Japanese Preschoolers’ Early Understanding of (Un)certainty:
A Cultural Perspective on the Role of Language in
Development of Theory of Mind

Tomoko Matsui¹, Peter McCagg¹, Taeko Yamamoto², and
Yohko Murakami³
International Christian University¹, Kyorin University², and
Stanford University³

1. Introduction

In this paper we present two studies that were motivated by considering three different perspectives on cognition and communication; namely, recent research on theory of mind, research on the manifestation of cultural values in communication styles, and research from a pragmatics theory perspective that suggests that inferential communicative capabilities are a subcategory of a more general theory of mind capability.

First, there would appear to be a fairly broad consensus about the maturational schedule involved in coming to know that others have intentions, desires and beliefs which motivate their behavior (Wellman et al. 2001). It is also widely agreed that mind-reading capabilities are essential for communication to function smoothly (Happe 1993). In addition, newer, and richer evidence suggesting the importance of verbal communication in theory of mind development have now become available (Astoning & Jenkins 1999; de Villiers & Pyers 2002; Hale & Tager-Flusberg 2003; Lohmann & Tomasello 2003; Peterson 2002; Ruffman et al. 2003). Possible cultural variation in this domain has been speculated, but research on non-western communities remains

¹ We have deeply benefited from stimulating discussions with Anne Fernald, who introduced the story-telling data to us and gave us permission to use the data for our analyses. We are indebted to Miki Kakinuma, who collected Japanese story-telling data that made our analysis possible. We also thank the members of the psychology department’s Infant Lab at Stanford University, especially, Tomoko Wakabayashi. We are grateful to the children and teachers of Nishino, Nozaki and Yamanaka nursery schools in Mitaka, Tokyo. Our special thanks also go to a group of Division of Languages’ students at International Christian University, led by Yui Miura, who assisted our experiment. This study was partly financed by JSPS Grants-in-aid for scientific research # 20296792 to Tomoko Matsui and Peter McCagg.

Second, there is a body of work suggesting that cultural values are manifested in communication styles. There is, for example, a persistent and popular literature in Japan and about Japanese speakers which essentially claims that Japanese is a “less-direct” language than English. Clancy (1986), for example, argues that Japanese are hesitant to say things explicitly that may cause any discomfort or loss of face for their listeners, and that from these characteristics grow heightened sensitivities to the unstated, to the implicit layers of communication. To name but one linguistic example of this phenomena, Ohta (1991) has pointed out that it is quite common in conversation among Japanese for speakers to question the truth of the propositions expressed in their own utterances. By using the sentence-final particle *kana* (“I wonder’), for example, speakers transform would be assertions into suggestions. *Empathy*, the term typically used to refer to this type of mind-reading ability, is considered to be indispensable part of communication in Japan.

Third, recent pragmatics approaches to studying theory of mind phenomena have helped develop our understanding of how inferential interpretation capabilities are related to theory of mind development (Sperber & Wilson 1986/1995, 2002).

In both naturalistic and experimental approaches to the development of theory of mind, considerable attention has been paid to the speech of mothers and children as offering the most relevant data for analysis. On the one hand, there is a body of work that suggests that mothers’ speech styles are culturally specific (Clancy 1986; Fernald & Morikawa 1993; Shapiro & Fernald 2003; Wakabayashi & Fernald 2000). On the other, findings that mothers’ use of mental state language correlates with children’s development of theory-of-mind understanding suggest that there are at least some individual differences in developmental paths in the domain (Dunn et al. 1991; Meins and Fernyhough 1999; Ruffman et al. 2002). Yet the questions of whether and how culturally-specific aspects of communication styles might play a role in shaping a child’s theory of mind, remain largely unexplored.

