditory condition, each item was presented as a voice (Amano and Kondo, 1999) from an audio speaker.

2.5. Memory ability

Each child underwent the Digit Span subtest of the Japanese version of the Wechsler Intelligence Scale for Children-Third Edition (Japanese WISC-III Publication Committee, 1998) for evaluation of their memory ability. This test requires the children to repeat a specific number of digits that are read aloud to them, in given and reverse order.

2.6. Experimental procedures

Each child performed a task sitting on a comfortable chair in a dimly lit silent room. The

paradigm consisted of two experimental conditions (Fig. 1). Every trial began with the presentation of a sequence of three pairs of visual and auditory stimuli. In the visual condition, each item was presented as a picture simultaneously with an auditory control stimulus (a beep sound). In the auditory condition, each item was presented as a voice simultaneously with a visual control stimulus (a gray circle). Each pair was presented for 1750 ms followed by a fixation cross for 250 ms, for a total of 2 s. The children were instructed to encode and maintain the three items silently in their head, without body movements such as head tilting and strong biting, which could produce motion artifacts in NIRS data. After a fixation cross of 6 s retention, a blue circle was presented for 6 s, signaling children to retrieve and say the three items aloud during the period. If children could not say all items within the period, they were told to stop the retrieval and attend to the next trial. A fixation cross of 6 s rest intervened between the end of the blue circle and the start of the next trial. Some body movements were allowed during the retrieval period and the early part of the rest period as children often cannot tolerate keeping still for extended periods. The children were asked to focus on the fixation cross between stimuli. Trial conditions were pseudo-randomly ordered such that each trial condition was equally often preceded and followed by each other. Across subjects, the assignment of items to the conditions was counterbalanced. Ten trials constituted a run, which lasted 4 min. Each child performed as many runs as possible from two to four runs (mean: 3.7), with a short break between the runs. Before the experiment, the children briefly practiced the

procedure.

2.7. Data analysis

2.7.1 Definition of Hb concentration change

Optical data derived from emitter-detector pairs were analyzed using the modified Beer-Lambert law (Cope et al., 1988). According to the law, a change in absorbance at a given wavelength (ΔA_λ) can be represented as:

$$\Delta A_\lambda = \varepsilon_\lambda (\text{DPF}) L \Delta C \quad (1),$$

where ε_λ is the extinction coefficient at the given wavelength (λ), DPF is the differential path length factor, L is the interoptode distance (= 3 cm), and ΔC is the concentration change in Hb. The concentration changes in oxy- and deoxy-Hb were calculated from the difference in the light absorbance (Okamoto et al., 2004b). Furthermore, the change in total-Hb was determined as the sum of the changes in oxy- and deoxy-Hb.

We discarded the trials in which children had failed to say all items. Furthermore, we removed the trials that included motion artifacts during the encoding and retention periods. The artifacts were detected by the analysis of abrupt changes in the raw NIRS data. For each channel, we averaged the NIRS data across trials of each condition for oxy-, deoxy- and total-Hb (Fig. 2). To suppress noise in the data, we used a three-point triangular smoothing filter, and we repeated this 10 times. We analyzed the NIRS data of the encoding and retention periods, which reflect recoding and rehearsal processes. We did not analyze the retrieval and rest periods as they often contained motion artifacts

produced by temporal muscle contraction in saying the items aloud and by body movements. For many children, the averaged data decreased during the first few seconds because the hemodynamic changes often did not return completely to baseline within the 6 s rest. Thus, Hb concentration changes (ΔC) were defined as maximum values minus minimum values within the time window of the encoding and retention periods (12 s). We then obtained the oxy-, deoxy- and total-Hb concentration changes for visual (ΔC_V) and auditory (ΔC_A) conditions. However, as these raw Hb change data cannot be quantified, they cannot be compared between subjects or between channels even within a subject because the exact differential path length factor (DPF) is not measurable (Hoshi, 2007).

2.7.2 Difference in Hb change between visual and auditory conditions

We investigated the difference in cortical activation between visual and auditory conditions, irrespective of age or the Digit Span score. Because analysis of raw data may be problematic, we utilized the method introduced by Sanefuji et al. (2007). The activations in visual and auditory conditions were represented by the following indices:

Visual index (VI) = $\Delta C_V / (\Delta C_V + \Delta C_A)$ and Auditory index (AI) = $\Delta C_A / (\Delta C_V + \Delta C_A)$.

