Children's Acquisition of Evidentiality

Ozge Ozturk and Anna Papafragou
University of Delaware

1. Introduction

This paper is concerned with the acquisition of the semantics and pragmatics of evidentiality. Evidentiality markers encode the speaker's source for the information being reported in the utterance. While languages like English express evidentiality in lexical markers (*I saw that it was raining* vs. *I heard that it was raining*), other languages grammaticalize evidentiality. In Turkish, for all instances of past reference there is an obligatory choice between the suffixes *-DI* (realized as *–di, -dı, -du, -dü, -ti, -tı, -tu, -tü* depending on the vowel harmony) and *–mIs* (realized as *-mis, -mıs, -mus, -müs* depending on the vowel harmony). These past-tense morphemes also carry evidential meanings: the morpheme *–DI* is used to describe witnessed events and the morpheme *–mIs* is used to describe information acquired from someone (hearsay) or some clue (inference):

(1)	Cocuk oyun oyna	-DI	/ –mIs
	Child game play	PAST direct evid.	/ indirect evid.
	'The child played'	(I saw it	/ I heard it or I inferred it)

The acquisition of evidentiality poses several challenges for the young learner. Children have to realize that their language encodes different types of information sources but they also have to identify and differentiate between these different types of evidential markers and map these markers onto different types of non-linguistic information sources. Since information sources are themselves abstract and unobservable, their identification as potential candidates for grammatical meaning may be a lengthy and complex process. Several studies of the acquisition of grammaticalized evidentiality, especially in Turkish (Aksu-Koc, 1988; Aksu & Slobin, 1986) and Korean (Papafragou, Li, Choi, & Han, in press), have pointed out that, even though production of evidential morphology appears around the age of two, genuine knowledge of evidentiality emerges much later (but cf. Choi, 1995).

2. Linguistic Evidentiality and Source Monitoring

A question of interest is whether and how the emergence of evidentiality depends on the nature of its cognitive prerequisites. It has been claimed in the literature that the early advantage of concrete over mental/abstract words in the child's lexicon is due to early difficulties with mentalistic concepts (Gopnik & Meltzoff, 1997). Following this *Conceptual Hypothesis* (cf. Smiley & Huttenlocher, 1995), one could argue that the acquisition of evidentiality is delayed because of the complexity and abstractness of the underlying source concepts: to monitor the source for a piece of information, children have to be able to differentiate between the possible events that could lead to beliefs, remember which event took place, and relate that event to a particular belief.. An alternative (perhaps complementary) explanation for the late emergence of mental words (and evidentials) comes from the *Mapping Hypothesis* (Snedeker & Gleitman, 2004): even if they have acquired the relevant concepts, children may have difficulties discovering the correspondence between these concepts and specific words/morphemes in their language, especially since this correspondence is hard to glean from individual contexts of linguistic use.

One way of teasing apart the contribution of conceptual and mapping factors to the acquisition of evidentiality is to conduct non-linguistic tasks of source reasoning with young language learners and compare results from such tasks to children's knowledge of linguistic evidentials. So far non-linguistic source monitoring studies have been conducted with English-speaking children and have produced somewhat mixed results (but see Papafragou et al., in press). These studies show that three-year-olds understand the relationship between seeing and knowing: someone who has not seen an object will not know about that object (Pillow, 1989; Pratt & Bryant, 1990). However, they have difficulty identifying the source of their beliefs (Wimmer, Hogrefe & Perner, 1988). In one study (Gopnik and Graf, 1988), children learned about the contents of a drawer in three different ways (they saw the content of the box, they were told about it or they inferred what is in the box from a clue). Next they were asked how they knew what is in the box. Then children were asked: "How do you know there is an *x* inside, did you see it, did I tell you about it or did you figure it out from a clue?". Findings show that 3- but not 5-year-olds had difficulty identifying the sources of their beliefs. Moreover, certain types of sources seem to be more difficult for the young children to identify than others: it is especially challenging for young children to identify inference as a source of beliefs before at least before the age of 5 (Sodian & Wimmer, 1987). Other research, however, shows that young children do encode the origins of mental representations to some extent. Three-year-olds' perform better with some sources (e.g., seeing) than with others (e.g., being told). In fact, when asked to report whether their beliefs were due to either seeing or telling, 3-year-olds' performance is well above chance (O'Neill & Gopnik, 1991; Whitcombe & Robinson, 2000).

