Information from Speech and Gesture is Integrated When Meanings of New Words are Categorized in Normal Young Children, but not in Children with Williams Syndrome

Nobuo Masataka

Six normal young Japanese children and six children with Williams syndrome (WMS) were presented a nonsense label to identify an unfamiliar target solid object that was either rigid (e.g., steel) or flexible (e.g., rubber). The presentation was accompanied with or without a gesture. When the gesture was performed, it emphasized either the shape and function (e.g., a rolling gesture), or the material (e.g., squeezing) of the target. Thereafter the children were presented with a pair of unfamiliar objects, one of them matched with the target in shape and function the other in material. They were asked, given the same label, to choose an object that matched the target in shape and function, or in material. When a gesture accompanied the preceding presentation of the target, the normal children were likely to choose that object which matched the target in shape and function if the gesture had emphasized the shape and function of the target, but tended to choose the object that matched the target in material if the gesture had emphasized the target material. When no gesture was provided, they chose either of the two objects at random. In contrast, the choices of the children with WMS were not influenced by the type of preceding gesture, but were determined randomly regardless of the presence of the gesture. These findings are discussed in terms of the possible impairment of working memory in children with WMS, which would make their language acquisition unique.

Keywords: word learning, gesture, categorization, Williams syndrome

1.0 Introduction

The traditional view of gesture has conventionally assumed a discontinuity between preverbal or nonverbal behavior, the activity typically expressed by the word ‘gesture’, and verbal production. Contrary to this dichotomous classification, McNeill (1985, 1992) proposed a close linkage between gesture and speech in which both should be regarded as aspects of a single underlying process of mental representation. According to his theory, both manifest themselves as a product of interaction between thought and language, with more imagistic information being coded in gesture and more semantic information in speech.

McNeill refers to several findings to support this notion. In particular, he stresses the synchronous character of gesture with speech. Almost all gestures under observation are reported to be temporally accompanied by speech. They convey semantically closely related meanings as those transmitted by the accompanying verbal utterances. Moreover, the specific linguistic segments that are coexpressive with a gesture are in fact cotemporal. He also argues that the time course of development of gesture and speech in children is identical. A child’s first gestures are reported to be concrete pointing, followed by other kinds of iconics, beats, metaphorics, and last of all, abstract pointing. This progression follows the same path as the development of speech, because children’s speech develops from
a largely referential focus, through descriptive elaboration, and finally to the ability to structure discourse. Nevertheless, McNeill’s description remains anecdotal; quantitative data are lacking and no explanation has been provided as to how or why gesture and speech come to be developmentally related to one another.

The present article is a preliminary attempt to answer this question. The data presented here indicate the possibility that the learning of meanings of new words in young children is enormously enhanced at an early stage of language development precisely by the synchronous nature of gesture and speech.

2.0 Segmentation and Mapping Problems in Word Learning

Word learning by young children is not a straightforward process. Typically, a caregiver of a child holds up a ball and says “This is a ball”, at which the child may appear to understand what the word ‘ball’ designates. Similarly, the word ‘cat’ is produced in the presence of a cat, and so on. At first glance, the child may appear to take the word uttered to be the label for the object pointed out, and there the matter ends. However, appropriate apprehension of the meanings of words presents the child with several difficulties. First, the child is required to find the word within the stream of speech. Whereas in written English individual words are separated by spaces, such that a word as a minimum unit constituting a sentence can be easily identified, in spoken language the phrase used to teach the meaning of ‘ball’ is “This is a ball.” The child is somehow required to find the word boundaries, a problem which has been referred to as the speech segmentation problem. Although a number of studies have investigated how a child tackles this problem, their findings are not to be described in detail here. What is noted is simply the fact that even if the child successfully identifies the sound sequence “ball” as a word, another problem yet remains. If a simple story of word learning is to work, the child must have the concept of things that are labeled by words in the language and lack only the label for them. Still, the question remains as to how the child knows that it is the ball that is being labeled rather than the ball’s color, the material the ball is made of, or the shape of the ball. Or perhaps the word being uttered in the presence of a ball is not a label at all but a command to do something, like play. The possibilities are of course limitless. This problem is known as the mapping problem, and is to be treated in this article.

This study is concerned with the issue of categorization by young normal children and by children with Williams syndrome (WMS). It addresses the question of what influences their categorization of novel referents presented in conjunction with new labels. It examines to what extent this categorization is affected by information coming from speech itself and that coming from gesture which accompanies the speech. By comparing the manner in which information affects categorization between normal and WMS children, the degree of flexibility with which categorization occurs is also examined.