Given that DPF is constant across measurements within a channel, these indices were independent of DPF as it was included in both the numerator and denominator (see Eq. (1)). Although the raw data (ΔC_V and ΔC_A) were highly scattered, likely because of variability in DPF or Hb changes among children, these indices (VI and AI) were less

scattered and approximately normally distributed, enabling parametric analysis (Supplementary Fig. 2). In addition to the resolution of the problem in quantification of Hb changes, these indices might be beneficial for detecting differences as parametric tests have more statistical power than nonparametric tests with small samples of a Gaussian population (Motulsky, 1995). For each corresponding channel across subjects, the difference between VI and AI was compared using Student's paired *t*-tests (two-tailed).[†] The levels of significance were adjusted to $p = 0.0125$ with Bonferroni correction for four channels.

2.7.3 Correlation analysis with age or memory ability

We investigated the relationship between the activation difference for visual material compared with auditory material and age or memory ability. To quantify the activation difference, we introduced another index which was represented as: Visual-auditory index (VAI) = $VI - AI = (\Delta C_V - \Delta C_A) / (\Delta C_V + \Delta C_A)$. VAI is independent of DPF, as for VI and AI. Next, we examined the coefficients of correlation between VAI and age or the Digit Span score. The correlation coefficients were calculated using Spearman's rank correlation (two-tailed),[†] because the age and the Digit Span score were not normally distributed. The adjusted levels of significance were $p = 0.0125$ with Bonferroni

[†] We also analyzed the raw data to test differences in activation between visual and auditory conditions (two-tailed Wilcoxon's signed rank test) and the correlation between the activation difference and age or the Digit Span score (two-tailed Spearman's rank correlation). However, the results were the same as those for the indices (data not shown).

correction.

2.7.4 Time course difference between visual and auditory conditions

We investigated the time course differences of hemodynamic changes between visual and auditory conditions. We predicted that visual material would require more time associated with recoding and rehearsal than for auditory material, which needs only rehearsal. The time (T) associated with recoding and/or rehearsal was defined as the difference in time between maximum and minimum values of Hb concentration change within the time window of the encoding and retention periods (12 s; Fig. 2). Next, we obtained the time for visual (T_V) and auditory (T_A) conditions. For each corresponding channel across subjects, the difference between T_V and T_A was compared using Wilcoxon's signed rank test (one-tailed). The adjusted significant levels were $p = 0.0125$ with Bonferroni correction.

Statistical analysis was performed using PASW 18.0 for Windows (SPSS, Tokyo, Japan).

3. Results

3.1 Behavioral data

The Digit Span scores (mean [SD]: 10.9 [2.3]) were appropriate to the children's ages (mean [SD]: 7.5 [1.6] years) as a score of 11 corresponds to 7.8 years of age according to the standardized data. There was positive, but not significant, correlation between age and the score (one-tailed Spearman's $r = 0.313$, $p = 0.096$, Supplementary Fig. 3).

Behavioral data of the NIRS task and its correlation with age or the Digit Span score were evaluated by nonparametric tests (two-tailed Wilcoxon's signed rank test and one-tailed Spearman's rank correlation). There were no significant differences between visual and auditory conditions in task performance (mean: 0.89 and 0.94, respectively, $p = 0.167$) and the number of analyzed trials (mean: 15.4 and 14.4, respectively, $p = 0.107$). The total task performance (mean: 0.91) was significantly correlated with the Digit Span score ($r = 0.534$, $p = 0.009$), but not with age ($r = 0.273$, $p = 0.129$). The correlations of task performance with the Digit Span score were significant in auditory condition ($r = 0.547$, $p = 0.008$), but not in visual condition ($r = 0.296$, $p = 0.109$; Supplementary Fig. 4). The numbers of presented (mean: 37.3) and analyzed (mean: 29.8) trials were not significantly correlated with age or the Digit Span score ($p > 0.2$).

3.2. NIRS data

3.2.1 Difference in Hb change between visual and auditory conditions

We investigated the difference in activation between visual and auditory conditions without the effect of age or the Digit Span score. No channels exhibited significant differences between VI and AI ($p > 0.15$) for oxy-, deoxy- and total-Hb.