Children also engage in source monitoring if they need to contrast and evaluate conflicting sources of information: for instance, they rightly trust their own visual perception more than conflicting verbal reports from others (Mitchell, Robinson, Nye & Isaacs, 1997). Furthermore, 3- and 4-year-olds are more likely to believe what they are told by an adult who has had visual evidence over an adult who has not (Robinson, Champion & Mitchell, 1998).

Focusing on evidential markers in Turkish, the experiments reported in this paper seek to distinguish between conceptual and mapping factors in the acquisition of grammatical evidentiality. To this end, we compare systematically Turkish children's acquisition of evidential markers (Exps. 1-3) and their non-linguistic source-reasoning development (Exp.4). To the extent that linguistic and conceptual development can be shown to proceed hand in hand, the Conceptual hypothesis will gain support; alternatively, a learning outcome that shows grasp of non-linguistic source monitoring before the corresponding linguistic-evidential distinctions are acquired should offer support to the Mapping hypothesis.

3. General Methods

3.1 Participants

A total of 30 monolingual Turkish-speaking children participated in this study. The children were assigned to one of the three groups on the basis of their age (**Group 1**, mean: 3;6, range: 35-48 mo; **Group 2**, mean: 4;8, range: 51-66 mo; **Group 3**, mean: 6;6, range: 68-86 mo). Each group included 10 children. All children came from upper-middle-class families and were recruited either from a preschool or a grade school in Istanbul, Turkey.

3.2 Stimuli and Procedure

Stimuli were presented on the screen of a laptop computer and consisted of animated scenarios. The digitized audio for the animations was recorded from the voice of a native Turkish speaker. Three linguistic experiments were followed by one non-linguistic experiment in a single session for each participant in the order presented below. Children were tested individually in a quiet room outside their classroom. All participants completed the experiment.

4. Experiment 1: Production of Evidential Morphology

4.1 Procedure

In the production experiment we attempted to elicit children's production of the evidential morphemes for direct evidence (*-DI*) and indirect evidence (*–mIs*). The participant's task was to tell Mickey what happened on the screen. We had three kinds of trials: 4 involved seeing, 4 hearing and 4 inference. In the seeing trials the participant watched something happen (e.g. a girl jump over a stone).

In the hearing trails, the participant heard the character in the animation utter a sentence (e.g. a woman said: "I went shopping today"). In the inference trials, the participant saw some hints indicating something must have happened. After each trial the experimenter encouraged the participant to tell Mickey what happened by beginning to utter a sentence[1]. However, she did not finish the sentence and let the participant finish it:

(2) Kiz tas -in ust -u -nden…
Girl stone-GEN above-3sg.-abl.…
'The girl over the stone…'

If the participant witnessed the event (seeing trials), he/she was expected to employ the evidential morpheme for direct evidence *-DI*. In the hearing and inferring trials, however, the participant was expected to use the indirect evidence morpheme *–mIs*. Two pseudo-random orders of presentation were employed for a total of 12 trials.

4.2 Results and Discussion

A 3 (Age: Group 1, Group 2, Group 3) x 3 (Item type: See, Hear, Infer) ANOVA with the proportion of correct responses as the dependent variable and Item Type as a within subjects factor revealed no significant main effect of Age. However, a significant main effect of Item type was found ($F(1, 117) = 9.941$, $p<.002$): overall, children performed better in the See and Hear Items (M_{see}= 66%, M_{hear}= 63.6%, M_{infer}= 43.3%). Moreover, the analysis revealed a significant interaction between Item Type and Age $F(2, 117) = 10.647$, $p<.000$): while children's performance improved with age for See and Hear items, their performance seemed to decrease with age for the Infer items (see Fig.1).[2]

Next, we had a closer look at the performance of each group. One-sample t-tests revealed that the performance of children in Group 1 was not significantly different from chance for any type of item. Children in Group 2 performed significantly differently from chance for the Hear items only ($t(39) = 2.333$, $p = .025$). Finally, performance of the oldest children (Group 3) was significantly different from chance for all item types (t_{see} (39) = 6.121, p=.000, t_{hear} (39) =

[1] Turkish is an SOV language so the verb's unmarked position is at the end of a sentence. Since the evidential markers are verbal suffixes, by not finishing the sentence the experimenter avoids using an evidential and gives the participant the chance to do so.

