Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Cognitive Development 21 (2006) 158–173
COGNITIVE DEVELOPMENT

On the role of language in children's early understanding of others as epistemic beings

Tomoko Matsui [a], Taeko Yamamoto [b], Peter McCagg [c,*]

[a] Kyoto University, Kyoto, Japan
[b] Kyorin University, Tokyo, Japan
[c] Division of Languages, International Christian University, 3-10-2 Osawa, Mitaka-shi, Tokyo 181-8585, Japan

Abstract

In the study reported here, Japanese-speaking children aged 3–6 were confronted with making choices based on conflicting input from speakers who varied in the degree of certainty and the quality of evidence they possessed for their opinions. Certainty and evidentiality are encoded in Japanese both in high-frequency, closed-class, sentence-final particles and also in low-frequency, mental state verbs. Our results suggest that children are able to make use of information encoded in the sentence-final particles earlier than information encoded in verbs, and that understanding of speaker certainty precedes understanding of quality of evidence. Furthermore, although the results generally support the position that children's overall understanding of epistemic vocabulary correlates with their understanding of false-belief, understanding of the sentence-final particles tested did not correlate with false-belief understanding. We argue that understanding of speakers' epistemic states as communicated by sentence-final particles prior to the fully-representational understanding of (false) belief should be taken as an indication of children's inchoate understanding of other's mental states.

Keywords: Theory of mind; Language acquisition; Epistemic modality

1. Introduction

For the most part, we come to know about the worlds we are born into—both physical and social worlds—either through direct observation of states, events and behaviors or by being told about things by others. Neither of these sources of information, of course, is infallible. Our eyes and other senses may be misled by optical or other "illusions" or they may simply not be adequately tuned in to the relevant aspects of a situation. We must also quickly learn—or else

* Corresponding author. Tel.: +81 422 333209; fax: +81 422 333550.
E-mail address: mccagg@icu.ac.jp (P. McCagg).

0885-2014/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.cogdev.2005.10.001

live the life of a dupe—that not all we hear from others can be equally relied upon. Learning to distinguish sources of information that can be trusted from those which cannot; understanding that information can be intentionally or unintentionally miscommunicated; and generally being able to evaluate the evidential bases, logical consistency and plausibility of ideas presented by other speakers to enable one to distinguish truth from falsity are crucial skills that allow human beings both ontogenetically as individuals and phylogenetically as a species to survive.

One of the more persistent questions in recent theory of mind research has been just when do children become able to competently assess the evidential quality of the information they are exposed to. Many deeply interesting experiments using both linguistic and non-linguistic tasks designed to help answer this question are discussed in the theory of mind literature. Results of some studies using non-linguistic tasks, for example, have shown that 3- and 4-year-old children generally understand that a person who has seen inside a container will have knowledge about the contents, whereas a person who has merely touched but not seen into the container, will not have such knowledge (Pillow, 1989; Pratt & Bryant, 1990). Such results suggest that, in the absence of any verbal communication, 3-year-olds appear to understand that seeing provides different information than touching does. Other studies, however, also suggest that children's ability to verbally report sources of information may develop more slowly (Gopnik & Graf, 1988; O'Neill & Gopnik, 1991).

Recent studies on children's suggestibility and ability to assess speaker reliability provide new evidence about children's early understanding of reliable speakers. They have shown that 3-year-olds who cannot provide a verbal report of sources of their belief, nonetheless are capable of deciding who to believe and who not to believe at the time of input (Koenig, Clements, & Harris, 2004; Robinson, Mitchell, & Nye, 1995; Robinson & Whitcombe, 2003; Sabbagh & Baldwin, 2001; Whitcombe & Robinson, 2000). These studies indicate that young children are capable of utilizing non-linguistic (e.g., plausibility of content of previous utterances, accessibility to evidence, etc.) clues to distinguish reliable speakers from unreliable ones. This, in turn, suggests that some, possibly implicit, understanding of speaker's epistemic states develops before explicit understanding of false-belief (Dienes & Perner, 1999). In the current study, we attempt to investigate the issue, focusing on how young children who typically fail the standard false-belief tasks can make use of *linguistic* clues to assess a speaker's epistemic stance (Chafe & Nichols, 1986; Thompson & Mulac, 1991).

2. Types of epistemic vocabulary: categorization based on linguistic form and modality

Study of children's acquisition of mental state verbs such as *think* and *know* has been a precursor to and has remained an important means for examining children's theory of mind development over the past two decades (Abbeduto & Rosenberg, 1985; Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982; de Villiers & de Villiers, 2000; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991; Furrow, Moore, Davidge, & Chiasson, 1992; Johnson & Maratsos, 1977; Johnson & Wellman, 1980; Olson, 1988; Shatz, Wellman, & Silber, 1983; Ruffman, Slade, & Crowe, 2002; Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003). Because mental state verbs refer directly to cognitive states, they have been considered one of the most reliable overt indicators of developing theory of mind—of thinking or knowing, for example. Functional analyses of mental state verbs, however, have revealed that verbs such as *think* and *know* can also be used to express degree of speaker certainty or uncertainty about the truthfulness of an utterance's propositional content, and it is this aspect of mental state verbs that is more relevant to the present research (Diessel & Tomasello, 2001; Shatz et al., 1983). Although some cross-linguistic studies of children's use of expressions of certainty broadly understood exist, data focused specifically on both production and comprehension of expressions of speaker (un)certainty currently only exist for English-speaking children. Existing studies investigating children's understanding of speaker (un)certainty concur; 4-year-olds are capable of differentiating degree of certainty (e.g., Moore, Bryant, & Furrow, 1989; Moore, Pure, & Furrow, 1990). 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 finer degrees of certainty, remains an open question.