3.2.2 Correlation analysis with age or memory ability

We investigated the relationship between the activation difference for visual condition compared with auditory condition and age or memory ability (Supplementary Table 1).

No channels showed significant correlations between VAI and age ($p > 0.05$) for oxy-,

deoxy- and total-Hb. However, the left VLPFC exhibited a significant positive correlation between VAI and Digit Span score for total-Hb ($r = 0.772, p < 0.001$). Partial correlation analysis adjusted for age showed similar results ($r = 0.719, p < 0.001$). Scatter diagrams indicated that the higher Digit Span score was associated with the greater VAI in the left VLPFC (Fig. 3). The diagrams also demonstrated that DLPFC showed different patterns compared with VLPFC. No channels showed a significant correlation between VAI and task performance ($p > 0.05$) for oxy-, deoxy- and total-Hb.

3.2.3 Time course difference between visual and auditory conditions

We compared the time courses between visual and auditory conditions. The left VLPFC showed a longer time course in visual (T_V) than in auditory (T_A) conditions for total-Hb ($p = 0.010$; Supplementary Table 2, Fig. 4). The difference was unlikely to be explained by the difference in the magnitude of Hb concentration changes between the two conditions, as shown above.

4. Discussion

By using event-related NIRS we explored the relationship between phonological recoding and rehearsal strategy for pictures and memory ability during childhood. The strategy was estimated by the activation difference between pictures (visual condition) and spoken words (auditory condition) in the left VLPFC including Broca's area. We found no effect of condition; i.e., children did not generally show higher activation of phonological strategy for the pictures. Some children displayed higher activation for the

visual than the auditory condition, while others exhibited lower activation. However, the activation difference for visual as compared with auditory condition (VAI) in the left VLPFC was positively correlated with the Digit Span score but not with age, suggesting that children with higher memory ability make more use of a phonological recoding and rehearsal strategy for pictures, as indicated by the brain activation pattern. To the best of our knowledge, this is the first functional neuroimaging study to provide functional evidence for this relationship.

The correlation of the phonological strategy with high memory ability is supported by a behavioral study investigating visual and phonological similarity effects, where children using phonological strategy tended to have higher memory ability than children relying on visual strategy, within the same age, or sometimes irrespective of ages (Palmer, 2000). In that study the close relationship between memory strategy and memory ability was demonstrated using behavioral methods, while in our study, we demonstrated this phenomenon using the left VLPFC activation that reflected the memory strategy. In addition, the study by Palmer (2000) suggested that the strategy change forms a continuum from visual encoding through a dual (visual and phonological) code to phonological encoding. This finding corresponds to the continuous change of the activation difference for visual material compared with auditory material in the left VLPFC in the present study (see Fig. 3).

The association of the phonological strategy for pictures with high memory ability

may be explained by maturation of rehearsal process. In the phonological loop, rehearsal process is a major determinant of memory ability both for children and adults (Baddeley, 1986; Hulme et al., 1984), and matures qualitatively as well as quantitatively during childhood (Gathercole and Hitch, 1993). As previously reported, the phonological loop is better at maintaining data including serially-presented material, while the visuospatial sketch pad is better at maintaining simultaneously-presented material (Baddeley, 1986). Therefore, in memory of serially-presented pictures, children with mature rehearsal may show high memory ability and tend to use the phonological strategy even with additional recoding process. On the other hand, children with primitive rehearsal may exhibit low memory ability and be likely to rely on the non-phonological strategy without the effortful recoding process.

The longer time course for visual compared with auditory conditions in the left VLPFC may reflect the recoding process. An event-related potential study investigated between-modality differences in spatio-temporal patterns of brain activity during retention of auditorily- or visually-presented verbal information (Ruchkin et al., 1997). In that study the onset of left frontal negativity, which reflects, at least in part, a rehearsal process, was later for visual than auditory presentation. These findings are supported by an anatomical-functional model of phonological short-term memory (Vallar, 2006). Auditory material has direct and automatic access to the phonological loop, and is rehearsed in the left VLPFC including Broca's area. By contrast, visual material requires

an additional recoding process, grapheme-to-phoneme conversion, in the left VLPFC to enter the loop and then is rehearsed.

