The Acquisition of Evidentiality and Source Monitoring

Ozge Ozturk and Anna Papafragou*

University of Delaware

* Corresponding author

Department of Psychological and Brain Sciences

University of Delaware

Newark DE 19716

annap@udel.edu
Author note: We would like to thank Louis A. Arena, Roberta Golinkoff, Gabriella Hermon, Paul Quinn, and Satoshi Tomioka, for comments on this research. We are also very grateful to Gyn. Op. Dr. Zübeyde Öztürk, Özden Lülü and Ömer Öz Lülü, Business Manager of Terakki Vakfi Schools in Istanbul, Turkey, for giving us the opportunity to conduct our experiments, as well as to all the participants and the parents of our child participants for their generosity. A.P. acknowledges support from NSF grant BCS-0749870. O.O. is now with Columbia University.
Abstract

Evidentiality in language marks how information contained in a sentence was acquired. For instance, Turkish has two past-tense morphemes that mark whether access to information was direct (typically perception) or indirect (hearsay/inference). Full acquisition of evidential systems appears to be a late achievement cross-linguistically. At present there are two distinct hypotheses about why this is so. According to the first hypothesis, the acquisition of evidentiality is delayed by conceptual factors related to source monitoring (the process of identifying and evaluating information sources). According to a different hypothesis, a substantial part of the learning difficulty comes from mapping evidential markers onto the underlying source concepts (even if these concepts are already available to the child), most likely because source concepts do not correspond to observable referents in the world. Here we tested these two hypotheses in a series of experiments comparing the acquisition of evidential morphology (Exps.1-3) and the development of source monitoring (Exps.4-6) in the same group of Turkish-speaking children. We found that the semantics and pragmatics of evidential morphology in Turkish are not acquired until age 6 or 7. A comparison between linguistic evidentiality and source monitoring experiments revealed that conceptual understanding of information access develops before the corresponding concepts are linked to evidential morphemes in Turkish, thereby demonstrating that mapping difficulties underlie the late acquisition of evidentiality in Turkish. Nevertheless, our data also suggest that conceptual limitations play an important role in the acquisition of evidentiality, since in both language and source monitoring direct evidence seems to be privileged compared to indirect evidence. This work has implications for the acquisition of mental-state language and the relation between children’s linguistic and conceptual development.
Keywords: evidentiality; theory of mind; source monitoring; information access; language acquisition; language and thought; Turkish

1 Introduction

It is widely acknowledged that the development of the ability to produce and understand language is intimately connected to the learner’s ability to entertain thoughts about the objects, relations and events that language encodes. However, there is considerable debate about the precise relationship between linguistic and cognitive development. According to some commentators, the way children use (or fail to use) language reveals rather straightforwardly whether children grasp (or fail to grasp) the underlying concepts; for instance, it has been argued that the timetable of emergence of various expressions in child language more or less directly reflects the degree of their conceptual complexity (see, e.g., Smiley & Huttenlocher, 1995; Gopnik & Meltzoff, 1997). Other researchers have proposed that what children say (or do not say) is determined not only by what they can (and cannot) think about but also by the internal mechanics of language acquisition that makes the discovery of the meanings for certain expressions easier than for others (Carey, 1982; Gleitman, 1990, among others).

Here we address the relation between language acquisition and conceptual development by focusing on evidentiality, i.e., the linguistic encoding of the source for the information reported in an utterance. In English, evidentiality is expressed through lexical means (I saw that it was raining vs. I heard that it was raining), but in about one sixth of the world’s languages, evidentiality is grammaticalized through specialized and often obligatory verbal affixes, particles or other devices (see Aikhenvald & Dixon, 2001; Anderson, 1986; Chafe & Nichols, 1986; Cinque, 1999; Davis, Potts
& Speas, 2007; DeLancey, 2002; Faller, 2001, 2002; Garrett, 2000; Givón, 1982; de Haan, 1998, 2001; Ifantidou, 2001; Izhvorski, 1998; Johanson & Utas, 2000; Lee, 2010; Matthewson, Davis & Rullmann, 2006; Mayer, 1990; McCready, 2008; McCready & Ogata, 2007; Mushin, 2001; Palmer, 1986; Papafragou, 2000; Sauerland & Schenner, 2007; Speas, 2004; Willett, 1988). For instance, in Turkish, for all past events encoded in a main clause, there is an obligatory choice between two verb suffixes, -DI and –mIş, that denote the direct vs. indirect past (the two suffixes are realized as –di, -di, -du, -diů, -ti, -ti, -tu, -tü and -miş, -miş, -muş, -müş respectively depending on phonological rules related to vowel harmony and final devoicing; Aksu-Koc & Slobin, 1986; Johanson, 2003; Aikhenvald, 2004):