Below, then, we present two studies that we have conducted to try to ascertain first whether there are important differences in how Japanese mothers and American mothers from otherwise similar socio-economic backgrounds speak to their children on a specific story-telling task, and second, whether Japanese children show any different tendencies in their abilities to recognize speaker attitudes, here, specifically, speakers’ attitudes of uncertainty or certainty with respect to novel object labels. Before introducing our studies, however, we discuss the notion of speaker (un)certainty as a subcategory of epistemic modality and briefly review what is known about young children’s understanding of (un)certainty.
2. Epistemic Reasoning and Children’s Understanding of Speaker (Un)certainty

Over the past decade and a half, investigations of children’s epistemic understanding have made significant contributions to our knowledge of developing theory of mind. In semantics, encoding of two types of epistemic reasoning are identified: (a) expressions that convey speakers’ attitudes or commitment to the truthfulness of a proposition expressed, and (b) expressions that concern the source of information described in a proposition expressed (Chafe & Nichols 1986; Fitneva 2001; Lyons 1977; Moriyama et al 2000; Palmer 1986). Here, we will use the term ‘speaker (un)certainty’ to refer to the former, and ‘evidentiality’ to refer to the latter.

In the standard theory of mind paradigm, development of children’s epistemic reasoning ability typically has been measured by testing their ability to make qualitative assessments of the source of their beliefs, in other words, their understanding of evidentiality (e.g. Gopnik & Graff 1988; O’Neill & Gopnik 1991; Montgomery 1992; Papafragou & Li 2002; Robinson et al. 1995). Results of experimental testing suggest that 4-year-olds, but not 3-year-olds, are typically capable of identifying and assessing the source of their belief appropriately.

Children’s understanding of speaker (un)certainty has also been investigated within this standard paradigm, although on a much smaller scale. Moore and his colleagues have carried out a series of experimental studies which tested whether young children can adequately distinguish the semantic/pragmatic difference between word pairs such as know and think (Moore et al. 1989, 1990). The results of their experiments also suggest there is a threshold between 3- and 4-year-olds: specifically, 3-year-olds are not capable of distinguishing between expressions of speaker certainty and expressions of speaker uncertainty.

Recently, however, there has been growing interest in investigating 3-year-olds’ and younger children’s capabilities in the domain of theory of mind understanding. Notably, it has been claimed that word-learning requires 2-year-olds to understand the intention of the speaker (e.g. Tomasello 2000). Furthermore, it has been claimed that young children also understand some epistemic modality in certain domains. Sabbagh & Baldwin (2001), for example, argue that in order for a child to learn the meaning of a new word, the child needs to be capable of judging if the speaker knows the word or not. Without this ability, a child would risk learning wrong word-referent links when the speaker is ignorant about the link. In one of two studies, Sabbagh & Baldwin investigated whether young children learn words more readily from knowledgeable as opposed to ignorant speakers. Their experimental results suggest that both 4- and 3-year-olds are capable of understanding speaker’s confidence about word-referent links, when such attitude is verbally expressed, as in ‘I know this toy’, or ‘I don’t know this toy’.
Their study is of interest here particularly for two reasons. First, unlike findings of Moore and his colleagues, Sabbagh & Baldwin’s study indicates that 3-year-olds may have some sensitivity to speaker uncertainty. Moreover, the successful performances by 3-year-olds may be partly due to the use of the word-learning paradigm. If so, the same paradigm should be effective in testing younger preschoolers’ understanding of speaker (un)certainty. However, their stimuli exclusively consisted either of explicit statements of relevant knowledge or explicit denials of it, and therefore, may not have been adequate to tap children’s understanding of degrees of speaker (un)certainty. If the stimuli contrasted between ‘think’ versus ‘know’ and ‘must’ versus ‘might’, 3-year-olds may not have responded as successfully as they did in the case of ‘know’ versus ‘don’t know.’ Thus, it seems worth exploring if in the same word-learning situation, 3-year-olds show any sensitivity to more fine-tuned speaker (un)certainty.

To summarize, studies investigating children’s understanding of speaker (un)certainty agree that 4-year-olds are capable of differentiating degrees of certainty. However, whether 3-year-olds, who are clearly capable of using ‘think,’ ‘know’ and ‘maybe’ to express speaker (un)certainty in their own speech, can properly differentiate fine degrees of certainty remains unclear. Overall, relatively little is known about younger preschoolers’ understanding of speaker (un)certainty, much less the ways that they come to grasp this concept. Robinson et al. (1995:683) speculate that exposure to oral expressions of uncertainty may facilitate children’s understanding of this concept. Given the evidence suggesting that mothers’ input influences children’s understanding of others’ minds, we also believe it is likely that mothers’ talk about speaker (un)certainty may affect children’s understanding of this concept. The present study aims to test this possibility.