2.1 Internal Constraints or Biases as a Possible Solution of Mapping Problem

The mapping problem was first formulated by the philosopher Quine (1960). He described the child’s problem as follows: an infinite number of hypotheses about word meaning are logically possible given the data the child has. Yet children tend to figure out the meaning of the words they hear. How do they do so? Since that time, there has been a growing literature on categorization by children, most of which has attempted to answer the question of how knowledge of the functional properties of an object is acquired, because this essentially concerns their comprehension of relationships between the name of the object and its referents. For instance, to relevantly understand what the label ‘cat’ designates, it is
insufficient to merely have the knowledge that the word refers to an attribute of ‘cat’, a momentary state of it, or some action it is undertaking. Children are required to know that the word refers to a whole object and to a class. This seemingly formidable task is in fact usually achieved very rapidly and successfully, long puzzling various psycholinguists and developmental psychologists since Quine (1960). Several explanations have been proposed, each relying on a different source of potential help to the child, on the basis of the assumption that word learning would be facilitated if children did not have to consider all of many possible meanings each time they heard a new word. They contend that children actually do not take into account all possible meanings but instead enter word learning situations with several assumptions about how the lexicon works. Since these assumptions limit, or constrain, the possible interpretations of new words that children must consider, several researchers have proposed that children observe certain constraints or show certain biases or strategies that narrow the possibilities they will initially entertain for the meanings of new words. These biases or principles are purported to facilitate acquisition by enabling the child to make a good first guess at the meanings of new words. The best-known conceptualization among such attempts would probably be Markman’s (1991) three proposed constraints on word learning, which have been the subject of a great deal of research; namely, the whole-object assumption, the taxonomic assumption, and the assumption of mutual exclusivity.

The whole-object assumption is the child’s assumption that words refer to whole objects. According to this proposal, children assume that every new word they hear refers to some whole object rather than to part of an object or to a property of the object. This eliminates “color”, “shape” and other attributes of a ball as meanings of ball. The presence of this assumption is confirmed from various experiments concerning word learning (e.g., Markman & Wachtel, 1988) as well as from errors young children make. For example, children commonly think that hot is the label for stove on the basis of the common experience of hearing, “Don’t touch it. It is hot”, in reference to the stove.

The taxonomic principle is the assumption that words are understood in terms of things of the same kind. The argument presupposes the existence of knowledge of a taxonomy, a system of classifying things into categories, in each child. By such knowledge, the child is led to think that cat will also refer to other cats, but not to things that are thematically related to cats such as collars, leashes, or bones. Evidence that children actually operate according to the taxonomic assumption in word learning has been provided by the finding that preschool children override their inclination to group thematically related things when presented with a new word (Markman & Hutchins, 1984).

Namely, in this study, a puppet was used when presenting a picture (e.g., a dog) to 2- to 3-year-old children, who were then presented a choice between a thematically related item (dog food) and a taxonomically related item (a dog of a different breed). In the “no word” condition, the puppet pointed to the first picture (of a dog) and said, “Look carefully now. See this?” When the next two choices were presented (the other dog and the dog food), the puppet said “Find another one that is the same as this.” In the “word” condition the puppet introduced the first picture by saying, “See this? It is a sud.” When presenting the other choices, the puppet said, “Find another sud that is the same as this sud.” Children in the “no word” condition were found to choose the taxonomically related choice 59% of the time. On the other hand, children in the “word” condition made the taxonomic choice 83% of the time. Even though these children think that thematically related things such as dogs and dog food “go together” (41% of the time), they do not regard words as labelling things that “go together”.
Children assume that words label things that are of the same kind.

The mutual exclusively assumption is the assumption that different words refer to different kinds of things. Thus, for example, members of the category labeled cat do not overlap with members of the category labeled dog. Evidence that suggests children in fact operate according to such an assumption comes from experimental studies in which children are given an array of familiar objects for which they have a label and one object that is novel and nameless. Given the instruction to “Show me the x,” where x is some nonsense syllable chosen to function as a new word, children pick the object for which they do not yet have a word. The response is interpreted to mean that the children assumed that the new word was not a synonym for any of the words they already know (Markman & Wachtel, 1988). Another consequence of the mutual exclusivity assumption is that it provides a basis for overriding the whole-object assumption, which children, having knowledge of what the word cup means, must do in order to learn terms for parts and properties of objects. By this a child is enabled to look for something else as the referent of the new term, and not, if its mother says, “This is a handle,” to take handle to be a synonym for cup.

2.2 Socio-Pragmatic Principles as an Alternative Solution

Even with the three constraints, however, word learning still appears extremely difficult for children. In particular, the taxonomic constraint is too ambiguous for systematic classification of objects. For, given an object such as an apple, the principle advises children “not to attend to objects that have thematic relations with the apple, not to exclusively attend to the apple, but to attend to the taxonomic relations of the apple, and to call only related objects under the same name, apple”. Nevertheless, the principle never advises the children what referents are related. Nor do the other two principles do so. By the principles themselves, nothing concrete is intrinsically provided as an actual content of criterion for classification by the children.

Meanwhile, it has been argued that the manner in which adults talk to children makes word meaning in context much less ambiguous than Quine had assumed. Observations of parental talk to children reveal that caregivers tend to label things the child is looking at. As a consequence, the child hears words when looking at things the words label. Children learn object labels better when the caregiver follows the child’s focus of attention and labels what the child is looking at rather than when they try to redirect the child’s attention or label things that are not in the child’s attentional focus (Dunham, Dunham, & Vurwin, 1993; Tomasello & Todd, 1983). Children apparently use the correspondence between the speech they hear and other information concerning the nonlinguistic context as a basis for figuring out the meanings of new words. To the extent that adults ensure that the words children hear will match their perception of the nonlinguistic context, the nature of the input facilitates word learning, and correspondingly, the mapping problem is attenuated.