(1) Çocuk oyun oyna -di.
 Child game play PAST. direct
‘The child played’

(2) Çocuk oyun oyna –miş.
 Child game play PAST. indirect
‘The child played’

More specifically, –DI is the unmarked form for the past tense. Even though it can be neutral with respect to information source, through the contrast with the indirect form, it is frequently used to convey the information that an event was experienced first-hand (typically, through visual perception). The morpheme -mIş is the marked past tense form and indicates indirect access to an event –
typically, through hearsay or inference.\(^1\) The past-tense morpheme –DI is vastly more frequent than –mIṣ in both adult and child-directed speech (Ozturk, unpublished data). Choice of evidential in Turkish gives rise to pragmatic effects associated with speaker reliability, with the direct marker indicating greater speaker reliability or lack of reservations about the factuality of the event compared to the indirect marker (Johanson, 2003).\(^2\) These two morphemes belong to a wider class of inflectional verb stem suffixes in Turkish that encode voice, modality, negation, tense-aspect-mood and person/number (Turkish is an agglutinating language with a rich morphological system; Kornfilt, 1997).

Even though little is known about how evidential systems are acquired, existing evidence (see next section) suggests that the acquisition of evidentiality poses significant challenges for learners. To understand the nature of these challenges, one needs to know whether the learners’ difficulties lie with the linguistic items encoding evidentiality themselves or with the underlying source monitoring ability to recognize and reason about sources of beliefs in oneself and others (usually referred to as source monitoring). Monitoring sources of information such as direct perception, verbal report, or inference interfaces in rather obvious ways with the ability to use evidential language. First, correct production of evidential markers presupposes the capacity to recognize and remember how one found out about an event. Second, accurate comprehension of evidentiality requires attributing to the speaker some type of information access. Third, there is an intuition that the direct/perception evidential markers generally indicate higher reliability compared to various indirect (hearsay, inferential etc.) markers (compare I saw that… vs. I heard/inferred that…, or the Turkish

\(^1\) In general, the indirect marker can be used to communicate the speaker’s dissociation from the event, and its extended uses include surprise, distance, irony, story-telling (Johanson, 2003) and ‘lack of preparedness’ for new information (Slobin & Aksu, 1982). There is considerable discussion about how best to capture these dimensions of evidentiality in Turkish (see Slobin & Aksu, 1982, among others).

\(^2\) The past-tense evidential –mIṣ should be distinguished from its non-evidential homonym –mIṣ used in participial adjectives (e.g., *kar -aI- mIṣ tabak* break-Caus-Perf plate ‘broken plate’; Johanson, 2003). The indirect evidential should also be distinguished from the copula particle *imiṣ* which cannot be added to verb stems but may follow nominals/nominalized verb stems and has indirect-evidential meaning (e.g., *gül* ‘rose’ + *imiṣ* = *gülmiṣ* ‘it is/was evidently a rose’; Johanson, 2003).
direct/indirect past distinction). This intuition is based on the fact that, other things being equal, first-hand access (e.g., visual perception) is ranked higher in the speaker’s consciousness than second-hand access (e.g., hearsay or inference); the same ranking has been observed in source reasoning studies (Koenig & Harris, 2005) and has a long philosophical tradition (Coady, 1992; Dancy, 1985; Fricker, 2006; Hume, 1748; Locke, 1689). However, these points of contact between linguistic evidentiality and source monitoring have not been explored systematically. For instance, even though children’s source monitoring has been extensively studied in the cognitive development literature, this strand of work has proceeded independently from work on the linguistic development of evidentiality. Furthermore, with few exceptions, the data on the development of source monitoring have so far come from English-speaking populations rather than children acquiring languages with grammaticalized evidential systems.