Far fewer studies of acquisition of epistemic modality in languages other than English are available (e.g., Aksu-Koc, 1988; Choi, 1995; Fitneva, 2001; Lee & Law, 2001; Papafragou & Li, 2002; Shirai, Shirai, & Furuta, 1999). Nonetheless, those that have been published provide intriguing data particularly on acquisition of evidential vocabulary, an area largely unexplored in the extant English studies. To date, experimental testing of young children's understanding of evidential particles has only employed questions that require metalinguistic analysis, or conscious reflection, of what is encoded by the evidential particles, and these studies indicate that metalinguistic awareness of linguistic indications of indirect evidence and hearsay develops later (not until about 5 years of age) than awareness and understanding of direct evidence markers. Clearly, metalinguistic understanding of word meaning needs to be distinguished from more spontaneous, or unconscious, understanding of what is communicated by an utterance (e.g., Karmiloff-Smith, 1992; Karmiloff-Smith, Grant, Sims, Jones, & Cuckle, 1996). Currently, very little is known about when and how young children can make use of information conveyed by evidential particles, in order to evaluate a speaker's knowledge on the matter at hand. Here, we attempt to investigate this relatively unexplored aspect of children's understanding of evidential particles.

It is important to note here that children's understanding of encoded speaker certainty has been exclusively studied by using verbs as target stimuli, while children's understanding of encoded evidential quality has been tested by use of particles as target stimuli. Whereas existing studies on children's understanding of English, for example, deal with mental state verbs of speaker certainty, studies on children's understanding of particles in Turkish and Korean, for example, deal exclusively with evidentiality. Japanese, which allows encoding of speaker certainty in both closed-class sentence-final particles and open-class verbs, and which also permits encoding of evidentiality in the same two types of linguistic form, permits researchers to look at the interaction among linguistic form and epistemic modality more directly.

In Japanese, speakers have a variety of linguistic forms with which they can convey speaker certainty (or uncertainty) as well as the directness of evidence available to the speaker (i.e., whether the speaker has direct knowledge of a state of affairs, or whether the speaker has only less reliable evidence, such as hearsay, for the information asserted in the propositions expressed). These differences can be communicated via sentence final particles such as *yo* and *kana*, which convey speaker certainty and uncertainty, respectively; via adverbials such as *kamoshirenai* (*perhaps*) and *tabun* (*maybe*); and via mental state verbs such as *shitteru* (*know*) and *omou* (*think*). As such, Japanese offers a highly useful set of expressions that allow investigation of children's developing abilities to base decisions on the degree of speaker certainty and the quality (or directness) of evidence upon which the oral testimony of others is based. Furthermore, the variety of expression available in Japanese also permits looking at how the ability to make correct choices develops depending on whether the relevant information is encoded in high frequency sentence-final particles or less frequent mental state verbs. Schematically, these distinctions may be represented as in Table 1.

Table 1
A taxonomy of linguistic form and epistemic modality combinations with Japanese exemplars

Linguistic form	Epistemic modality	
	Certainty	Evidentiality
Particles	(Contrast 1) *yo* vs. *kana*	(Contrast 3) *yo* vs. *tte*
Verbs	(Contrast 2) *shitteru* vs. *omou* (*know* vs. *think*)	(Contrast 4) *miru* vs. *kiku* (*see* vs. *hear that*)

3. Encoding speaker certainty and evidentiality in Japanese

For the research reported here, we selected four contrasting pairs of Japanese linguistic forms to examine—two pairs of sentence-final particles, and two pairs of mental state verbs. To look at speaker certainty, we tested understanding of the particle pair of certainty *yo* and uncertainty *kana*, and the verb pair of certainty *know* (*shitteru*) and uncertainty *think* (*omou*). To look at evidentiality, we tested understanding of the particle of direct speaker knowledge *yo* and the particle of hearsay *tte*, and the verbs of direct knowledge *see* (*miru*) and of hearsay *hear that* (*kiku*). In each epistemic category, the contrast is encoded once in sentence-final particles, and once in the verbs. The first item in each contrastive pair is the relatively higher certainty (or more direct evidence) marker, and the second item is the relatively lower certainty (or more indirect evidence) marker.

3.1. Speaker certainty pairs

Sentence-final particles, *yo* and *kana* (Contrast 1) communicate speakers' relative certainty about the truth of the propositions expressed in their utterances. *Yo*, when affixed to assertions, expresses strong speaker conviction about the information communicated. In the Contrast 1 stimulus utterances used in our studies, the particle *yo* follows the unmarked assertive form *da*, which communicates that the speaker believes the proposition in the expressed utterance is true. (*Da*)*yo* strongly commits the speaker to the truth of whatever is asserted. *Kana*, in contrast, encodes information that the speaker cannot make a judgment as to the truth or falseness of the statement (Masuoka, 2001). The common mental state verbs, *shitteru* and *omou* (*know* and *think*) (Contrast 2), also communicate information about the speaker's relative certainty about expressed propositions. *Know* (*shitteru*) communicates that the speaker has evidence for the information to be true, and it expresses that the speaker presupposes the truth of the proposition. On the other hand, *think* (*omou*) does not presuppose the truth of the proposition; it merely communicates that the speaker is presenting her own thought on her own judgment (Moriyama, 1992).