Of course, correspondence between speech and context is not in and of itself sufficient for appropriate word learning. It is not enough that a child merely match words with the nonlinguistic context; he must also match words relevantly with the speaker’s intended meaning. For this task, the presence of some internal principles should be assumed. These might include the three constraints mentioned above, the content of which needs to be qualified to some extent by socio-pragmatic principles. This is because various lines of evidence show that children are capable of inferring a speaker’s semantic intentions socio-pragmatically, and then use these inferences to assign meaning to newly introduced words. For instance, Baldwin (1993) found that by 18 months of age, children are able to follow a speaker’s gaze and use it as a clue to word
meaning, while Akhtar, Carpenter, & Tomasello (1996) found that 2-year-olds will infer that a novel label produced with an expression of surprise refers to an object that is also novel to the speaker.

Moreover, Tomasello & Barton (1994) found that 2-year-olds could distinguish accidental from intentional actions. In their study, if an experimenter said, “Let’s go twang it” and then did something accidentally (e.g., clumsily) followed by doing something intentionally, children took twang to refer only to the intentional action. The input view for the word learning underlying this study is basically that information necessary for the task is available in the environment. The social behavior of others is added to the richness of the environment. The view adds children’s understanding of other people to their abilities to pick up that information.

Recently this notion was clearly demonstrated experimentally by Kobayashi (1997). She examined whether 2-year-olds attend to the relevant properties (shape or material) of solid objects in making inferences when adults demonstrate appropriate action. In the first experiment, the investigator presented 24 Japanese 2-year-olds a nonsense label with an unfamiliar target solid object that was either rigid (e.g., steel) or flexible (e.g., sponge). The investigator performed an action on the object that emphasized the shape or the material of these targets, and then asked the children, given the same label, to choose an object that matched the target in shape or in material. The results showed that the children used the information about the type of action the experimenter performed in order to make their choices regardless of the presence of labels: if the action performed emphasized the shape of the object, the children chose the target so as to match it in shape subsequently, and vice versa.

Kobayashi concluded that an adult’s information about action alone serves as a basis to direct children to attend to relevant object properties in making inferences about solid objects. In other words, in solving the mapping problem, the information children glean from the environment comes mostly from the action others perform, simultaneous with the utterance of a word novel to the children. In this situation, whether the action is a pragmatic behavior or not is important because what is crucial is that certain elements of the action be decoded by the child as inferring the meaning of the message the other is attempting to convey to it. Initially, the meaning of a novel object must be encoded in the speech and in the simultaneously occurring action. This in turn gives rise to the possibility that the “action” described in this socio-pragmatic approach to the mapping problem (e.g., Kobayashi, 1997) can be rephrased as the gesture, as argued by McNeill (1985, 1992). As he noted, gesture and speech occur synchronously, and are both realized as a product of the underlying interaction between thought and language. From this, while speech conveys the message more semantically, gesture does it more imagistically. Taking this notion with the above findings about the mapping problem, a hypothesis can be formulated that owing to these characteristics, children might succeed in the relevant semantic categorization of words totally unfamiliar to them.

In order to explore this possibility, the following experimentation was conducted. It examined the effects of action information as a source of gesture occurring synchronously with speech on the categorization of meanings of new words, in both normal young children as well as children with Williams syndrome.

### 3.0 Background on Williams Syndrome

Williams syndrome (WMS) is a rare congenital disorder associated with a hemizygous deletion on chromosome 7 (7q11.23), the locus of a gene that programs elastin, a protein that gives elastic properties to arteries, lungs, intestine and skin, and accounts for the supravalvular aortic stenosis, hernias and premature aging wrinkles of the face that patients with this disorder ex-
perience. However, the interest of this genetic disorder to cognitive scientists and linguists does not stem from neurologic complications of aortic stenosis, but rather because the deletion on chromosome 7 in this disorder extends a variable distance beyond the boundaries of the elastin gene. Although some are as yet unidentified, the deleted pieces are assumed to contain approximately 20 genes, some of which are known to be expressed in the brain in normal development. Their defect has been considered to be associated with the mental retardation and unusual patterns of cognitive abilities and deficits characteristic to this disease. The cognitive hallmark of WMS is said to be a dissociation between language (relative strength) and visuo-spatial cognition (profound impairment). Consistent cognitive deficits are found in this syndrome, and typically, standard Full Scale IQ scores range from 40-100, with means of around 60. These include commensurate impairments in spatial, quantitative and reasoning abilities. Severe deficits in WMS have been documented in conceptual reasoning tasks, with most patients failing Piagetian tests (Piaget, 1963) for conservation of number, space, substance, weight, and quantity (Bellugi, Lichtenberger, Mills, Galaburda, & Kornberg, 1999; Karmiloff-Smith, 1998). In contrast to the overall impairment seen in this area of cognitive abilities, one striking aspect of the WMS profile is the strength in language abilities in adolescence and adulthood. In fact, the first individuals with WMS to be cognitively investigated ranged in age from 11 years to young adulthood. They all displayed a striking discrepancy between language skills and level of general cognitive functioning. Although IQs ranged from 40 to 70, they gave the impression of being even more verbal and more conversational than typically developing children. Adolescents with WMS tell coherent and complex stories with great emotional expression (Reilly, Klima, & Bellugi, 1990). They use advanced and unusual vocabulary in their spontaneous speech — words like surrender, nontoxic, commentator, and brochure — and, when asked, are capable of providing appropriate definitions of these terms. Their spontaneous language is grammatically complex in terms of both grammatical morphology and sentence structure, although with occasional errors of overgeneralizing morphology (Bellugi, Wang, & Jernigan, 1994).