In what follows, we report a series of experiments with Turkish-speaking children investigating both the acquisition of evidential morphology and the development of source monitoring. The experiments were designed to reveal the relationships between linguistic and cognitive development in this domain. To introduce the specific hypotheses motivating our experimental work, in the remainder of the Introduction we review prior work on the development of evidentiality and the relationship between evidentiality and the corresponding non-linguistic concepts. ³

The Acquisition of Linguistic Evidentiality

In a pioneering study, Aksu-Koç (1988) explored the acquisition of the Turkish evidential system in children between the ages of 3 and 6. In one of her experiments, she presented children with

³ Formal linguistic treatments of evidentiality are only beginning to emerge (see McCready, 2008, for an overview), and recent theoretical treatments of evidentiality have used different mechanisms to account for the behavior of evidential devices in different languages (e.g., presupposition: McCready & Asher, 2006; speech-act operators: Faller, 2007; probabilistic modality: McCready & Ogata, 2007). Our experimental efforts are neutral with respect to specific theories of linguistic evidentiality. Future empirical work will be needed to ascertain whether specific properties of evidential mechanisms (e.g., whether an evidential morpheme operates on speech acts or on the basic propositional content) raise additional complications for the learner mapping evidential meanings onto the corresponding linguistic forms. Future work would also have to address the possibility that evidential source might not be a semantic primitive (see Speas, 2010).
scenarios in which an event (e.g., the popping of a balloon) was fully accessible to perception and other scenarios where it was possible to infer the occurrence of an event through some clues (e.g., the popped balloon). Children were asked to describe the events. It was expected that directly perceived events should lead to the use of $-DI$, whereas inferred events should be described with $-mI\$_ş. Results show that young children do not appear to differentiate the two morphemes on the basis of their evidential meaning: even 6-year-olds produced the correct morpheme only 72% of the time for witnessed events and 68% of the time for non-witnessed events. To test whether the two morphemes were understood by young children, Aksu-Koç presented children with a sentence and asked them to guess which character in a story said it. Depending on the evidential marking of the sentence ($-DI$ or $-mI\$_ş), children were supposed to pick a character who had or had not witnessed the event. Younger children performed at chance in this task. Even 6-year-olds were correct 71% of the time with $-DI$ and only 56% of the time with $-mI\$_ş.

In another study, Aksu-Koç explored children’s knowledge of $-mI\$_ş on its hearsay interpretation. Children listened to a sentence containing the direct marker $-DI$ and were asked to report what happened, thus having to convert the direct marker $-DI$ into the indirect marker $-mI\$_ş. On other trials, participants directly experienced an event and were expected to report it using the direct marker $-DI$. Even the oldest children performed poorly in this task (although some features of the design complicate the interpretation of the results: Aksu-Koç, 1988, p. 159ff.). In sum, experimental studies on Turkish found that children’s production of evidentials was around chance levels before the age of 4 and was still non-adult-like by the age of 6, especially for the indirect marker; moreover, the ability to link the indirect marker to non-witnessed events was poor even at the age

Cross-linguistic data add to the picture of how evidential systems are acquired. A study on the acquisition of Korean evidential markers found that 3-year-old children appear to have productive command of grammaticalized evidentials for direct evidence and verbal report but their comprehension of these morphemes is not stable (Papafragou et al., 2007; cf. Choi, 1995). Moreover, 3- and 4-year old Korean children are unable to conclude that hearsay statements are less reliable than statements with the default evidential marker (Papafragou et al., 2007; cf. also Aksu-Koç & Alici, 2000 for related data on Turkish). Comprehension tasks with Tibetan 2- to 5-year-olds show that children do not seem to differentiate between the neutral and the indirect evidential marker (de Villiers, Garfield, Speas & Roeper, 2007; cf. de Villiers et al., 2009). Additional studies examined the ability of 3- to 6-year-old Japanese children to adjudicate between information marked with particles encoding direct vs. indirect evidence (Matsui, Yamamoto & McCagg, 2006): only 6-year-olds performed differently from chance, preferring utterances marked with direct evidence particles. Importantly, Japanese children succeeded earlier in a conceptually identical task involving particles encoding speaker certainty. Finally, Bulgarian 9-year-olds can use evidential markers about both how information was acquired and by whom to decide which of two utterances to believe but 6-year-olds can only use information about source type (Fitneva, 2008; cf. also Lee & Law, 2000 on Cantonese). Taken together, the cross-linguistic suggest that the acquisition of grammaticalized evidentiality is a late achievement.