3. Study 1: Analysis of naturalistic data

3.1. Participants

Japanese and US mothers were recorded as they told stories to their 3-5-year-old children during home observations in Okinawa and Kyoto (n=13, \( M=3.9 \)) and in San Francisco (n=13, \( M=4.6 \)). Stories were based on four pictures depicting emotionally ambiguous situations, such as a group of children playing together while another child is off to the side alone.

3.2. Coding categories and procedures

We classified mothers’ mental state utterances into ten categories. Most of the categories correspond to standard English mental state term categories (Bartsch & Wellman 1995; Dunn et al. 1991; Ruffman et al. 2002), but four new categories were added for two reasons. First, we were interested in looking at mothers’ expressions used to communicate speaker attitudes, as well as expressions typically used to refer to genuine mental states. Second, we needed
to accommodate expressions in Japanese that do not fit any of the standard English mental state categories. The four new categories are: “Sharing uncertainty”, “Asking for confirmation”, “Intentionality” and “Attributed speech”. Figure 1, below, introduces and illustrates the coding scheme.

We shall restrict our comments on coding categories to those that are most relevant for the present paper. All uses of “think”, ‘know’ and ‘believe’ are combined in a single category, labeled “Knowledge.” Even relatively formulaic uses of these predicates such as ‘you know what?’ (which would be coded as “Directing the interaction” by Shatz et al. 1983) or ‘I think he is pretty happy’ (which would be classified as “Modulation of Assertion” by Shatz et al. (ibid.)) were included here. Although they do not refer to genuine mental states, these formulaic, or conversational, uses of ‘think’ and ‘know’ are of interest to us as they are used to communicate various speaker attitudes.

<table>
<thead>
<tr>
<th>Category Name</th>
<th>Lexical Exemplars</th>
<th>Sample Japanese Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Knowledge</td>
<td>think, know, believe</td>
<td>Ii to omou?</td>
</tr>
<tr>
<td>2. Desire</td>
<td>want, like, hope,</td>
<td>Doko ikitai?</td>
</tr>
<tr>
<td></td>
<td>hoshii, suki, negau</td>
<td>Where do you want to go?</td>
</tr>
<tr>
<td>3. Emotion</td>
<td>happy, sad, angry</td>
<td>Sabissso da ne.</td>
</tr>
<tr>
<td></td>
<td>unreshi, kanashii, okotteiru</td>
<td>[He] looks sad.</td>
</tr>
<tr>
<td>4. Modulation</td>
<td>might, perhaps, looks like</td>
<td>Tsuniki no oshiro mitai da ne.</td>
</tr>
<tr>
<td></td>
<td>kamoshirenai, tabun, mitai</td>
<td>It looks like a block castle.</td>
</tr>
<tr>
<td>5. Sharing Uncertainty</td>
<td>maybe, wonder</td>
<td>Kenka shita no kana?</td>
</tr>
<tr>
<td>6. Seeking</td>
<td>kana, daroka</td>
<td>I wonder if [they] had a fight.</td>
</tr>
<tr>
<td></td>
<td>right, isn’t it</td>
<td>In da yo ne.</td>
</tr>
<tr>
<td>7. Intentionality</td>
<td>mean to, try to</td>
<td>It’s okay, isn’t it?</td>
</tr>
<tr>
<td></td>
<td>tsamori, (shiy)ou-tosuru</td>
<td>I meant to go, but…</td>
</tr>
<tr>
<td>8. Attributed Speech</td>
<td>say, ask</td>
<td>Gomennasai te.</td>
</tr>
<tr>
<td></td>
<td>to/tte iu, to/tte tanomu</td>
<td>[He] said “sorry”.</td>
</tr>
<tr>
<td>9. Causal Talk</td>
<td>why, because, since</td>
<td>Waruikoto shita kara.</td>
</tr>
<tr>
<td></td>
<td>doshite, kara, node</td>
<td>Because I did something bad.</td>
</tr>
<tr>
<td>10. Orienting Utterances</td>
<td>oh, look, see, wow</td>
<td>Nee, dare to iku?</td>
</tr>
<tr>
<td></td>
<td>are, hora, nee, ara,</td>
<td>Say, who are you going with?</td>
</tr>
</tbody>
</table>