3.2. Evidentiality pairs

Contrast 3 forms, (*da*)*yo* and *tte*, convey information about evidentiality through use of sentence final particles. *Yo* indicates that the proposition expressed is believed by the speaker to be true, and as such suggests that the speaker has compelling evidence for his or her beliefs. *Yo* marks what is old information for the speaker, but new to the listener. *Yo*'s counterpart in this pair, *tte*, is a hearsay particle, and it indicates that the proposition expressed is not based on the speaker's direct experience. The level of the speaker's certainty about the statement must be inferred indirectly. In theory, information based on indirect (hearsay or secondhand) evidence is communicated with relatively less certainty than information that is based on direct evidence. As with speaker certainty, evidentiality in Japanese can be expressed both by particles and by evidential verbs such as *miru* (see with one's own eyes) and *kiku* (hear from someone else, i.e., *hear that*). Seeing something, prototypically, is reason for believing it. Claiming to have seen something is claiming to have direct evidence for the truth of what is being claimed. Hearing from someone else that they have directly experienced something is reason for believing only that the person claims to have knowledge, not necessarily reason for believing that the described state of affairs actually occurred. Contrast 4, then, is conveyed in the experiments by the verbs *miru* and *kiku* (*see* and *hear that*) which differ in the test stimuli in terms of what they suggest about how direct the evidence is for the speaker's claims.

In the current study, we address directly the four-way distinction represented in Table 1, hypothesizing that the difference in acquisition timing among various epistemic vocabulary items may depend on factors of linguistic form and function, and cognitive complexity (i.e., steps of reasoning called for or employed in comprehension), as well as frequency of input and perceptual salience.

Our first hypothesis is that the particles and verbs may invoke different cognitive functions. The intuition that particles and morphemes (i.e., closed-class items) have a somewhat distinct function from that of verbs and nouns (i.e., open-class items) in verbal communication is widely shared by researchers who are interested in linguistic meaning (Blakemore, 1987; Karmiloff-Smith, 1992; Pinker, 1999; Talmy, 2001; Wilson & Sperber, 1993). Definite and indefinite articles, regular morphology, and auxiliary verbs in English are among prototypical examples of such closed-class items. Japanese case-particles and sentence-final particles are also considered to be closed-class items (Matsui, 2000). Typically, closed-class items are seen to encode non-representational (i.e., procedural) rule-like information, which manipulates representational (i.e., conceptual) information, and as such, are not accessible to consciousness. The well-known fact that most native-speakers can provide no immediate answer, and even after some thought, can provide only imprecise or vague responses, when being asked what any closed-class item means, such as the definite article *the* in English and the sentence-final particle *yo* in Japanese, suggests that the information is not readily accessible to conscious awareness. More recently, it has been suggested that learning of such rule-like information involves implicit, or procedural learning through gradual accumulation of experience, rather than explicit, or declarative learning (Ullman, 2004). At present, the psychological distinction between explicit versus implicit or declarative versus procedural knowledge/learning is much more controversial than linguistic distinction between open- versus closed-class items, and therefore, it would be unwise to connect the two distinctions without any-first-hand evidence to confirm it. Thus, our suggestion here is a modest one, mainly based on widely shared linguistic intuitions, leaving psychological explanation of the distinction open.

With regard to cognitive complexity, it is our position that understanding speaker certainty, whether it is encoded in sentence-final particles or in verbs, requires less cognitive processing than does comprehension of evidential certainty. Essentially, we suggest that there are extra steps required to ascertain how likely a speaker is to be committed to the content of the expressed proposition when the quality of evidence that the speaker has is embedded in an utterance. With the former (understanding speaker certainty), one needs to understand the meaning encoded in the linguistic forms used to convey that attitude. With the latter (understanding evidential certainty), one needs to understand not only the quality of evidence (direct versus indirect) encoded in the linguistic indicators (sentence-final particles or predicates), but also how the quality of evidence is likely to affect the speaker's commitment to the truthfulness of the content of the proposition. Moreover, a hearer cannot simply base a judgment in such cases on a determination of how committed the speaker may be to the truth of the proposition, but needs to calculate (at least implicitly) how likely the proposition is to be true according to heuristics about what it means for someone to hear that something is the case, to see, and therefore know that something is the case, and then weigh both in conjunction with what it means when someone says something. The interactions among these knowledge and belief sets are complicated, and at least partially explain differences that we observe in the behavior of child comprehenders.

4. Study: On early understanding of speaker certainty and evidentiality

The research presented here was conducted to further explicate the nature of developing theories of mind as it may be observed in the behavior of Japanese children between the ages of 3 and 7 with respect to their abilities to respond sensibly to linguistic indications of speaker certainty and uncertainty and linguistic markers of evidentiality. Our study proceeded in three stages. First, we examined the available naturalistic data of mother–child talk in the JCHAT corpus to get a sense of frequency of occurrence of the specific linguistic forms selected for use in our experimental studies. The JCHAT data are presented in Section 4.1. Next, we tested the four sets of contrastive pairs selected for use in the main experiment with college-age adults to verify our expectations about adult reliance on more certain and more direct evidence in making choices. This preparatory work is described in Section 4.2. Then, we conducted the core study in which we tested 3-, 4-, 5-, and 6-year-old children's ability to make correct decisions based on information presented in the selected sets of contrastive pairs. We also collected false-belief data on all children tested to allow comparisons among performances on the comprehension of speaker certainty and evidentiality recognition tasks and performance on these widely researched indicators of theory of mind.

Our hypotheses are as follows:

1. The ability to make correct choices based on information about speaker certainty and evidentiality when that information is encoded in sentence-final particles will develop earlier than when the relevant information is encoded in mental state verbs.
2. The ability to make correct choices based on (a) being able to distinguish degrees of speaker certainty will develop earlier than (b) ability to do so based on evaluation of the quality of evidence the speaker has for making his or her assertions.
3. Performances on the hidden object tasks—tasks that require at least some comprehension of other's mental states—will correlate positively with performances on standard false-belief tests.