[2] One explanation for this potentially puzzling fact is that it might have been hard for these children (who have otherwise acquired the direct evidence marker) to decide how much of an event should be observed in order for the event to count as a Seeing rather than an Inferring event. In other words, the oldest children seem to have treated Inference trials as Seeing trials. We do not know whether this confusion is due to the experimental materials or is a more general characteristic of the acquisition of the inferential but we are addressing these possibilities in ongoing experimental work.

2.333, p=.025, t_{infer} (39) = -4.684, p=.000). In sum, our results indicate that even the youngest children have acquired the distribution of the two Turkish morphemes under study and they correctly select them to denote past-tense meaning; however, the identification of the evidential meaning of the morphemes increases significantly with age and is still developing in 6-year-olds.

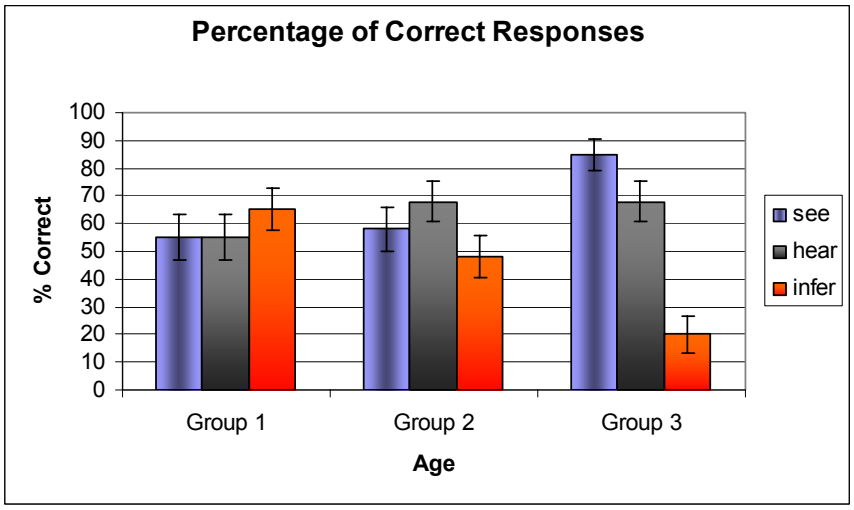

Figure 1. Percentage of correct responses (Experiment 1).

5. Experiment 2: Semantic Comprehension of Evidentiality

5.1 Procedure

This experiment was conducted to see if children can attribute a sentence with an evidential morpheme to a speaker depending on the speaker's access to information. In two separate (within subject) conditions we contrasted **seeing** vs. **inferring** and **seeing** vs. **hearing** in separate types of story. Each story involved two animals and one human character. In the **seeing vs. inferring** trials, an animal watched a character do something (e.g. a bird came in and watched a girl knock over a glass of lemonade). Then the animal left and a second animal came in and saw some evidence indicating what might have happened in the first scene (e.g. the bird left and a cat came in; the cat saw the girl looking sadly at the knocked over glass of lemonade). Next, the two animals reappeared on the screen. Then the participant heard a pre-recorded voice uttering one sentence containing either the direct evidence morpheme *–DI* or the indirect evidence morpheme *-mIs* :

(3) Kiz limonata -yi devir -DI. / -mIs
Girl lemonade-Acc. knock over-PAST direct. / indirect
'The girl knocked over the lemonade' (direct /indirect)

The experimenter then asked the participant: "Which animal said that?" It was expected that if participants understood the difference in the kind of evidence associated with the two morphemes, they would pick the animal that saw what happened when the sentence included the direct evidential morpheme *–DI* and the animal that inferred what happened when the sentence included the indirect morpheme *–mIs* .