Later research on WMS explored the early stage of language development, where it became apparent that WMS is associated with an initial delay in language development. First words usually emerged between 18 and 24 months of age whereas they are recorded around 12 months in normal children. However, when grammar begins to emerge, children with WMS diverge from the developmental course followed by age-matched and Full-Scale IQ-matched subjects with other mental retardations, and come to perform far better on a wide variety of grammar probes, such as reversible passives, negation, tag questions, sentence repetition, sentence completion, sentence correction and conditionals (Bellugi, Klima, & Wang, 1996; Bellugi et al., 1994). The findings appear to indicate a remarkable strength in language ability at all levels in older individuals with WMS in light of the cognitive deficits they have generally. In fact, the language abilities of WMS children, which are often at a level higher than their overall cognitive abilities, can often mislead their school teachers into thinking the children are in fact progressing.

3.1 Characteristics of Early Language Development in WMS

Findings on the relationship between language and cognition characteristics in WMS are controversial (Bellugi et al., 1999; Karmiloff-Smith, 1998). Some researchers consider the syndrome a remarkable example of the modularity of language as a system which is separate from general cognitive abilities. Others argue that as adults with WMS are said to function in some ways at the five- to seven-year-old level, there
is a sufficient substrate of cognitive abilities for the development of complex syntax, and that WMS does not thus represent a dissociation between language and general cognitive functions, and about the intactness of levels of language in WMS. Yet, researchers agree generally that language (for example, morphology and syntax) is a relative strength in WMS, in apparent contrast to other syndromes that involve mental retardation.

As a possible mechanism underlying such strength, an unusual degree of musicality and engagement with musical stimuli has been pointed out in many people with WMS. Anecdotal evidence implies that some Williams people possess extraordinary musical talent. Even though their attention span for most tasks is short, many will listen to music, sing and play instruments with astonishing persistence. Most cannot read musical notes, yet some have perfect or nearly perfect pitch and an uncanny sense of rhythm. One boy quickly learned to play an extremely complex drumbeat in 7/4 time with one hand while drumming in 4/4 time with the other hand. A number of individuals retain complex music for years, remembering melodies and verses of long ballads; one even sings songs in 25 languages. Experienced Williams musicians also sing harmonies, improvise and compose lyrics readily (Lenhoff, Wang, Greenberg, & Bellugi, 1997, pp. 42). Such anecdotes have recently led to the systematic investigation of musical ability in children with WMS. Don (1996) tested 19 individuals using the tonal and rhythm subtests of Gordon’s Primary Measures of Music Audiation (1986). “Audiation” is defined as the ability to hear music in one’s head, akin to visual imagery. His subjects demonstrated relatively preserved audiation ability, at levels commensurate with their relatively preserved language. Levitin & Bellugi (1998) assessed eight individuals with the syndrome in terms of rhythm production, and compared their achievements with those of normal children using an echo clapping task. Their results revealed that, despite serious deficits in other cognitive domains and generally poor coordination, the Williams subjects demonstrated equivalent abilities in meter change and beat maintenance.

Taken together, the available data indicate that hyperacusis in the auditory and, in particular, musical domain is significantly related to relative strength in language in WMS. Yet, as to how the hyperacusis relates to language development, virtually no work has been done. The present study was undertaken partly to answer this question.

4. Comparison of Word Learning between Normal Children and Children with WMS

4.1 Participants

The present study was conducted with 6 normal children (4 males and 2 females) and 6 children with WMS (4 males and 2 females) after informed consent was obtained from their parents. The mother of each child was a monolingual woman who spoke only Japanese, and was a full-time housewife between the ages of 24 and 31. Each of the 6 normal children was healthy and had uncomplicated prenatal and perinatal histories. Three were first-born and the remaining three were second-born. Mean age when tested was 27.3 months (range: 24-29).