4.1. Naturalistic data

There is little doubt that generally speaking Japanese sentence-final particles occur with far greater frequency in adult speech than do any individual verb forms. To see whether this widely observed frequency pattern holds with respect to the specific forms selected for investigation here as they are found in the speech of mothers to their young children, we consulted the JCHAT corpus, a Japanese children's speech corpus (Miyata, 2000). The JCHAT corpus consists of data from three children. Unfortunately, in two of the cases, full-data on the mothers' utterances was not recorded. Here, then we focus primarily on the JCHAT data for only one child, Tai. Tai's utterances were collected between ages 1;5 and 3;2 and were gathered during once-a-week recording sessions that lasted for about 40 min each. A total of 75 sessions make up the Tai corpus. Generalizations on the basis of a single case study are unwise, but the patterns of use discernable in the Tai corpus

Table 2
Frequency counts for certainty and evidentiality markers in the Tai corpus

		Child	Mother
Certainty			
Verb	*shitteru* (*know*)	34	70
	omou (*think*)	12	51
Particle	*yo*	3317	3955
	kana	145	970
Evidentiality			
Verb	*miru* (*see*)	109	410
	kiku (*hear*)	6	34
Particle	*tte*	270	1603

closely mirror the generally observed frequency patterns in adult Japanese speech (see Table 2). It is also the case that the data available for the other two children reflect the same frequency patterns observed in the Tai data.

The Tai data was coded by two independent groups of tabulators (university students in an undergraduate pragmatics class). Cronbach alpha reliability tests were conducted, and reliability coefficients were higher than .9 for all forms in both the mother's and in Tai's utterances except for the hearsay particle *tte*. A third group of raters tabulated the occurrences of *tte* forms in both the mother's and in Tai's utterances, and the Cronbach alpha reliability test was run on the two closest groups, revealing a correlation of over .9 for this form as well. Table 2 presents the frequencies with which the coded forms were observed in this corpus.

All three of the sentence-final particles occur with higher frequency in both the speech of Tai's mother and Tai himself than do any of the verbs under investigation here. With the exception of the relatively frequent use of *miru* (*see*) in Tai's speech, the difference between use of the sentence final particles and verbs considered is dramatic. Tai's use of *miru* (*see*), as we believe is the case with many children, is commonly found in his requests of his mother to look at something, not as an evidentiality marker as is the case in the stimuli for the studies discussed below. In fact, functional analyses of the verb utterances in the data would tease apart uses of the verbs that are not the same as those used in the experimental stimuli, and likely reduce the frequency counts of these items, and thereby further expand the input gap between the verbs tabulated and the sentence-final particles. Even without conducting such analyses, though, there are clear differences in the input frequency (mother's utterances) and the production frequencies (Tai's utterances).

A Pearson correlation test was run on the input and production frequencies; the correlation coefficient ($r = .98$) was significant at the $p < .01$ level.

4.2. Adult performance on the experimental tasks

To establish a base-line adult response pattern to the specific stimuli that would be used in the core study, we tested understanding of the four sets of contrastive pairs with 47 university students who attended an undergraduate pragmatics class and who participated in the study as a requirement for getting credit in the class. The same animation as was used in the actual study (see Section 4.3.3) was used to gather this data. Instead of individual testing, however, the animated sequences were projected on a large screen so that all participants could watch the animation at the same time. The experimenter operated the computer giving instructions as she did in the actual experiment. The participants wrote down their answers on answer sheets distributed in advance.

In four trials with the speaker certainty verb pair (*know-think*), all 47 participants answered in every case that the toys were hidden in the boxes that had been identified by the speaker claiming to know the location of the hidden objects and not the boxes identified by the speaker claiming only to think the toys were hidden in them. With the speaker certainty particle pair (*yo-kana*), the evidentiality particle pair (*yo-tte*), and evidentiality verb pair contrasts (*see-hear that*), one of the participants chose the box presented by a speaker using the latter form in each pair in one out of the four trials. All other participants, 98%, chose to believe in every case the speakers using *yo* not *kana*, *yo* not *tte*, and *see* not *hear that*. With these results confirming our own intuitions that adult Japanese speakers would choose to trust information presented by *yo* (rather than *kana*), *know* (rather than *think*), *yo* (rather than *tte*), and *see* (rather than *hear that*) we now turn to the main study in this investigation.

4.3. Core study

4.3.1. Participants

A total of 97 normally developing Japanese-speaking children—25 three-year-olds and 24 each in the age groups four, five and six—participated in this study. Mean ages for each group were 3;6, 4;6, 5;5, and 6;5, respectively. There were approximately the same numbers of boys and girls in each group. All participants were recruited from nursery schools and kindergartens in a west Tokyo city with a primarily middle-class population. Results for an additional six children were determined to be unusable either due to inability to conduct both the false-belief tests and the experimental trails on the same day, or due to inability of the children to complete all tasks.

4.3.2. Stimuli

We generated a set of eight utterance types using the seven forms introduced in Table 1 above. *Yo* serves both as a speaker certainty marker (in contrast with *kana*) and as a marker of better evidence (in contrast with *tte*). The eight utterance types were organized into four contrasting pairs of utterances as illustrated in Table 3.

All stimuli were presented on a laptop computer screen by animated characters created for the experiment. Thus, all participants were presented with the same stimuli in terms of the sound quality, volume, speed, and intonation contours, as well as visual information communicated by facial expressions and other non-verbal gestures.

Table 3
Illustrative stimulus utterance examples

Contrasts	Degree of certainty/quality of evidence	
	Relatively stronger	Relatively weaker
1	The one the apple is in is the red box *dayo*.	The one the apple is in is the blue box *kana*.
2	I *know* the one the car is in is the yellow box.	I *think* the one the car is in is the green box.
3	The one the hat is in is the blue box *dayo*.	The one the hat is in is the pink box *datte*.
4	I *saw* it. The one the socks are in is the white box.	I *heard* (of) it. The one the socks are in is the orange box.