Although we found no correlation between the task performance and the activation difference that represented the phonological strategy, the Digit Span score correlated with task performance in the auditory condition as well as with the activation difference. The correlation with the auditory task performance suggests an analogy between the Digit Span subtest and the auditory NIRS task, both of which require children to repeat a sequence of items that were presented aloud to them. The lack of correlation with the visual task performance may be accounted for by the additional process such as visual processing or recoding for pictures in the visual NIRS task, although the task and the Digit Span subtest were the same in memory of sequential items.

In our study, DLPFC, unlike the left VLPFC, showed no significant correlations between the activation difference for visual material compared with auditory material and memory ability. The disparity between VLPFC and DLPFC may reflect the distinct roles of these regions in verbal short-term memory, as demonstrated by functional neuroimaging studies. The left VLPFC is associated with recoding and rehearsal processes (Awh et al., 1996; Henson et al., 2000; Paulesu et al., 1993; Smith et al., 1998), although Henson et al. (2000) suggested the dissociation between the two processes. DLPFC is involved in various processes including encoding (Rypma and D'Esposito, 1999) and active maintenance of information (Cohen et al., 1997). There is little

information on the developmental change in DLPFC and VLPFC for short-term memory in childhood. Our data suggest that there is no apparent developmental difference in DLPFC activation between visual and auditory presentation.

There is controversy as to whether oxy-, deoxy- and total-Hb using NIRS are more closely related to cortical neuronal activity. Deoxy-Hb is suggested to be much less sensitive than oxy- and total-Hb (Strangman et al., 2002). Simultaneous measurements of NIRS and PET demonstrated that changes in total-Hb are physiologically coupled to changes in regional cerebral blood flow (Hock et al., 1997; Villringer et al., 1997). However, a perfused rat brain model showed that oxy-Hb was the most sensitive indicator of these changes (Hoshi et al., 2001). In the present study, we only found effects for total-Hb, suggesting that total-Hb may be a better marker of cortical neuronal activity than oxy-Hb for event-related NIRS, as previously suggested (Sanefuji et al., 2007). A study on temporal hemodynamics with event-related design also demonstrated that cerebral blood flow measured by arterial spin labeling MRI had the highest correlation with total-Hb monitored by NIRS (Huppert et al., 2006).

There are some limitations associated with the NIRS methodology used in our study. First, we investigated only VLPFC and DLPFC, but not other brain areas such as the posterior parietal cortex which mediates phonological short-term storage (Awh et al., 1996; Paulesu et al., 1993), as there was no anatomical marker corresponding to the cortex in the cranio-cerebral correlation study (Okamoto et al., 2004a). Furthermore,

placement of many channels to cover wide areas requires prolonged times, during which the children can tire before experimentation. Second, the channel position might not exactly correspond to the regions of interest in our study of 5 to 11 year old children because of morphologic differences in brain structures of children (Gaillard et al., 2001). The cranio-cerebral correlation is not yet established for children. However, normalization into a standard stereotactic space is acceptable for children over 6 years of age even in statistical parametric mapping analysis (Muzik, 2000) for PET or functional MRI, which has much higher spatial resolution than for NIRS of 2-3 cm (Villringer et al., 1993). Third, unlike PET and functional MRI, there are no standard methods of NIRS data analysis (Hoshi, 2007). Analyses of both the raw data and the DPF-independent indices (i.e., the proportion of Hb changes under visual and auditory conditions) produced similar results, but may vary under specific circumstances. We previously reported that activation differences between Hit and CR in R15 (Sanefuji et al., 2007) were significant for the indices ($p = 0.016$, paired Student's *t*-test), but not for the raw data that were highly scattered and skewed ($p = 0.053$, Wilcoxon's signed rank test). The index-based analysis may be beneficial for statistically modest results (Meeting of Japan Optical Functional Brain Imaging Society, 2006, Tokyo). Finally, in the present study, we did not separate the direct left VLPFC activation of rehearsal or recoding process, and thus did not demonstrate their individual developmental changes. As described above, we can only obtain proportional activation due to problems in quantification of Hb changes

(Hoshi, 2007). Future technical advances in NIRS will resolve these limitations.

In conclusion, we provide functional evidence that strategy in short-term memory for serially-presented pictures shifts gradually from a non-phonological to phonological method as memory ability increases in childhood. The present paradigm revealed a functional role of the articulatory control process and may shed light on neural mechanisms of autism and Down syndrome, which exhibit dysfunction of this process.