Of the 6 children with WMS, three were first-born and the remaining three were second-born. Older siblings of the three second-born were healthy and normal. The parents of the six children were all normal and healthy. All 6 had received a diagnosis of WMS on the basis of clinical, physical and neuropsychological examinations. All had the facial dysmorphology typical of WMS. All agreed to undergo fluorescent in situ hybridization (FISH) testing for the genetic deletions implicated in WMS, and all results were positive. When tested, mean age was 40.2 months (range: 37-43) and the children were thus older than the normal children. The children with WMS were part of a much broader
Table 1 Eight object sets used in the experiment

<table>
<thead>
<tr>
<th>Set</th>
<th>Target</th>
<th>Test item 1</th>
<th>Test item 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Yellow, glass, egg-shaped object</td>
<td>Yellow, glass pyramid-shaped object</td>
<td>Yellow, styrofoam, egg-shaped object</td>
</tr>
<tr>
<td>B</td>
<td>Orange, steel, wing-bolt</td>
<td>Orange, steel, ruler</td>
<td>Orange, plastic, wing-bolt</td>
</tr>
<tr>
<td>C</td>
<td>White, rubber, ring</td>
<td>White, rubber, rope</td>
<td>White, metal, ring</td>
</tr>
<tr>
<td>D</td>
<td>Purple soap, imitating the figure of a character in a cartoon</td>
<td>Purple, common, soap</td>
<td>Doll of the same character as that of the target</td>
</tr>
<tr>
<td>E</td>
<td>White, paper, napkin</td>
<td>White, letter paper</td>
<td>White, clothe napkin</td>
</tr>
<tr>
<td>F</td>
<td>Red, rubber, frisbie</td>
<td>Blue, rubber, rectangular cube</td>
<td>Silver, plastic, frisbie</td>
</tr>
<tr>
<td>G</td>
<td>White, plastic, knife</td>
<td>Black, plastic, floppy disc</td>
<td>Iron knife with a wooden handle</td>
</tr>
<tr>
<td>H</td>
<td>Green, rubber, pot cleaner</td>
<td>Yellow, rubber, shoe sole</td>
<td>Stainless steal, wire, pot cleaner</td>
</tr>
</tbody>
</table>

The longitudinal study of the uneven cognitive profile in this syndrome in which they were entered at the age of 26 weeks, much earlier than the onset of first words. After entry, a series of observations was undertaken on two consecutive days biweekly, during which the cumulative lexicon was recorded. By the time the present study was conducted, the average number of words included in the lexicon of the six children was 224. This number is nearly equivalent to that of normal Japanese children at around 2 years (Ogura, 1999). Thus, the 2-year-old normal children and 3-year-old children with WMS included here were considered to be matched in terms of the size of lexicon they had acquired.

4.2 Methods

Overall Design of the Experiment

The experiment was undertaken essentially according to the protocol developed by Kobayashi (1997). Each experimental session consisted of two different phases, name training and testing. During the name training phase, each child was presented an unfamiliar solid object (referred to below as the ‘target object’) with a nonsense label. In the subsequent testing phase, two other solid, unfamiliar objects were presented to the child (referred to as ‘test objects’). Each of the test objects differed from the target object, but matched the target with respect to either material, or shape and function. The child was then instructed to choose either of the two, given the same label as that heard during the preceding name training phase. When the choice was made, the experimental session was regarded as having ended.

Stimulus: A total of 8 sets of stimuli were prepared, each consisting of a target object and two test objects (Table 1). Each child saw a set only once, thus undergoing a total of 8 experimental sessions throughout the entire study.

Procedure

The target objects of the 8 prepared sets of stimuli were presented to the children with one of the following labels: memu, ruto, rami, pito, muta, papu, omi, and agu. All are nonsense words and novel to Japanese speakers, but all are composed phonologically of syllables common to the Japanese language. Assignment of
Table 2  Gesture accompanying the presentation of targets in 8 stimulus sets

<table>
<thead>
<tr>
<th>Set</th>
<th>Material Emphasizing</th>
<th>Shape &amp; Function Emphasizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Pretend to hold the target in front of her eyes and to look through it at the child's face</td>
<td>Pretend to hold the target and roll it repeatedly in the air to express the movement of the egg rolling on the ground</td>
</tr>
<tr>
<td>B</td>
<td>Show the action of lifting the target as if it were heavy</td>
<td>Pretend to hold the target and to show the action of screwing it into the hole</td>
</tr>
<tr>
<td>C</td>
<td>Pretend to stretch the target</td>
<td>Pretend to hang the target around one's wrist</td>
</tr>
<tr>
<td>D</td>
<td>Pretend to wash hands with the target</td>
<td>Show the action of walking, represented by wriggling the index- and the middle-fingers</td>
</tr>
<tr>
<td>E</td>
<td>Show the action of tearing the target to pieces</td>
<td>Show the action of wiping one's mouth with the target</td>
</tr>
<tr>
<td>F</td>
<td>Pretend to bend the target</td>
<td>Show the action of spinning by the index-finger</td>
</tr>
<tr>
<td>G</td>
<td>Pretend to hold the target and look into it</td>
<td>Show the action of cutting with the target</td>
</tr>
<tr>
<td>H</td>
<td>Pretend to squeeze the target</td>
<td>Show the action of washing with the target</td>
</tr>
</tbody>
</table>

labels was varied among targets and the targets were presented in random order, but each child heard every label once. Typically, the researcher presented each child with one of the 8 targets (e.g., a yellow, glass, egg-shaped object), and then told them verbally in Japanese to look at it using the assigned label (e.g. in the case of muta “Muta wo mite ne”, renderable in English as “Look at muta”). The manner used in name training actually varied according to three different experimental conditions: presentation was accompanied with or without a gesture, and presentation with gesture was further divided into two prearranged gestures, one emphasizing the material of the target and the other its shape and function. These three are referred to below as ‘Naming Alone Condition’, ‘Material Emphasizing Condition’ and ‘Shape & Function Emphasizing Condition’, respectively.