---

4 Aksu-Koç (1988) reports longitudinal data where children younger than 3 appear to be producing evidentials correctly in spontaneous speech. However, given the subtlety of the distinctions involved, these data do not reveal the extent to which evidentials are truly understood (e.g., they might be interpreted as confidence markers; see de Villiers et al., 2009 on Tibetan; Robinson, 2009). Indeed, in Aksu-Koç’s elicited production data, 3-year-olds (alongside much older children) overextended the direct marker to cases of indirectly acquired information (cf. also Papafragou et al., 2003, for a similar point on spontaneous production data from Korean reported in Choi, 1995).
Conceptual vs. Mapping Difficulties in the Acquisition of Evidentiality

What might account for children’s difficulty with the acquisition of evidentiality? One possibility is that the emergence of evidentials is delayed because of their abstract and complex conceptual presuppositions. Specifically, it has been argued that children are initially more attentive to “concrete, referential and objective characteristics of situations” than to “subjectively relevant distinctions such as the speaker’s attitude to the proposition asserted” (Aksu-Koç, 1988, p.195). Similarly, within the evidential class, the order of appearance of individual morphemes has been taken to reflect at least in part the child’s developing abilities to handle various information sources. For instance, the fact that, in Turkish child speech, the indirect evidence morpheme appears later than the direct evidence morpheme has been attributed to “the further complexity of making an inference… as compared to simply accessing an experienced event from memory” (Aksu & Slobin, 1986, p.165). A related explanation has been suggested for the fact that, in Turkish, the hearsay interpretation of the indirect evidential is acquired after the inferential interpretation (ibid.). Such proposals echo similar views in studies of early lexical development, where it is widely held that much, perhaps most, of the difficulty in learning mental-state vocabulary comes from mastering the relevant concepts (cf. Bartsch & Wellman, 1995; Gopnik & Meltzoff, 1997; Smiley & Huttenlocher, 1995, on the acquisition of verbs such as think or know).

According to an alternative hypothesis, the acquisition of linguistic evidentials poses mapping problems for learners. On this hypothesis, children may be delayed in finding the correct meaning for evidentials even if they have the underlying source concepts because discovering the correspondence between evidential language and concepts is complex. There are several factors that may contribute to the language-to-concept mapping complexities. Most obviously, perhaps, evidential meanings do not correspond to observable referents in the world. Despite the fact that at
least some evidentials connect to external contingencies (e.g., witnessing an event, or hearing about
an event in a conversation), the mentalistic notion of information access is only indirectly related to
such contingencies. Similar difficulties have been observed in simulated learning environments
with both adults (Gillette, Gleitman, Gleitman & Lederer, 1999; Snedeker & Gleitman, 2004) and
children (Papafragou, Cassidy & Gleitman, 2007) trying to acquire mental verbs such as think and
know: the meanings of these verbs prove hard to glean from extra-linguistic (observational) cir-
cumstances alone. In the case of Turkish evidentials, the mapping task becomes more complex by
the fact that the forms used to communicate evidentiality also encode tense meanings (and need to
be distinguished from other similar-sounding forms, see fn.2). There may be other potential sources
of difficulty during the mapping process, such as discovering the structure of the linguistic domain
of evidentiality (e.g., the number of evidential distinctions within the linguistic system), etc.

To evaluate the role of conceptual vs. mapping factors in the acquisition of linguistic evidentia-
ality, one would require independent evidence from source monitoring tasks. Overall, previous lit-
erature shows that source monitoring abilities develop over time, with visual perception being un-
derstood early as a knowledge source. For instance, even young children seem to grasp the connec-
tion between seeing and knowing: 3-year-olds select the character who had visual access to an ob-
ject hidden inside a box as the one who knows what is hidden inside the box over another character
who simply lifted or pushed the box (Pratt & Bryant, 1990; Pillow, 1989; but cf. Povinelli &
deBlois, 1992; Wimmer, Hogrefe & Perner, 1988 for different results in slightly more complex
tasks). Similarly, 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). Fur-
thermore, it is recognized that visual perception (generally) offers very reliable access to inform-
ation: young children trust their own visual perception more than conflicting verbal reports from
others (Mitchell, Robinson, Nye & Isaacs, 1996) and have higher confidence in the location or
properties of an object after seeing it for themselves over being told from a speaker (Clément, Koenig, & Harris, 2004; Robinson, Haigh, & Nurmsoo, 2008).

In contrast, young children may not always realize that verbal report leads to knowledge: 3- and 4-year olds have difficulty understanding that a person who has been verbally informed about the contents of a box knows what is in the box (Wimmer, Hogrefe & Perner, 1988; but cf. Pratt & Bryant, 1990). Relatedly, 3-year-olds do not understand that a person who has not heard a particular statement is ignorant in relation to someone who has (Mossler, Marvin & Greenberg, 1976; see also Robinson, 1994; even though see Jaswal, 2010; Jaswal, Croft, Setia, & Cole, 2010). Young children sometimes continue to believe an informant who has repeatedly misled them (Mascaro & Sperber, 2009).