4.3.3. Procedure

All data collection activities were conducted in a playroom designed to put children at ease. Participants were tested individually and each session was videotaped with two cameras—one focused on the child, and one focused on the main experimenter. The children's task was to indicate the location of hidden objects based on the animated and ambiguous hiding scenes they witnessed and the verbal clues they received. For each hidden object (each trial), the participants heard two conflicting statements about the hidden location using one of the four pairs of contrasting terms under investigation. The objects—typical food items, articles of clothing and toys—were all familiar to the children.

Following a practice run to familiarize participants with the nature of the "game," two sets of tasks were presented to each child. Each set contained eight trials which were further divided into sets of four—one for each of the four contrasting pairs. After the first eight tasks (two sets), the procedure was interrupted to administer a version of the "Sally Anne" false-belief test (Wimmer & Perner, 1983) and a version of the "Smarties" false-belief test (Perner, Leekam, & Wimmer, 1987). In these now widely known tests, the overall standard procedures were followed. The contrastive pairs in each set of trials were presented to the participants in a counterbalanced order to offset potential fatigue and practice effects. All other experimental design variables (colors of containers, pairs of colors of containers, number of times correct answer was provided by each animal, position of the animal providing the correct answer, position of the colored containers, color of the correct answer, order of presentation of correct answer, order of the modality contrasts tested, and the order of the trials) were also randomized.

Each individual trial consisted of two parts and began with the child watching a thief surreptitiously and, from the child's viewpoint, ambiguously, hide four different objects—one at a time—in one of four different pairs of containers. Once all four objects had been hidden, the thief exited the scene. At this point, the experimenter said to the child, "See? The thief has stolen many things. Now, let's ask the animals where they are. Listen to what the animals say very carefully and tell me which colored box the things are in. First was a car, wasn't it? Let's ask the rabbit and the frog which box the car is in." Once the child's attention seemed focused, the experimenter then addressed the cartoon animals on screen in turn saying, for example, "Hello rabbit. Hello frog. Which box is the car in?" When addressed by the experimenter, the animals, which were activated by the experimenter using the cursor, offered conflicting statements about the location of the toy in question. In the presentation of Contrast 1 stimuli, for example, the rabbit says, "The one the car is in is the red box *dayo*," and the frog claims that, "The one the car is in is the blue box *kana*." Once the conflicting information had been presented, the experimenter addressed the child by asking, "Which container is the car in?"

The procedure with the second set of four trials was identical to the first. The animated thief and the hidden objects, though were different, as were the nature (shape and color) of the containers. As stated earlier, all experimental design variables were counterbalanced. By conducting two, two-part trials, each consisting of four contrast tests for a total of 16 different decisions, each child had a chance to respond to each of the contrastive pairs four times. A perfect score for each contrastive pair then was four—one point for each correct decision based on that contrastive pair.

The entire experimental procedure, including greetings, warm-up activities, and the false-belief tests, required approximately 30 min per child.

4.3.4. Results

4.3.4.1. Understanding of speaker certainty and evidentiality. The mean scores for the four-modality contrasts by age group are shown in Table 4 (in graphic form in Fig. 1). The figure

Table 4
Mean scores for each contrast by age group

Age group	*N*	Contrast (maximum = 4)			
		Certainty particles (*yo-kana*)	Certainty verbs (*know-think*)	Evidentiality particles (*yo-tte*)	Evidentiality verbs (*see-hear that*)
3-year-olds	25	3.0	2.2	2.5	2.1
4-year-olds	24	3.3	2.9	2.5	2.1
5-year-olds	24	3.3	3.2	2.8	2.2
6-year-olds	24	3.5	3.3	3.3	2.8

suggests that the children, in general, performed better on the certainty contrasts than the evidentiality contrasts, and that they understood speakers' epistemic states better when they were conveyed by particles than by verbs. We can see that the 3-year-olds already have a comparatively good understanding of certainty as conveyed by particles (*yo-kana*). As for certainty conveyed by verbs (*know-think*), though the 3-year-olds' performance was poor, it was understood fairly well by the 4-year-olds. It is also characteristic that evidentiality is slow to be understood; the understanding of the two evidentiality contrasts remain poor with only a subtle development till the age 5, but between ages 5 and 6, it shows a sharp rise in both forms. Certainty and evidentialty were both understood better when they were conveyed by particles than when they are encoded by verbs by all age groups.

A three-way analysis of variance was performed on the data, with age (four levels: 3, 4, 5, and 6) as a between-subjects variable and modality type (two levels: certainty, evidentiality) and linguistic form (two levels: particle, verb) as within-subjects variables. Results showed significant main effects for age, $F(3, 93) = 8.15$, $p < .001$, modality type, $F(1, 93) = 29.00$, $p < .001$, and linguistic form, $F(1, 93) = 22.31$, $p < .001$. No significant interactions were revealed in the analysis of variance. These results indicate that evidentiality contrasts were significantly more difficult

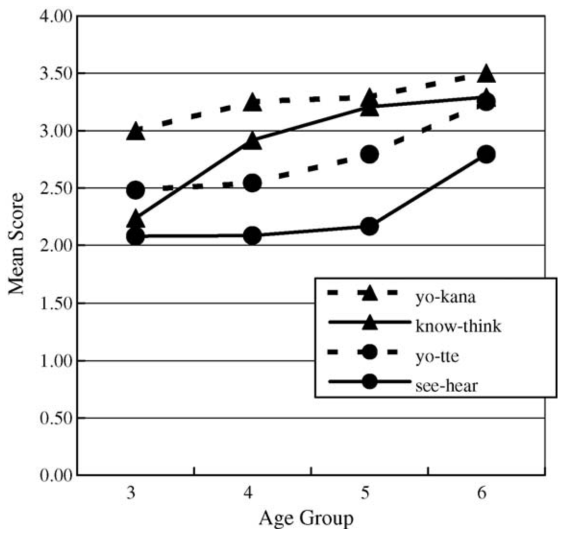

Fig. 1. Relative performances on the four tested contrasts.