In the **seeing** vs. **hearing** trials, the participant again watched an animal come in and watch a character do something (e.g. a dog came in and watched a boy pick apples). Then the animal and the character left and a second animal (e.g. an elephant) came in. Next, a new character, named Ali, came in, was introduced to the participant, turned to the second animal and whispered to him revealing what happened in the previous scene. Next, curtains were lowered and the two animals (e.g. the dog and the elephant) reappeared on the screen. Then, the participant heard a pre-recorded voice uttering one sentence with either the direct evidence morpheme *–DI* or the indirect evidence morpheme *–mIs*:

(4) Cocuk elma topla -DI. / –mIs
Child apple pick PAST direct / indirect
'The child picked apples' (direct /indirect)

The experimenter then asked the participant: "Which animal said that?" The participant was expected to match the sentence with the evidential morpheme *–DI* with the animal which had witnessed the event and the indirect evidence morpheme *–mIs* with the animal which had heard what happened from someone.

There were 12 trials in total (6 involved seeing vs. inferring and 6 hearing vs. seeing)[3] administered in two different random and fixed orders.

5.2 Results and Discussion

A 3 (Age: Group 1, Group 2, Group 3) x 2 (Item type: See vs. Hear, See vs. Infer) ANOVA with the proportion of correct responses as the dependent variable and Item Type as a within subjects factor revealed no significant main effect of Age. However, a significant effect of Item type was revealed ($F(1, 183) = 4.862$, $p = .029$): overall, children performed better in the See vs. Infer Type of Items than in See vs. Hear Type of Items ($M_{see/infer} = 61\%$, $M_{see/hear} = 48\%$). There was no significant interaction between Item type and Age.

A closer look at the performance of each age group on each type of story (one-sample t-tests) revealed that only the oldest group's performance on the See vs. Infer type of story was significantly different from chance (mean: 65%, $t_{see/infer}(65) = 2.564$, $p=.013$).

[3] The experiment did not involve a condition comparing hearing and inferring since the same morpheme ***–MIS*** is employed for both purposes in Turkish.

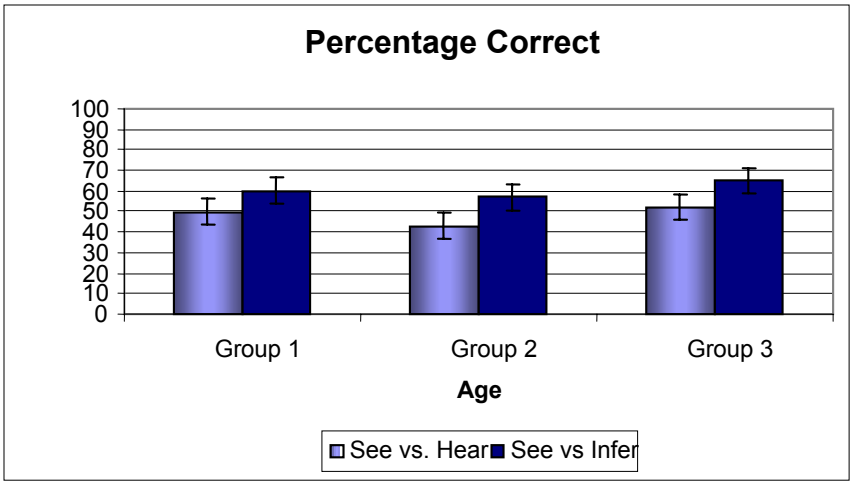

Figure 2. Percentage of correct responses (Experiment 2).

Our results for the semantic comprehension task show that children in the two younger age groups cannot match the evidential markers to the right speaker in any of the trial types. Even the children in the oldest age group (mean = 6;6) have limited success in this task. Note that, to be successful in this task, the child is required to take into account the information perspective of the speaker. This is a multi-dimensional task: first, the child has to have knowledge about the mode of acquisition of information about an event; second, she has to have assumptions about what type of information is a legitimate basis for talking about an event. Finally, she has to coordinate these two dimensions with one another. Other studies have also found difficulties with the process of attributing an utterance to a potential speaker, especially if evidential marking is the basis of the attribution (Papafragou et al., in press).