Acknowledgments

We thank Mr. Yasuhiro Maeda for permission to use the children-familiar pictures. We also thank the parents and children who participated in the study. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (No. 21790995, No. 20591225, No. 70304805).

References

- Alloway, T. P., & Gathercole, S. E. (2006). *Working memory and neurodevelopmental disorders*. New York: Psychology Press.
- Amano, S., & Kondo, T. (1999). *NTT database series Nihongo-no Goitokusei (Lexical properties of Japanese)*. Tokyo: Sanseido.
- Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppel, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. *Psychol Sci*, 7(1), 25-31.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat Rev Neurosci*, 4(10), 829-839.
- Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.
- Casey, B. J., Cohen, J. D., Jezzard, P., Turner, R., Noll, D. C., Trainor, R. J., et al. (1995). Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. *Neuroimage*, 2(3), 221-229.
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, 386(6625), 604-608.
- Cope, M., Delpy, D. T., Reynolds, E. O., Wray, S., Wyatt, J., & van der Zee, P. (1988). Methods of quantitating cerebral near infrared spectroscopy data. *Adv Exp Med Biol*, 222, 183-189.

Gaillard, W. D., Grandin, C. B., & Xu, B. (2001). Developmental aspects of pediatric fMRI: considerations for image acquisition, analysis, and interpretation. *Neuroimage*, 13(2), 239-249.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, 3, 410-418.

Gathercole, S. E., & Hitch, G. J. (1993). Developmental changes in short-term memory: a revised working memory perspective. In A. Collins, S. E. Gathercole, M. A. Conway & P. E. Morris (Eds.), *Theories of memory* (pp. 189-210). Hove: Erlbaum.

Halliday, M. S., Hitch, G. J., Lennon, B., & Pettipher, C. (1990). Verbal short-term memory in children: The role of the articulatory loop. *European Journal of Cognitive Psychology*, 2, 23-38.

Henson, R. N., Burgess, N., & Frith, C. D. (2000). Recoding, storage, rehearsal and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia*, 38(4), 426-440.

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London Series B*, 302, 325-340.

Hitch, G. J., Woodin, M. E., & Baker, S. (1989). Visual and phonological components of working memory in children. *Memory and cognition*, 17(2), 175-185.

Hock, C., Villringer, K., Muller-Spahn, F., Wenzel, R., Heekeren, H., Schuh-Hofer, S., et al. (1997). Decrease in parietal cerebral hemoglobin oxygenation during performance of

a verbal fluency task in patients with Alzheimer's disease monitored by means of near-infrared spectroscopy (NIRS)--correlation with simultaneous rCBF-PET measurements. *Brain Res*, 755(2), 293-303.

Hoshi, Y. (2007). Functional near-infrared spectroscopy: current status and future prospects. *J Biomed Opt*, 12(6), 062106.

Hoshi, Y., Kobayashi, N., & Tamura, M. (2001). Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model. *J Appl Physiol*, 90(5), 1657-1662.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of experimental child psychology*, 38, 241-253.

Huppert, T. J., Hoge, R. D., Diamond, S. G., Franceschini, M. A., & Boas, D. A. (2006). A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage*, 29(2), 368-382.

Japanese WISC-III Publication Committee. (1998). *Nihonban WISC- β chinou kensahou (the Japanese version of the Wechsler Intelligence Scale for Children-Third Edition)*. Tokyo: Nihon Bunka Kagakusha.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalogr Clin Neurophysiol*, 10, 367-380.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K., & Gruber, O. (2009).

Functional architecture of verbal and tonal working memory: an fMRI study. *Hum Brain Mapp*, 30(3), 859-873.

Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural attunement processes in infants during the acquisition of a language-specific phonemic contrast. *J Neurosci*, 27(2), 315-321.

Moriguchi, Y., & Hiraki, K. (2009). Neural origin of cognitive shifting in young children. *Proc Natl Acad Sci U S A*, 106(14), 6017-6021.

Motulsky, H. (1995). 3 The standard deviation. In H. Motulsky (Ed.), *Intuitive Biostatistics* (pp. 22-30). New York: Oxford University Press.

Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., et al. (2004a). Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10-20 system oriented for transcranial functional brain mapping. *Neuroimage*, 21(1), 99-111.