Figure 1: Coding scheme for mother’s mental state utterances

The category “Sharing uncertainty,” was established—following Masuoka (2002)—to allow expressions of speaker Uncertainty to be grouped separately from expressions of speaker certainty, which we classified as “Modulation.” In other words, here the category “Modulation” includes only expressions used to convey speaker certainty. There is general agreement that the Japanese sentence-final particle kana indicates a speaker’s uncertainty about the articulated proposition. In our analysis, we have included ‘maybe’ and ‘wonder’ in English in the same category, as both expressions typically convey a speaker’s
uncertainty. The English expression ‘looks like’ was also included in the “Modulation” category. We felt that it was used to express speaker certainty concerning visual perception, as in ‘it looks like they’ve been having a really busy day’. One of the motivations for this was that the Japanese sentence-final adverb mitai conveys meaning that is similar to the English phrase ‘looks like’.

Initial coding was carried out independently by two of the authors. Where there were discrepancies, discussion was held to reach agreement.

3.3. Results and Discussion

Table 1 shows the number of times lexical items in each category was used, by Japanese and American mothers (JM and AM) as well as Japanese and American children (JC and AC). The total number of mental state utterances by each group of mothers was actually very similar (1150 vs 1205).

There are some notable differences between Japanese and American mothers’ mental state utterances. For example, American mothers used the verb ‘think’ frequently, whereas Japanese mothers rarely used the Japanese equivalent, omou. On the other hand, Japanese mothers used the verb iu (‘say’) to mention attributed speech rather often. Japanese mothers also used ne (a sentence-final confirmation marker to elicit sympathy or agreement) with high frequency (Category 6).

Table 1: Frequencies of mental state term use by Japanese mothers, American mothers, Japanese children and American children

<table>
<thead>
<tr>
<th>Mental state terms</th>
<th>JM</th>
<th>AM</th>
<th>JC</th>
<th>AC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Knowledge</td>
<td>25</td>
<td>293</td>
<td>5</td>
<td>65</td>
</tr>
<tr>
<td>2. Desire</td>
<td>22</td>
<td>74</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>3. Emotion</td>
<td>121</td>
<td>146</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>4. Modulation</td>
<td>27</td>
<td>173</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>5. Sharing uncertainty</td>
<td>255</td>
<td>109</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>6. Asking for confirmation</td>
<td>366</td>
<td>34</td>
<td>21</td>
<td>1</td>
</tr>
<tr>
<td>7. Intentionality</td>
<td>60</td>
<td>98</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>8. Attributed speech</td>
<td>92</td>
<td>52</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>9. Causal talk</td>
<td>73</td>
<td>74</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td>10. Orienting utterances</td>
<td>109</td>
<td>152</td>
<td>28</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>1150</td>
<td>1205</td>
<td>166</td>
<td>214</td>
</tr>
</tbody>
</table>

Of all the quantitative differences observed, however, those in Category 4 “Modulation” and 5 “Sharing uncertainty,” appear to be most worthy of further analysis if evidence of the effect of culturally influenced style differences on the development of theory of mind is sought. Japanese mothers used expressions of speaker uncertainty significantly more often (22%) than American mothers (9%), p=0.02. At the same time, American mothers used expressions of speaker certainty, which are categorized under “Modulation” here, more often than Japanese mothers. A difference of this magnitude suggests that further study is
warranted, and led us to conduct the elicitation study introduced below as Study 2.

A closer look at the data revealed that for the American mothers, ‘maybe’ was the most common expression of (un)certainty. The expression that is most commonly used by Japanese mothers in this data set is *kana*. Japanese mothers’ frequent use of this particle compared with any other expressions, is striking (76.8% of the total number of utterances in Category 4 & 5). Examples of typical use of *kana* that appeared in our data of Japanese mothers’ utterances are presented below.

(1) *Undookai de hashite n no kana.*
   Sport festival at are-running NOM I-wonder
   ‘I wonder if maybe they are running in an athletics festival.’