Finally, it has been argued that children do not fully realize that inference can be a source of knowledge until at least the age of 6 (O’Neill & Gopnik, 1991; Pillow, 2002; Sodian & Schneider, 1990; Sodian & Wimmer, 1987; Wimmer et al., 1988). In one study (Sodian & Wimmer, 1987), the child participant and one other observer were shown a transparent jar containing only blue balls. As the child watched, one ball was removed from the jar and placed in a plastic bag. The other person did not witness this transfer but was told that there was a ball in the bag and that it had been taken from the jar. Four- and 5-year old children claimed that the other person did not know the ball’s color. In contrast, 6-year olds understood that the observer would be able to infer the color of the ball in the bag without viewing it directly. Other studies have confirmed that 6-year-olds understand the role of evidence: for instance, children of this age (unlike younger children) realize that people with access to ambiguous visual (Taylor, 1988) or verbal (Sodian, 1988) information lack definitive knowledge. Furthermore, 6- and 7-year old children are able to recognize that evidence can provide a basis for inference about unseen events; they are also able to evaluate the informativeness of potential evidence in deciding between a pair of hypotheses (Sodian, Zaitchik, &
However, even children of this age do not always distinguish justified inference from mere guessing in others (Pillow, Hill, Boyce, & Stein, 2000).

A separate strand of studies looking at children’s ability to justify their own knowledge in terms of one of many potential sources has also found development of this ability between the ages of 3 and 5 (even though 3-year-olds could reliably report how they knew about something; Gopnik & Graf, 1988; O’Neill & Astington, 1990; O’Neill & Chong, 2001; O’Neill & Gopnik, 1991; Woolley & Bruell, 1996). In these studies, children typically discover the contents of a container (e.g., they are allowed to see them, they are told, or they are shown a clue such as a thematically related object), and are then asked how they found out. This procedure includes an initial training phase in which the experimenter explicitly points out different sources of information to children (e.g., “This time I’m going to give you a clue so you can figure out what’s inside […]”; O’Neill & Gopnik, 1991, Exp.3), asks them about these sources and gives them feedback. These studies found that inference was identified as a source later than perception and communication. Other tasks using deductive inference show that the tendency to explain one’s own knowledge by spontaneous appeal to inference becomes robust only after the age of 5 (Sodian & Wimmer, 1987).

Taken together, these results suggest that the source monitoring ability to recognize different information sources and reason about them in oneself and others develops along a lengthy timetable. This finding seems to support the idea that the acquisition of linguistic evidentials is delayed because of the conceptual complexity of the underlying source concepts (rather than the intricacies of mapping known source concepts onto grammatical items). It is particularly likely that difficulties with verbal report and inference impact the acquisition of reportative or inferential markers (such as the Turkish -mİṣ). However, given the variability of tasks used to test knowledge of grammaticalized evidentiality vs. source monitoring, these conclusions can only be speculative.
A more direct way of teasing apart the contribution of conceptual and mapping factors to the acquisition of evidentiality would be to conduct source monitoring tasks of source reasoning with young language learners and compare results from such tasks to the same children’s knowledge of linguistic evidentials. In an early study that took this approach, Papafragou et al. (2007) showed that 3- and 4-year-old Korean speakers who do not comprehend hearsay morphology accurately nevertheless succeed in identifying hearsay as a source of knowledge for themselves and others. This result is important since it suggests that a critical part of the difficulty with evidential morphology in Korean is not due to conceptual factors (see also Matsui, Miura & McCagg, 2006 for results on Japanese). However, this earlier study did not track the acquisition of a full, obligatory evidential system (such as the Turkish system) but of an individual hearsay morpheme which is frequently (but not obligatorily) used to report prior speech. Relatedly, the study did not address a broader range of information sources (including inference). Moreover, it did not trace the full developmental trajectory of either linguistic or conceptual evidentiality but focused on very young learners who had not yet fully acquired evidential morphology. Finally, the study used similar but not identical tasks to test linguistic vs. conceptual source knowledge. Other studies have addressed the development of source monitoring in speakers of Turkish but without using fully matched linguistic and non-linguistic tasks (see Ozturk & Papafragou, 2007; Ögel, 2007, summarized in Aksu-Koç, Ögel-Balaban, & Alp, 2009; cf. also Lucas et al., 2013).