6. Experiment 3: Pragmatic Comprehension of Evidentiality

6.1 Procedure

This experiment was conducted to investigate if children trusted a character that used either the direct evidence morpheme *-DI* or its full verb counterpart ("I saw that…") more than a character who employed the indirect evidential morpheme *-mIs* or its full verb counterpart ("I heard that…"). The experiment consisted of 8 stories, each involving one box and two animals. The experimenter informed the participant that they were going to play a game to find the content of a box. The participant was told that all of the boxes were going to be opened at the end of the game to see whether or not the participant was right in her choice. In the beginning of each story, both animals and the box appeared on the screen. The animals took turns and uttered conflicting statements about the content of the box. In 4 of the 8 trials the sentences the

animals produced included the main verbs *gor-mek* 'to see' and *duy-mak* 'to hear':

(5) Bu kutu-da bir helikopter ol-dug -u-n-u gor-du -m
This box –loc. one airplane be-Nomin.-3sg.-acc. see-past.-1sg.
'I saw that there is a helicopter in this box'

(6) Bu kutu-da bir ucak ol-dug -u-n -u duy-du -m
This box –loc. one airplane be-Nomin.-3sg.-acc. hear-past.-1sg.
'I heard that there is an airplane in this box.'

The participant was expected to trust the animal which employed the main verb 'see' more than the animal employing the main verb 'hear' -hence to conclude that there is a helicopter in the box. In the remaining 4 stories, the animals produced sentences with either the direct evidence or the hearsay morpheme:[4]

(7) Bu kutu-da bir ucak var -di
This box-loc. one airplane to.be-past & direct evidential
Intended reading: "I saw that there is an airplane in this box."

(8) Bu kutu-da bir helikopter var -mIs
This box-loc. one helicopter to.be-past & indirect evidential
Intended reading: "I heard that there is a helicopter in this box."

The participant was expected to trust the animal which employed the morpheme *-DI* more than the animal employing the morpheme *-mIs* - hence to conclude that there is an airplane in the box. The left-right position of the animals producing the correct answer was counterbalanced throughout.

6.2 Results and Discussion

A 3 (Age: Group 1, Group 2, Group 3) x 2 (Item type: Full verb vs. Morpheme) ANOVA with the proportion of correct responses as the dependent variable and Item Type as a within subjects factor revealed no significant main effect of Age. However, a significant main effect of Item type was found ($F(1, 117) = 5.563$, $p=.020$): overall, children performed better on the Full verb items (M_{full}= 55%, $M_{morpheme}$= 39%). The analysis revealed no significant interaction between Item Type and Age.

Next, we had a closer look at the performance of each group. One-sample t-tests revealed that children's performance in each of the groups was never significantly different from chance for either Full verb or Morpheme items.

[4] E.g. the first animal said: "There is a car in the box ***–DI***" and the second said: "There is a motorcycle in the box –***MIS.***"

Moreover, there was no difference between the two types of item within each age group.

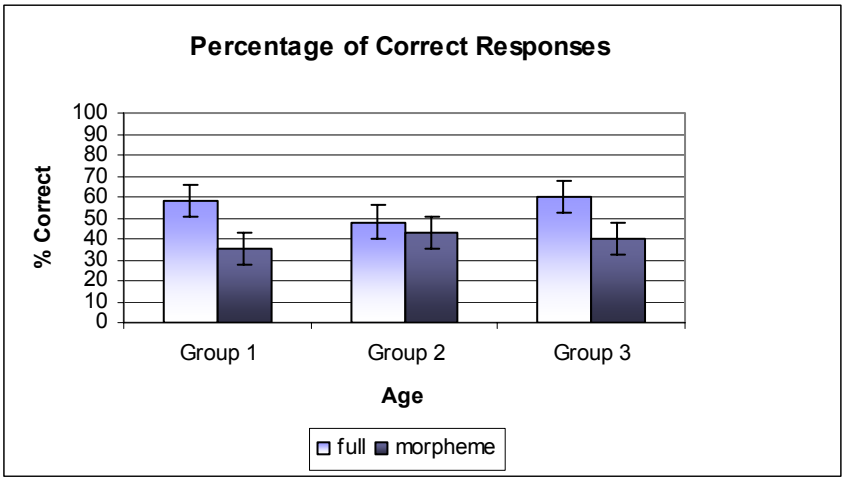

Figure 3. Percentage of correct responses for full verbs and evidential morphemes (Experiment 3).