(2) *Nande okotte ru no kana*
   why is-angry NOM I-wonder
   ‘I wonder why he’s mad’

The finding that Japanese mothers use the sentence-final particle *kana* to express speaker uncertainty nearly 80 percent of the time when they express any attitude of certainty towards proposition expressed in their utterances led us to select this lexical item for use in testing experimentally Japanese children’s understanding of speaker (un)certainty. One data-based study (Shirai et al. 1999) reports that children acquire *kana* soon after their second birthday, so we assumed that the particle would be in normal 3-year-old Japanese children’s productive vocabulary.

4. **Study 2: Elicited Data of Japanese Children**

Given that American 3-year-olds were wisely reluctant to assign labels when speakers signaled uncertainty about the word-referent link in Sabbagh and Baldwin’s study, we decided to adapt their experimental procedure to test Japanese children’s understanding of speaker uncertainty. We revised slightly the procedure, and more crucially, we altered the stimuli so that they included no explicit statements about speaker knowledge, and making the only difference between the two conditions overt expressions to mark speaker’s attitude of certainty/uncertainty. We were interested in young children’s understanding of propositional attitude of uncertainty, which is distinguished from understanding of propositional content, and in our opinion, an explicit statement of knowing and not knowing something belongs the latter. Our expectation here was that the change of stimuli might influence children’s performance: for example, children may not sense an attitude of uncertainty when they do not hear an explicit statement that the speaker does not know the right link between the label and the toy.
4.1. Participants

Fifty-six normally developing Japanese-speaking children from two age groups participated. The 3-year-olds ranged from 3 years, 0 months (3.0) to 3,11 (M=3.6); and the 4-year-olds ranged from 4.0 to 4,11 (M=4.5). There were equal numbers of boys and girls. Participants were recruited from three nursery schools in a west Tokyo city with a primarily middle-class population.

4.2. Design

In a between subjects design, children were randomly assigned to one of two experimental conditions: (1) certain speaker condition and (2) uncertain speaker condition. The only difference between the two conditions was the sentence-final particles used by the experimenter to introduce each set of three toys and the name of the target toy to the participants. In the certain speaker condition, the Japanese particle dayo (equivalent to English adverbial sure) was used; in the uncertain speaker condition kana (equivalent to English maybe) was used. The particle dayo (or the combination of two particles da and yo, to be more precise) was chosen as the acquisition period of the particles more or less coincides with that of kana, and it has the same number of syllables as kana.

Children participated in two trials. The central materials were two sets of three colorful novel plastic toy objects which did not fit neatly into known named toy categories. Two hand puppets, i.e. Mickey and Donald, always introduced the same set of toys, and the target toy in Mickey’s set was always named mappi, and the target toy in Donald’s set was named toma.

4.3. Procedure

In both conditions, participants were introduced to a puppet (Mickey or Donald) who quickly excused himself to go play outside with Minnie (Mouse), but before leaving gave the child participant permission to play with his toys while he was away. Once the puppet had gone, the experimenter suggested that they play with the toys. The first critical experimental manipulation started when the experimenter showed how to play with the toys. In the certain speaker condition, the experimenter communicated her confidence about how to play with each toy by using the sentence-final particle dayo. In the uncertain speaker condition, the experimenter communicated her confidence about how to play with each toy by using the sentence-final particle dayo. In the uncertain speaker condition, the experimenter’s uncertainty about how to use each toy was expressed by the particle kana. This activity ended when a phone rang, and the child started listening to a recorded message from Mickey (or Donald) asking for someone to send him their mappi or toma. To check each child’s understanding, the experimenter asked the child to repeat the message. The second critical experimental manipulation began when the experimenter named the target object. In the certain speaker condition, the experimenter named the toy using the particle dayo. In the uncertain speaker condition, the toy was named along with the particle kana. The child then was asked to take the target
toy to the mailbox so that Mickey/Donald could play with it. Next, the experimenter suggested that they look at Mickey's/Donald's photo album and see how many of the pictures they could name. Each album contained photographs of familiar objects, unfamiliar and non-presented objects, and two photographs of each of the novel objects presented earlier. The photo albums offered each child two opportunities to name the novel objects.