Table 5
Guttman scalogram patterns for a four-item scale

Pattern	1	2	3	4	5	Other patterns	*N*
yo-kana (78%)	−	+	+	+	+		
know-think (65%)	−	−	+	+	+		
yo-tte (57%)	−	−	−	+	+		
see-hear (34%)	−	−	−	−	+		
3-year-olds	2	3	5	3	1	11	25
4-year-olds	0	2	6	2	3	11	24
5-year-olds	2	1	4	6	5	6	24
6-year-olds	0	1	1	9	7	6	24
Total	4	7	16	20	16	34	97

Note. A minus sign means a child failed the task in question; a plus sign means the child passed. The percentage figures in the parentheses indicate the proportion of children correct on the task.

than certainty contrasts and verb contrasts significantly more difficult than particle contrasts. A Bonferroni multiple comparison test conducted for age revealed that the 6-year-olds performed significantly better than the 3- and 4-year-olds ($p < .05$).

In order to examine directly the order in which the children become able to handle the four-modality pairs, we carried out a Guttman scalogram analysis. Guttman's (1944) basic idea is that a set of ranked items is considered to establish a unidimensional continuum if a respondent who answered positively on a given item in the list also answered positively on all previously ranked items. We regarded a child as correct on the task when he/she answered correctly on three or four out of altogether four trials of each modality task. Then the four tasks are ordered from the easiest to the hardest in terms of the proportion of children correct on the tasks. The four tasks were ranked in the order of *yo-kana* (the easiest), *know-think*, *yo-tte*, and *see-hear* (the hardest). An ideal response pattern was constructed so that if a child was correct on a given item, he or she must be correct on all easier items. We then examined to what extent the children's response patterns fit this ideal model. Table 5 shows the resulting Guttman scalogram for the four tasks. As shown in the table, the responses of 63 children out of 97 (65%) fit the Guttman scale exactly. The coefficient of reproducibility was .91. (Values greater than .90 indicate the items constitute unidimensional series and that they are scalable.) The scale indicates that before the children understand speaker certainty when it is conveyed by the verbs tested, they are sensitive to speaker certainty as conveyed by particles. The results also support our predictions that (1) the ability to distinguish degrees of speaker certainty develops earlier than the ability to evaluate the quality of evidence a speaker possesses; and that (2) the ability to understand the epistemic state of a speaker when conveyed by sentence-final particles precedes ability to understand a speaker's epistemic state when that information is conveyed by mental state verbs.

4.3.4.2. Comprehension of epistemic modality and false-belief understanding. In order to test our third hypothesis (comprehension of speaker certainty/evidentiality correlates positively with understanding of other's false-belief), Pearson's correlation test was conducted between total FB test scores and each modality task score. It revealed significant correlations between FB test score and certainty particle pair (*yo-kana*) score ($r = .22$, $p < .05$), certainty verb pair (*know-think*) score ($r = .039$, $p < .001$), and evidentiality verb pair (*see-hear*) score ($r = .22$, $p < .05$).

We divided the subjects based on their FB test performance. The subjects were grouped as "pass" when all three trials were correct, as "transitional" when one or two trials were correct, and

Table 6
Performance on false-belief tests and mean modality scores

FB results	Modality mean scores			
	yo-kana	*think-know*	*yo-tte*	*see-hear*
Pass ($n = 41$)	3.5	3.3	2.9	2.4
Transitional ($n = 38$)	3.1	2.6	2.7	2.2
Fail ($n = 18$)	2.9	2.3	2.2	1.9

as “fail” when none of the three trials was correct. Table 6 presents mean scores in each modality tasks according to performance on the false-belief tests regardless of age.

To determine if children’s comprehension of epistemic modality is a discriminating factor in whether or not children understand false-beliefs, we conducted Wilks’ lambda tests in which the dichotomous dependent variable was whether the children passed or failed the FB tasks, and the independent variables were the scores on the four-modality tasks. The results revealed that comprehension of epistemic modality conveyed by verbs (*know-think*, Wilks’ lambda = 0.826, $F(1, 57) = 12.01$, $p < .001$, *see-hear*, Wilks’ lambda = 0.827, $F(1, 57) = 11.96$, $p < .001$) significantly relates to whether or not children pass false-belief tasks. On the other hand, the statistical analyses revealed that comprehension of modality as conveyed by particles has no significant relation with false-belief understanding.

5. General discussion

Knowing just how children become able to determine the reliability of potential informants is an important part of the epistemic puzzle that makes up our understanding of what it means to have a fully explicit theory of mind. The reliability of linguistically conveyed information may be determined on the basis of multiple factors, but two related and fundamental ones would appear to be understanding of a speaker’s degree of certainty about or commitment to the content of an uttered proposition, and the nature of evidence upon which the expressed proposition is based. Indeed, quality of evidence is one factor in determining speaker certainty. Speakers ought to be more certain the better the quality of their evidence is. If we rely on our own direct perceptual experiences more than we do on the testimony of others, when we hear from a speaker who herself claims to only have hearsay evidence for the uttered proposition, we ought to be more leery than when the speaker claims to have directly witnessed the state of affairs about which she speaks.