Okamoto, M., Dan, H., Shimizu, K., Takeo, K., Amita, T., Oda, I., et al. (2004b). Multimodal assessment of cortical activation during apple peeling by NIRS and fMRI. *Neuroimage*, 21(4), 1275-1288.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.

Palmer, S. (2000). Working memory: a developmental study of phonological recoding. *Memory*, 8(3), 179-193.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, 362(6418), 342-345.

Ruchkin, D. S., Berndt, R. S., Johnson, R., Jr., Ritter, W., Grafman, J., & Canoune, H. L. (1997). Modality-specific processing streams in verbal working memory: evidence from spatio-temporal patterns of brain activity. *Brain Res Cogn Brain Res*, 6(2), 95-113.

Rypma, B., & D'Esposito, M. (1999). The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. *Proc Natl Acad Sci U S A*, 96(11), 6558-6563.

Sanefuji, M., Nakashima, T., Kira, R., Iwayama, M., Torisu, H., Sakai, Y., et al. (2007). The relationship between retrieval success and task performance during the recognition of meaningless shapes: an event-related near-infrared spectroscopy study. *Neurosci Res*, 59(2), 191-198.

Schroeter, M. L., Zysset, S., Kupka, T., Kruggel, F., & Yves von Cramon, D. (2002). Near-infrared spectroscopy can detect brain activity during a color-word matching Stroop task in an event-related design. *Hum Brain Mapp*, 17(1), 61-71.

Smith, E. E., Jonides, J., Marshuetz, C., & Koeppel, R. A. (1998). Components of verbal working memory: evidence from neuroimaging. *Proc Natl Acad Sci U S A*, 95(3), 876-882.

Strangman, G., Culver, J. P., Thompson, J. H., & Boas, D. A. (2002). A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain

activation. *Neuroimage*, 17(2), 719-731.

Vallar, G. (2006). Memory systems: the case of phonological short-term memory. A festschrift for cognitive neuropsychology. *Cognitive Neuropsychology*, 23, 135-155.

Villringer, A., Planck, J., Hock, C., Schleinkofer, L., & Dirnagl, U. (1993). Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neurosci Lett*, 154(1-2), 101-104.

Villringer, K., Minoshima, S., Hock, C., Obrig, H., Ziegler, S., Dirnagl, U., et al. (1997). Assessment of local brain activation. A simultaneous PET and near-infrared spectroscopy study. *Adv Exp Med Biol*, 413, 149-153.

Figure legends

Fig. 1. Schematic illustration of a trial.

Children-familiar items were either presented visually as pictures (visual condition) or auditorily as spoken words (auditory condition; see section 2.6 for details).

Fig. 2. NIRS waveforms for total-Hb in the left VLPFC of one child.

Each waveform was averaged across trials of visual (dark brown) or auditory (green) condition. The Hb concentration at the onset of each trial ($= 0$ s) was defined as the initial concentration ($= 0$ $\mu\text{M}\cdot\text{cm}$). The definition of the Hb concentration change (ΔC) and the time (T) are shown for the auditory condition.