For example, for name training involving stimulus set A in Table 1, the gesture displayed by the investigator in the Material Emphasizing Condition was to pretend to hold the target (a yellow, glass, egg-shaped object) in front of her eyes and to look through it at the child’s face. In the Shape & Function Emphasizing Condition, on the other hand, the gesture was to pretend to hold the object and roll it repeatedly in the air to express the movement of the egg rolling on the ground. Gesture accompanying the presentation of targets in the total of eight stimulus sets used in the present study are listed in Table 2. Note that in each condition the target was never actually held; rather, the investigator simply pretended to hold it by gesture.

After hearing the label of the target, the child was allowed to maintain contact with it for the following 15-s period in all name training conditions. Thereafter the phase was terminated by the investigator asking e.g., “Can you give me the muta?” (“Muta wo kureru?” in Japanese). After the investigator received back the target, she put it out of the child’s sight and the experiment progressed to the testing phase. In this phase, two choice objects were presented to the child. As shown in Table 1, one of the two was
of a different shape but of the same material as the target object, and the other was of a different shape but of the same material. The same material items were located on the left in one half of all test trials and on the right in the other half. Having encouraged the child to touch the two by saying “You can touch these things, too”, the experimenter asked which of the objects was e.g., muta to the child. The child’s response was determined by its pointing to, attempting to reach, or actual touching either of the objects.

For a given stimulus set, the experimental condition in which the target and choice objects were presented to each child was determined randomly. However, the child received presentations under each of the three conditions at least twice throughout the experiment.

4.3 Results

During each testing phase, 1 point was scored when the participating child chose the object that matched the target with respect to material and 0 if the response was made on the basis of shape and function. Occasionally both objects were chosen during the test phase, in which case the response was explained by a statement renderable as “Both the same”. Under such circumstances, 0.5 point was scored. Table 2 shows mean values of points the children received. Logically, consistent selection of the choice object by material throughout the experiment would give a score of 8 points.

Results were analyzed in terms of whether the child was normal or had WMS, and on the basis of condition of stimulus presentation. The main effect was seen by experimental condition, $F(2, 30) = 5.54, p < .01$. The effect of the participant group was not significant, $F(1, 30) = 1.39, p > .10$. However, there was a significant interaction between the factors, $F(2, 30) = 4.96, p < .05$. Post-hoc comparisons revealed that the scores differed significantly by experimental condition in the normal children, $H(2) = 6.16, p < .05$, but not in the children with WMS, $H(2) = 2.57, p > .10$. In the normal children, mean score was lower for Shape & Function Emphasizing Condition and greater for Material Emphasizing Condition. Score for Shape Emphasizing Condition was significantly greater than that of the children with WMS, $U(n_1 = n_2 = 6) = 3, p < .01$. In contrast, the score of the normal children for Material Emphasizing Condition significantly exceeded that of the children with WMS, $U(n_1 = n_2 = 6) = 4, p < .05$. Scores did not significantly different between the normal and WMS children for Naming Alone Condition, $U(n_1 = n_2 = 6) = 14, p > .25$.

4.4 Discussion

Kobayashi (1997) demonstrated that two-year-old normal children make use of an adult’s actions to guide their attention to various object properties, and that based on the information obtained this way, they make inferences about novel word meanings derived from nonsense syllable names without other linguistic information. The results obtained in the present experiment with normal children clearly confirm her finding. Moreover, they further extend the ‘content’ of the information utilized, and strongly indicate that what is necessary for the making of an inference is not necessarily the actual witnessing of

\begin{table}[h]
\centering
\caption{Mean scores (SDs) of normal children and children with WMS under three experimental conditions}
\begin{tabular}{lcc}
\hline
Condition & Normal children & Children with WMS \\
\hline
Material Emphasizing & 6.74 (1.82) & 4.59 (2.38) \\
Naming Alone & 4.11 (1.67) & 4.40 (3.14) \\
Shape & Function Emphasizing & 1.92 (1.55) & 4.62 (3.37) \\
\hline
\end{tabular}
\end{table}
the action with the object per se. Rather, watching pretention alone in simulation of such action with the object is sufficient for the children to solve the mapping problem in learning the correct meaning of the word labelled to the object. This pretention can be taken as a gesture serving as a vehicle expressing semantic representation of its performer as well as a word uttered simultaneously.