The speaker certainty tasks in our study required that the participants determine reliability of information based on speaker attitude when it was encoded as an attitude toward the information conveyed by the proposition. In contrast, the evidentiality tasks employed required that the participants determine reliability of information when speaker certainty needed to be inferred from the quality of evidence possessed by the speaker for the state of affairs described in the proposition.

Results of our study showed that children as young as 3 can make use of information about a speaker’s attitude toward the expressed proposition when that attitude is indicated in sentence final particles, but that even 6-year-olds continue to have difficulty in evaluating the quality of evidence (first hand versus hearsay) as it is indicated in mental state verbs such as *see* and *hear that*. The results then suggest that particles are easier for young children to make use of in identifying the location of hidden objects than are verbs, and that understanding of speaker certainty is easier than evaluating the quality of evidence (directly witnessed being better than hearsay).

Existing studies of young children's knowledge acquisition suggest that 3-year-olds understand that seeing leads to knowing. Children of this age also seem to have intuitive or implicit understanding that hearsay evidence is less reliable than direct perceptual evidence (Robinson & Whitcombe, 2003; Zaitchik, 1991). However, we do not yet know how robust 3- and 4-year-olds' understanding of the nature of hearsay evidence is. Preschool children tend to accept what other people say at face value, unless there are specific clues indicating otherwise. This tendency to trust the speaker and accept what was said is essential for efficient early knowledge acquisition, where children do not have enough knowledge to make their own judgments about the plausibility of the statements made by others. We suggest that this tendency toward acceptance in early verbal communication may explain the late development of evidential reasoning in interpreting utterances. Thus, even though a speaker indicates by the use of a hearsay particle that she only has indirect (hearsay) evidence for P, children may interpret the utterance as communicating that the speaker believes P is true. Our results suggest that by the age of 5, children seem to be able to understand that the speaker's propositional attitudes can be inferred on the basis of verbal expressions of hearsay evidence.

Further analysis of results revealed that both particles that indicate speaker certainty about the proposition and particles that indicate the quality of evidence available to the speaker are understood before basic verbs that encode roughly the same meaning are. That particles are understood earlier than verbs are can be accounted for in a number of ways. It has been suggested that linguistic items that appear with high frequency in child-directed speech will receive special cognitive salience in child's mind, and as such, may have significant influence to a child's semantic and cognitive development (Choi & Gopnik, 1995; Gopnik, Choi, & Baumberger, 1996). In the present study, the analysis of Japanese corpus data confirmed the high frequency of sentence-final particles in the mother's speech, and thus, the result tends to support this hypothesis. However, as our data consists of only one child–mother pair, further data is needed to confirm the exact causal relations.

In addition, although verbs in Japanese appear at the end of sentence and so the saliency based on the sentential position may not differ between verbs and sentence-final particles, the latter may generally receive higher cognitive saliency than the former for socio-cultural reasons. It has been observed that Japanese child-directed speech tends to put more emphasis on sharing of social and interactional (or attitudinal) information, rather than on exchange of propositional information (Clancy, 1985; Fernald & Morikawa, 1993). In other words, in addition to accurately grasping the main information conveyed by an utterance, both adult and child hearers are expected to pay attention to, and respond appropriately towards, the speaker's various attitudes that accompany the utterance. Japanese sentence-final particles encode speaker's epistemic or illocutionary attitudes, and as such, they are believed to play a highly significant role in Japanese-style conversation.

It is our belief that early understanding of particles provides important information about children's understanding of other's epistemic mental states in general. One of the most intriguing results of our study is that understanding of particles does not correlate with understanding of false-belief. It is now widely assumed that passing false-belief tasks involves explicit representational theory of mind (Perner, 1991). However, there is currently little agreement as to whether children can grasp others' mental states prior to that, and if they do, how. It has been reported that children who fail false-belief tasks show procedural, unconscious grasp of other's mental states through eye-gaze, and such understanding has been called "implicit" understanding of another's mind (Clements & Perner, 1994; Ruffman, 2000). We are inclined to believe that Japanese 3-year-olds' understanding of other's knowledge states may similarly be of an implicit kind, though the concept of implicit understanding itself requires further clarification. What the current study has shown, however, is that a consistent, working understanding of knowledge states precedes fully representational understanding of (false) beliefs. Moreover, our study indicates that the early working understanding of others as epistemic beings is deeply situated in frequent, continuous, and largely verbal interaction. The question of which aspects of verbal communication have the most direct causal relation to the development of such understanding remains a topic of future research.

Acknowledgments

The authors would like to thank Yui Miura for her help in constructing the animations used in our study. We also would like to thank Nobuyuki Kawai and Akira Nakayama for their help with the statistical analyses, and Peter Bryant for helpful and constructive comments. Lastly, we are very grateful for the comments provided by two anonymous reviewers whose recommendations we have tried to incorporate into the final version of the paper.

References

Abbeduto, L., & Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. *Journal of Child Language*, *12*, 621–641.

Aksu-Koc, A. (1988). *The acquisition of aspect and modality: The case of past reference in Turkish*. Cambridge: Cambridge University Press.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford: Oxford University Press.

Blakemore, D. (1987). *Semantic constrains on relevance*. Oxford: Oxford University Press.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, *18*, 906–921.

Chafe, W., & Nichols, J. (Eds.). (1986). *Evidentiality: The linguistic coding of epistemology*. Norwood, NJ: Ablex.

Choi, S. (1995). The development of epistemic sentence-ending modal forms and functions in Korean children. In J. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 165–204). Amsterdam: John Benjamins.

Choi, S., & Gopnik. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, *22*, 497–529.

Clancy, P. (1985). The acquisition of communicative style in Japanese. In B. B. Shieffelin & E. Ochs (Eds.), *Language socialization across cultures*. Cambridge University Press.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377–395.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and false belief. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind*. Psychology Press.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808.