The results of the pragmatic comprehension experiment showed that children cannot compute the pragmatic effects (speaker reliability/certainty) associated with the use of evidential markers. This finding was not surprising considering the problems Turkish children had in our earlier production and semantic comprehension experiments. These results are also in line with previous research (Papafragou et al., in press; Aksu-Koc, 1988).

Taken together, findings from Exps. 1-3 show that Turkish children produce the morphemes *–DI* and *–mIs* appropriately for past events. However, they seem to have difficulty differentiating between the two morphemes in terms of their evidential semantic functions and computing the pragmatic effects associated with these functions. These findings are remarkable given the fact that these suffixes are obligatory for past-tense reference and hence very frequent in the input. What is it that makes the acquisition of the evidential features of these markers so hard? We next turn to the two hypotheses introduced in the beginning of this paper and present results from a non-linguistic source monitoring task with the same participants designed to adjudicate between them. Recall that, according to the Conceptual hypothesis, the underlying abstract source concepts constitute the difficulty behind the acquisition of evidentiality: this hypothesis predicts that the same participants who failed on the linguistic tasks would fail on a non-linguistic source monitoring task in which they are asked to report how they acquired a piece of information. Alternatively, on the Mapping hypothesis, the problem of acquiring evidentiality is not (exclusively) conceptual in nature: this hypothesis leaves open the possibility that participants

who have not acquired linguistic evidentials might show success on non-linguistic counterparts of our linguistic tasks.

7. Experiment 4: Reporting One's Own Sources

7.1 Procedure

This experiment asked whether children were able to report their own sources of information. There were three types of trials. In the seeing trials, the participant saw something happen on the screen (e.g. a fairy fly from the flower to the tree). In the hearing trials, the participant heard a character on the screen utter a sentence (e.g. "I went shopping today"). In the inference trials, the participant saw some hints indicating that something had happened (e.g. a sad-looking boy next to a knocked-over glass) and she was expected to infer what happened (here, the boy had knocked the glass over). After each trial, the experimenter asked the participant to report what happened (using a question that did not include an evidential morpheme):

(9) Ne ol- dug-un -u bil-iyor mu-sun?
What happen-nominalizer-3.sg-Acc.know-Pr.Prog.Q.-2.sg
'Do you know what happened?'

Immediately after the participant's answer, the experimenter asked about the participant's source of information:

(10) Ner -den bil -iyor -sun?
where-abl. know-Pr.Prog -2sg
'How do you know?'

(11) Gor-du -n -mu? Duy-du -n -mu? Anla -di -n -mi[5]?
see-Past-2sg.-Q. Hear-Past-2sg.-Q. Understand-past-2sg.-Q
'Did you see? Did you hear? Did you understand?'

There were 12 trials in total (4 involving seeing, 4 hearing and 4 inference) administered in two pseudo-random orders.

7.2 Results and Discussion

A 3 (Age: Group 1, Group 2, Group 3) x 2 (Item type: See, Hear, Infer) ANOVA with the proportion of correct responses as the dependent variable and Item Type as a within subjects factor revealed no significant main effect of Age. However, a significant main effect of Item type was found ($F(1, 117) = 29.425$, $p=.000$): overall, children performed better in the See and Hear items than in the

[5] There is no Turkish counterpart for the English verb *infer* so we decided to use the verb *anlamak* 'understand' in this study.