Conceptually, semantic representation is regarded as having two dimensional structures, as argued by Kita (1997), the analytic dimension and affecto-imaginistic dimension. While the analytic dimension is a dimension of decontextualized predication, according to his definition, the other corresponds to “iconic and affective representation of a proto-eventuality”, which is a unit produced as a consequence of temporal contoured unification of bundled information. Obviously, the meaning of the symbolic nature of language is denoted mostly in the analytic dimension of semantic representation. On the other hand, mental information borne in those gestures performed in the present study is denoted in the affecto-imaginistic dimension, in which characteristics of language other than the symbolic one (linguistic meaning) e.g., sensory, motor and affective information is recognized. Developmentally, what young children are required to do to learn a word is to bundle various modality-specific information stemming from recurring experience with the hearing of the word, and to somehow combine the information to represent a coherent and amodal notion. When this is achieved, they are considered to understand the correct meaning of the word. In other words, the integration of information from speech and that from gesture is the process in which, according to Kita (1997)’s argument, “sensory, motor, and affective information has to go through the affecto-imaginistic dimension to reach the analytic dimension, making the affecto-imaginistic dimension the only interface for” (pp. 410) word learning for normal young children, though Kita himself does not argue the developmental implication of the phenomena.

Kita points out as a future topic the so-called “interface question”, that which concerns the actual mechanism by which linguistic information becomes accessible through the modality-specific information available in the affecto-imaginistic dimension. In this regard, the results obtained in the present study with children with WMS should be intriguing because in these children information from gesture did not affect the decoding of information from speech.

5. General Discussion

5.1 Overextension and WMS

Although most normal children’s word meanings are very much adult-like from the beginning of lexical development, the fact must be admitted that some are in essence still far from adult meanings. Typically, the child has a broader meaning, as anecdotally reported by Bowerman (1978) in which her daughter used the word moon to refer to the moon, to a ball of spinach, to hangnails she was pulling off, to half a Cheerio, to curved steer horns mounted on a wall, and to a magnetic capital letter D she was about to put on the refrigerator. Another baby first used the word baby to refer to a framed picture of himself his mother had just labeled baby. He then used baby in reference to his own reflection in a mirror, and then to refer to any framed photograph. In discussion of the meaning representations in children’s mental lexicons, a great deal has been made of this sort of error, which is called overextension.

Seemingly the extent to which young children overextend the meaning of a word when the word is heard is superpressed, due to the integration of information from speech and that from gesture. However, this notion cannot be empirically tested as long as only normal children are under investigation. In this regard, the present results provide the possibility that examination of the hypothesis could be plausible by studying chil-
dren with WMS, whose ability to immediately perform correct repetition of the verbal stimuli provided by an investigator is markedly superior to that of normal children, who mispronounce the stimuli more frequently (Karmiloff-Smith et al., 1997). The WMS people are most often characterized by their overperformance in encoding the phonological form of the presented sound sequences and in their repetition, in the absence of the influence of the type of ongoing gesture when meanings of new words are categorized. Given the fact that, in contrast to the situation of individuals with Down syndrome, the chromosomal defect in WMS individuals does not significantly disrupt language faculty, the present findings indicate that word learning as an important aspect of such faculty is not performed exclusively by auditory processing mechanisms such as phonological memory alone, but that visual processing also participates.

Generally, WMS is also said to be characterized by a striking deficit in global processing of visual stimuli. For instance, when copying from memory a letter D that is built from a collection of small Y’s, WMS adolescents are reported to reveal impairment in integrating details into a large configuration (Lenhoff et al., 1997). Such feature renders difficult the prevention of overextension, which consequently is taken, at least partly, to account for the high degree of loquaciousness of WMS people.

5.2 Possible Role of Working Memory in Word Learning

The verbal and auditory features of WMS may somewhat contrast with the characteristics of some linguistically disabled children reported to date. Children with deficits in aspects of language development but with normal nonverbal intelligence are typically described as having a developmental disorder. Studies with this disorder have reported the co-occurrence of impairment of short-term memory (e.g., Graham, 1980; Kirchner & Klatsky, 1985). A more detailed analysis (Gathercole & Baddeley, 1990) revealed that they were poorer at repeating single nonwords and recalling word lists than even younger children of matched verbal abilities. As there was no evidence suggesting that the poor memory performance was due to either impaired perceptual processing or to slow articulation rate, the poor memory performance of these language-disordered children was considered to derive from an impairment of phonological storage in working memory. This was taken to play a central role in their disordered language development. The results were confirmed later by Swanson & Berninger (1995) and Swanson & Alexander (1997).

A central principle of the working memory approach is that the temporal storage of information may play an important role in a range of complex cognitive tasks such as learning, comprehension and reasoning. Baddeley & Hitch (1974) suggested that many short-term memory phenomena could be best conceptualized within working memory framework. They proposed a system that had three main components, an attentional system known as the central executive, and two subsidiary slave systems, the visuospatial sketchpad and the phonological loop. The visuo-spatial sketchpad is concerned with constructing and manipulating visual imagery, and the phonological loop is specialized in the temporary storage of phonological material. The model is described in detail by Baddeley (1986). The argument developed in the series of studies by Swanson is obviously based on this conceptualization. Since then, the function of the phonological loop has become the most actively-investigated aspect of working memory.