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics*, *12*(2), 97–141.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, *62*, 1352–1366.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variation in Japanese and American mothers' speech to preverbal infants. *Journal of Child Language*, *16*, 477–501.

Fitneva, S. A. (2001). Epistemic marking and reliability judgments: Evidence from Bulgarian. *Journal of Pragmatics*, *33*, 401–420.

Furrow, D., Moore, C., Davidge, J., & Chiasson, L. (1992). Mental terms in mothers' and children's speech: Similarities and relationships. *Journal of Child Language*, *19*, 617–631.

Gopnik, A., Choi, S., & Baumberger, T. (1996). Cross-linguistic difference in early semantic and cognitive development. *Cognitive Development*, *11*, 197–227.

Gopnik, A., & Graf, P. (1988). Knowing how you know: Young children's ability to identify and remember the sources of their beliefs. *Child Development*, *59*, 1366–1371.

Guttman, L. (1944). A basis of scaling quantitative data. *American Sociological Review*, *9*, 139–150.

Johnson, C. N., & Maratsos, M. P. (1977). Early comprehension of mental verbs: Think and know. *Child Development*, *48*, 1743–1747.

Johnson, C. N., & Wellman, H. M. (1980). Children's developing understanding of mental verbs: Remember, know and guess. *Child Development*, *51*, 1095–1102.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Karmiloff-Smith, A., Grant, J., Sims, K., Jones, M.-C., & Cuckle, P. (1996). Rethinking metalinguistic awareness: Representing and accessing knowledge about what count as a word. *Cognition*, *58*, 197–219.

Koenig, M. A., Clements, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, *10*, 694–698.

Lee, T. H., & Law, A. (2001). Epistemic modality and the acquisition of Cantonese final particles. In Mineharu Nakayama (Ed.), *Issues in East Asian language acquisition* (pp. 67–128). Tokyo: Kuroshio Publishers.

Masuoka, T. (2001). Epistemic modality. *Nihongogaku* [Japanese Linguistics], *21*(2), 6–16.

Matsui, T. (2000). Linguistic encoding of the guarantee of relevance: Japanese sentence-final particle *yo*. In G. Andersen & T. Fretheim (Eds.), *Pragmatic markers and propositional attitude* (pp. 145–172). Amsterdam: John Benjamins.

Miyata, S. (2000). The TAI corpus: Longitudinal speech data of a Japanese boy aged 1:5.20-3;1.1. *Bulletin of Shukutoku Junior College*, *39*, 77–85.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, *60*, 167–171.

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development*, *61*, 722–730.

Moriyama, T. (1992). On a sentence final mental state verb 'omou'. *Nihongogaku* [Japanese Linguistics], *11*(9), 105–116.

Olson, D. (1988). On the origins of beliefs and other internal states in children. In J. Astngton, P. Harris, & D. Olson (Eds.), *Developing theories of mind* (pp. 414–426). New York: Cambridge University Press.

O'Neill, D., & Gopnik, A. (1991). Young children's understanding of the sources of their beliefs. *Developmental Psychology*, *27*, 390–397.

Papafragou, A., & Li, P. (2002). Evidential morphology and theory of mind. In *Proceedings of the 26th Annual BUCLD*, pp. 510–520.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year-old's difficulty with false-belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125–137.

Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, *47*, 116–129.

Pinker, S. (1999). *Words and rules: Ingredients of language*. New York: Basic Books.

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development*, *61*, 973–982.

Robinson, E. J., Mitchell, P., & Nye, R. (1995). Young children's treating of utterances as unreliable sources of knowledge. *Journal of Child Language*, *22*, 663–685.

Robinson, E. J., & Whitcombe, E. L. (2003). Children's suggestibility in relation to their understanding of sources of knowledge. *Child Development*, *74*(1), 48–62.

Ruffman, T. (2000). Nonverbal theory of mind: Is it important, is it simulation, is it relevant to autisum? In J. W. Astington (Ed.), *Minds in the making: Essays in honour of David R. Olson* (pp. 250–266). Oxford: Blackwell.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mother's mental state language and theory-of-mind understanding. *Child Development*, *73*(3), 734–751.

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development*, *18*, 139–158.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*(4), 1054–1070.

Shatz, M., Wellman, H., & Silber, S. (1983). The acquisition of mental terms: A systematic investigation of the first reference to mental states. *Cognition*, *14*, 301–321.

Shirai, J., Shirai, H., & Furuta, Y. (1999). Acquisition of sentence-final particles in Japanese. In M. Perkins & S. Howard (Eds.), *New directions in language development and disorders* (pp. 243–250). New York: Plenum Press.

Talmy, L. (2001). *Toward a cognitive semantics* (Vol. 1). Cambridge, MA: MIT Press.

Thompson, S., & Mulac, A. (1991). A quantitative perspective on the grammaticization of epistemic parentheticals in English. In E. C. Traugott & B. Heine (Eds.), *Approaches to grammaticalization: vol.2* (pp. 313–329). Amsterdam: John Benjamins.

Ullman, M. (2004). Contribution of memory circuit to language: The declarative/procedural model. *Cognition*, *92*, 231–270.

Whitcombe, E. L., & Robinson, E. J. (2000). Children's decisions about what to believe and their ability to report the source of their belief. *Cognitive Development*, *15*, 329–346.

Wilson, D., & Sperber, D. (1993). Linguistic form and relevance. *Lingua*, *90*, 1–25.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 41–68.

Zaitchik, D. (1991). Is only seeing really believing? Sources of the true belief in the false belief task. *Cognitive Development*, *6*, 91–103.