Infer items (M_{see}= 73%, M_{hear}= 73%, M_{infer}= 37%). The analysis also revealed a significant interaction between Item Type and Age ($F(2, 117) = 5.213$, $p=.007$): participants performed better in the See and Hear type of items with age, but their performance decreased in Infer items as they grew older. (As in Exp. 2, this effect may be spurious: children may overestimate the visual component of the Inference trials, or they may be less willing to accept the verb "understand" as a rough synonym of "infer").

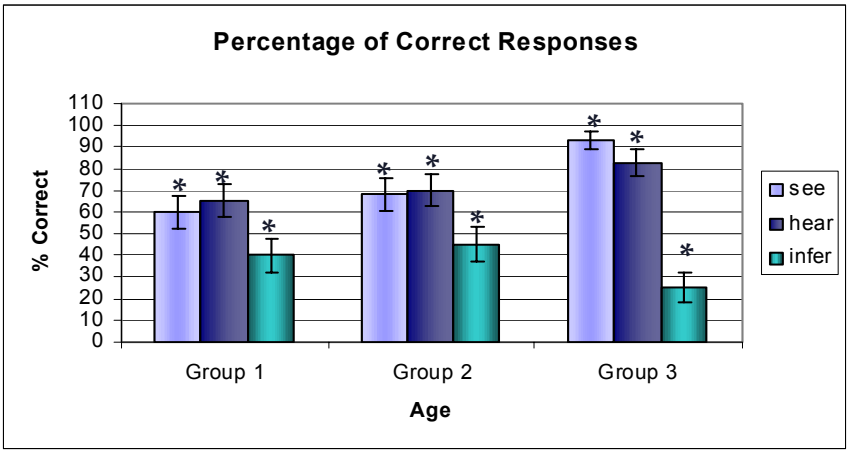

Figure 4. Percentage of correct responses (Experiment 4).

One-sample t-tests revealed that performance in each group was significantly different from chance for the See and Hear items, whereas performance for the Infer items was significantly different from chance in the oldest group only. In sum, the same children who participated in our linguistic experiments (and performed poorly in all three tasks) were able to successfully report the sources of their information, at least for cases involving visual perception and communication.

8. Conclusion

In the experiments described in this paper, we investigated linguistic evidentiality and non-linguistic source monitoring in young Turkish-speaking children to explore the role of both (i) the subtleness and abstractness of the underlying concepts (Conceptual hypothesis), and (ii) the indirectness of the correspondence between an evidential morpheme and the surrounding circumstances (Mapping hypothesis) in the acquisition of evidential morphology. Our results demonstrate that the acquisition of evidentiality poses considerable difficulties to learners of Turkish. Nevertheless, these difficulties do not seem tied to the nature of the source concepts themselves: children who cannot produce or comprehend evidential morphology accurately can

nevertheless report on their own sources of information (especially when those sources are visual or verbal). Thus the learning problem posed by the category of evidentiality seems to be best characterized in terms of mapping source concepts onto the corresponding morphemes in the language. In ongoing work, we probe the kinds of linguistic and extra-linguistic context that might create 'epiphany points' for the learner engaged in solving the mapping problem for evidentiality.

References

Aksu, A., & Slobin, D. (1986). A psychological account of the development and use of evidentials in Turkish. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology* (pp. 159-167). Norwood, NJ: Ablex.

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

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: Benjamins.

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.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Mitchell, P., Robinson, E., Nye, R, & Isaacs, J. (1996). Contamination in reasoning about false belief: an instance of realist bias in adults but not children. *Cognition*, 59, 1-21.

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

Papafragou, A., Li, P., Choi, Y., & Han, C-H. (in press). Evidentiality in language and cognition. *Cognition.*

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

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., Champion, H., & Mitchell, P. (1998). Children's ability to infer utterance veracity from speaker informedness. *Developmental Psychology*, 35, 535-546.

Smiley, P., & Huttenlocher, J. (1995). Conceptual development and the child's early words for events, objects and persons. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: young children's acquisition of verbs* (pp. 21-61). Hillsdale, NJ: Erlbaum.

Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. In G. Hall & S. Waxman (Eds.), *Weaving a lexicon*. Cambridge, MA: MIT Press.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 58, 424-433.

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

Wimmer, H., Hogrefe, G.-J., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development*, 59, 386-396.