Although the phonological loop is no doubt a prerequisite for word learning, my findings in the present experiment indicate the importance to the categorization of meanings in novel word learning of integration (probably in the central executive) of information from speech (stored in the phonological loop) and that from gesture
(stored in the visuo-spatial sketchpad). While the phonological loop itself does not appear impaired in children with WMS, they still have serious difficulty in such categorization. In this regard, the findings of a recent neuropsychological study (Iwata, 1997) with WMS should be intriguing, which found patients with this disorder to have no difficulty in simply “tracing” out a copy of a presented visual stimuli despite having difficulty in “drawing” the stimuli. Iwata subsequently conducted a series of detailed neurological and neuropsychological examinations on the patients, but found no evidence indicating that the difficulty was due to the impairment of the capacity of the visuo-spatial sketchpad per se, or due to the impairment of the ability to immediately recall the information. Rather the difficulty is taken to occur owing to the impairment in access from the visuo-spatial sketchpad to the central executive. Otherwise, as an alternative possible (probably more central) deficit, it might be assumed that they cannot associate object properties with properties of action afforded from information that comes from the visuo-spatial sketchpad. In either case, the storage of information in one of the two subsystems appears inefficient for central processing of word categorization, which relatively enhances the importance of information that comes from the other subsystem, i.e., the phonological loop.

5.3 Possible Neurological Basis for Word Learning

The possibility of the impairment of input from the visuo-spatial sketchpad to the central executive in the working memory in children with WMS appears to be supported by the latest neurological findings in this disorder. Rae, Karmiloff-Smith, Lee, Dixon, Grant, Blamire, Thompson, Styles, & Radda (1998) succeeded in unraveling the biochemical bases of cognition in healthy people and in WMS. They identified particular biochemical abnormalities in the brain that may be correlated with the level of cognitive performance in an array of standard measures of cognition. The authors used two sets of biochemical measures, both derived from magnetic resonance spectroscopy (MRS), but each distinct in biochemical specificity and cerebral locus. The first set, localized in the left cerebellar hemisphere, was 31H MRS-derived concentrations or ratios, such as creatinine/N-acetylaspartate. The second set, localized bilaterally in the frontal cerebrum, consisted of 1P MRS-derived concentrations such as phosphomonoester/total phosphorus. Significant differences were found between WMS patients and controls in both the cerebellum and cerebrum, and significant correlations were shown between cognitive measures and cerebellar biochemistry. Such a relationship is intriguing, given the recent finding that the activity of the cerebellum is significantly associated with anticipated or imagined movement in time and space (Decety & Ingvar, 1990). Their pioneering study of such association reported that even simple mental practice (mental imaging) of certain motor activity (e.g. playing tennis) activates the cerebellum (as measured by the amount of the blood flow as an indicator), and that the extent of activation was as high as when the motor activity was actually performed. This sort of mental imaging can be taken as that which healthy children are required to perform to integrate information about gesture they see with hearing speech sound in word learning. In other words, they have to perceive affordance in the gesture as the visual stimuli for categorizing the meaning of speech. The perception can be impaired by abnormality in the cerebellum, which could be, according to the cognitive term, specified in the impairment of the working memory.

6.0 Summary and Future Direction of Research

The present study suggests the possibility that the visuo-spatial component of working memory may be directly involved in at least one important aspect of language acquisition in children.
— word learning. Children with impaired visuospatial memory skills, and consequently with difficulty in integrating such information with phonological information, should experience at least one problem in vocabulary development, namely overextension to an unusually large degree. To overcome this, they would be expected to show a high degree of verbal activity and auditory hyperacuity.

Of course, this argument must be empirically examined, but fortunately testing is in practice possible. In the field of working memory research, various measures have been developed, each of which evaluates different aspects of memory as a system. Using these, therefore, measures of phonological processing accuracy, visuospatial processing accuracy as well as executive processing accuracy can be compared between normal children and children with WMS. The above reasoning should lead to the hypothesis that compared to normal children, children with WMS have superior phonological processing accuracy but inferior visuo-spatial accuracy, whereas no difference exists in terms of executive accuracy.

Acknowledgments

This research was supported by Grant-in-Aid 09207106 from the Japan Ministry of Education, Science, Culture, and Sports. I am grateful to Sotaro Kita and Hirofumi Saito for reading the earlier version of the article and making a number of invaluable comments.

References


Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental dis-


(Received 15 Oct. 1999)

(Accepted 15 Jan. 2000)

---

**Nobuo Masataka**

Nobuo Masataka obtained his B.Sc. (1977), M. Sc.(1979) and Ph.D. (1982) in Psychology from Osaka University, all concerning the topic of communication in nonhuman primates. He is currently an Associate Professor of Primate Research Institute, Kyoto University. His main research interests are in evolution of language with special emphasis of the comparison between development of vocal behavior of human infants and that of nonhuman primates.