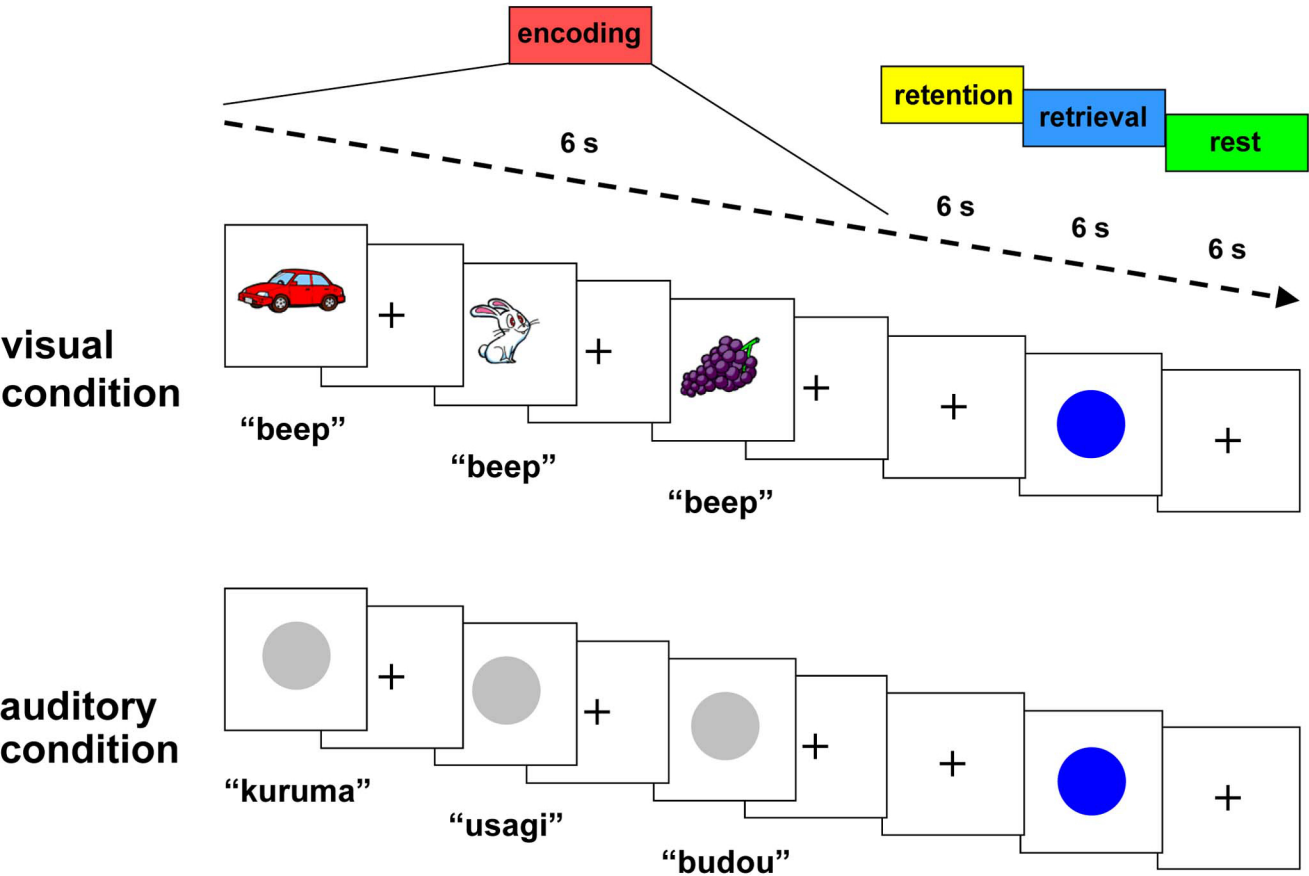
Fig. 3. Correlations between visual-auditory index (VAI) and Digit Span score for total-Hb in the four channels.

Linear regression lines are depicted in the diagrams. Approximate channel positions are superimposed on a MRI scan of a 6 year old child.

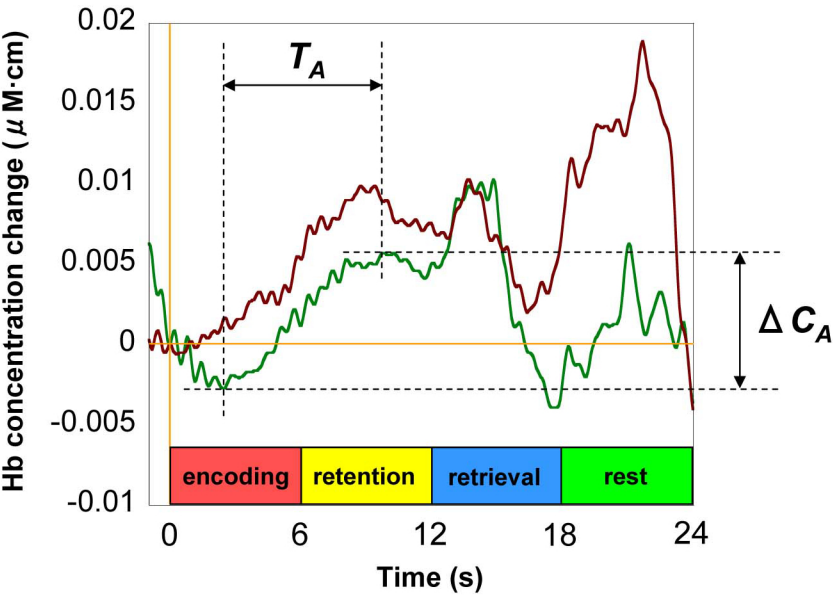
Fig. 4. Time course differences between visual (T_V) and auditory (T_A) conditions.