Immediately after the second trial, children were given a comprehension test by asking them to identify the named novel objects (mappi and toma) from among all six toys used across both trials displayed randomly on the table.

4.4. Hypotheses and Predictions

Past studies of children’s understanding of uncertainty generally suggest that 4-year-olds are capable of differentiating degree of speaker certainty. Although Sabbagh & Baldwin’s study indicates that American 3-year-olds are capable of understanding speaker’s confidence about word-referent links when they are explicitly told whether the speaker does or does not have the knowledge about the target object, whether the same children understand uncertainty communicated only by markers of propositional attitudes such as particles or modal adverbs (e.g., ’maybe’) is not yet known. In this respect, studies carried out by Moore and his colleagues are more indicative: for example, the result of one of their experiments strongly indicates that American 3-year-olds do not aptly distinguish between ‘might’ and ‘must’. Thus, we hypothesized that if Japanese 4-year-olds, but not 3-year-olds, capably distinguish kana, an uncertainty marker, from dayo, a certainty marker, it would suggest that Japanese children’s understanding of speaker uncertainty develops at a similar pace as does that of American children, both of which would coincide with the so-called “watershed” at the age of four (Astington 1991). However, if Japanese 3-year-olds are able to distinguish between the two particles, that can be interpreted as an indication of earlier understanding of speaker uncertainty. Such early understanding of an aspect of theory of mind would be of significance particularly in the context that Japanese children appear to pass standard false-belief test at later point than English-speaking children (Wellman et al. 2001). Furthermore, such earlier understanding of the concept can be seen to be correlated with frequency of input. Ideally, a direct comparison should be made using two experiments following the same protocol, one with Japanese 3-year-olds and the other with American 3-year-olds, and that is our longer term aim. However, our current prediction about Japanese children is based on the very general assumption that four is also the ‘universal’ threshold age for adequate understanding of speaker (un)certainty.

4.5. Results and Discussion

The main result of the experiment was that Japanese 3- and 4-year-olds learned words better in the certain speaker condition than in the uncertain
speaker condition \(p<0.01\) both for the production performance and the comprehension task. Thus, the participant age difference wasn’t a factor for Japanese 3- and 4-year-olds. Figure 4 shows the mean production scores (maximum score being 4) of 3- and 4-year-olds in the two conditions. A 2 (age) X 2 (condition) ANOVA with elicited production performance as the dependent variable revealed a significant main effect for the condition; children learned words better in the certain speaker condition, \(p<0.01\). No main effect was found for age, or Age X Condition interaction. Similarly, Figure 5 shows mean comprehension scores (maximum 2). Again, a 2 (age) X 2 (condition) ANOVA with the comprehension score as the dependent measure revealed a significant main effect for condition, \(p<0.01\). Children showed better word learning in the certain speaker condition. No significant effect was found for age, nor Age X Condition interaction.

![Figure 4: Mean performance on the elicited production test](image1)

![Figure 5: Mean performance on the comprehension test](image2)

The results were in accordance with our predictions that Japanese 3-year-olds are capable of distinguishing \textit{kana} from \textit{dayo}, given mothers’ frequent use of sentence-final particles. At this stage, we cannot say that the only possible explanation for the result is input frequency. However, we can at least say that in our data there is correlation between input frequency of expression of uncertainty and children’s early understanding of the concept.

5. Final remarks

The two studies presented in this paper indicate that mothers’ speech style, considered here as a manifestation of cultural values of a community, is one of the possible factors that contribute to shaping children’s inferential-communicative ability. This, in turn, suggests that an individual’s resulting socio-communicative ability may also be specific enough to fit certain cultural
demands in the particular society to which she belongs. Understanding of propositional attitude of (un)certainty is part of one’s ability to understand speaker intention, and our elicited study revealed that Japanese 3-year-olds have ability to understand such an attitude. Investigation of children’s intention-reading ability in communication is still relatively new in theory of mind research, but the present study suggests that this aspect of theory of mind ability may potentially be more susceptible to cultural variation, as ability to read communicative intention is inevitably mastered only through communication itself.

References