The box plots illustrate medians (horizontal lines), the first and third quartiles (box boundaries), and 1.5 times the upper and lower quartiles (whiskers).

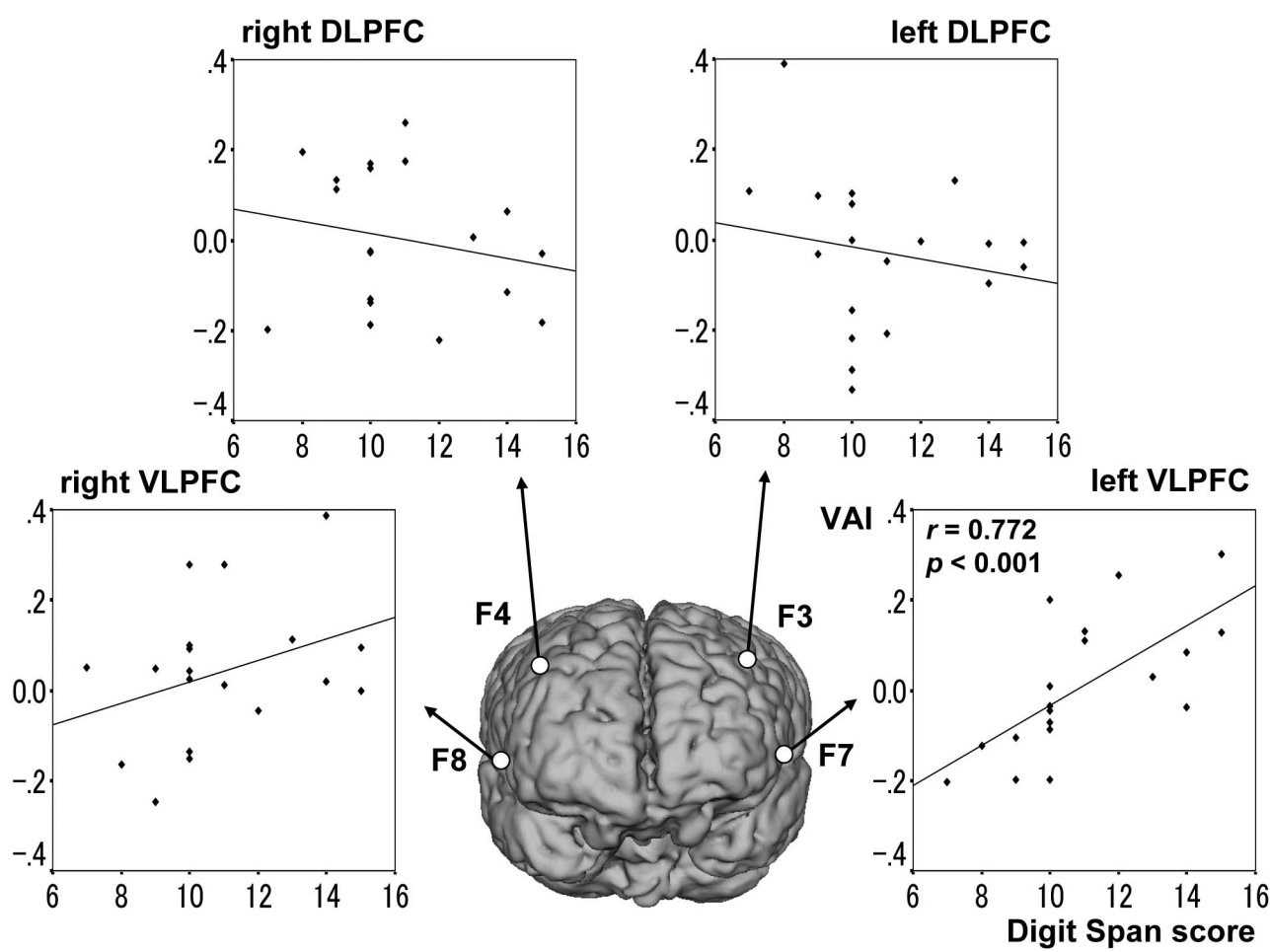
(Fig.1)



(Fig.2)



(Fig.3)



(Fig.4)