Below we report findings from six experiments conducted with Turkish learners between the ages of 5 and 7 to probe further into the relationship between evidential language and source monitoring. We chose to look at somewhat older children, since earlier work indicated that 6-year-old Turkish learners are still undergoing important linguistic development in the domain of evidential morphology. Three of our experiments test children’s acquisition of the semantics and pragmatics of linguistic evidentiality (Exp.1-3) and the remainder test the same children’s source monitoring
abilities (Exp.4-6). The linguistic experiments target the full range of evidential meanings (direct vs. indirect: hearsay/inference) encoded in the Turkish past tense system and attempt to chart their developmental timetable. The source monitoring experiments examine the corresponding sources of information (direct evidence, i.e., perception, vs. indirect evidence, i.e., verbal report or inference) and ask how children reason about these sources and calculate their reliability. Task structure and difficulty are carefully matched across evidentiality and source monitoring studies. As far as we know, this is one of the most systematic comparisons of linguistic and conceptual development available.

Our overall goal is to contribute to theories of how evidentiality and source monitoring develop in children. Crucially, we are also interested in whether and how the conceptual basis of evidentiality constrains the acquisition of evidential grammar. Two potential outcomes are of particular interest here. If children’s performance on source monitoring tasks turns out to be higher than their performance on the corresponding linguistic evidentiality tasks, this will offer support for the role of mapping factors in the acquisition of evidentiality. In contrast, if performance is tightly coupled between linguistic evidentiality and source monitoring tasks (and especially if, in both cases, there is an asymmetry favoring direct over indirect sources, as suggested by prior work), this would reveal considerable influence of the source concepts on the developmental timetable of linguistic evidentials.

**General Methods**

**Participants**

A total of 96 monolingual Turkish-speaking children participated. Children fell into three age groups: 32 5-year-olds (mean: 5;10, range: 64-72 months), 32 6-year-olds (mean: 6;6, range: 73-84
months) and 32 7-year-olds (mean: 7;8, range: 85-96 months). All children came from upper-middle-class families and they were recruited either from a preschool or a grade school in Istanbul, Turkey, and tested individually in a quiet room outside their classroom.

**Design Overview**

Three experiments measured children’s performance with linguistic evidentiality. These linguistic experiments tested children’s production (Exp.1) and comprehension (Exp.2) of evidential morphology, as well as the pragmatic ability to link evidentiality to speaker reliability (Exp.3). An additional set of three experiments measured the same children’s performance in source monitoring tasks. These tasks tested children’s ability to report their own (Exp.4) and others’ (Exp.5) information sources, as well as the ability to relate information source to reliable knowledge (Exp.6). Note that these studies do not involve either interpreting or producing the critical evidential morphology, however, they do involve a verbal report component due to the nature of these tasks. Methodologically, the linguistic evidentiality and source monitoring experiments were designed to be as parallel to each other as possible. Theoretically, the two kinds of experiments were also tightly coupled in the sense that each source monitoring experiment targeted the conceptual underpinnings of an aspect of evidentiality tested in a specific linguistic task: the ability to remember one’s own information sources (Exp.4) is a prerequisite to producing linguistic evidentials appropriately (Exp.1); the ability to remember others’ access to information (Exp.5) is a prerequisite to attributing correct evidential language to speakers (Exp.2); finally, the ability to recognize that certain information sources are more reliable than others (Exp.6) is a prerequisite to understanding that utterances marked for such information sources convey higher speaker reliability (Exp.3).

Across all linguistic evidentiality and source monitoring experiments, half of the children with-
in each age group were randomly assigned to the *See vs. Hear* and the other half to the *See vs. Infer* condition. The *See vs. Hear* group was tested on linguistic knowledge of the evidential markers – *DI* and –*mIṣ* (on its hearsay interpretation) and on the conceptual understanding of perception and communication. The *See vs. Infer* group was tested on linguistic knowledge of –*DI* and –*mIṣ* (on its inference interpretation) and on the conceptual understanding of perception and inference. (We did not include a *Hear vs. Infer* group because the same morpheme –*mIṣ* is used for both hearsay and inference in Turkish.) The split allowed us to compare direct and indirect sources while keeping the testing sessions reasonably short. Materials for the *See* trials were identical in the two conditions.

To ensure that results were fully comparable between source monitoring and linguistic evidentiality studies, materials were swapped for half of the participants between Experiment 1 (Linguistic Production study) and its source monitoring counterpart (Experiment 4: Reporting One’s Own Sources), as well as between Experiment 2 (Linguistic Comprehension study) and its source monitoring counterpart (Experiment 5: Reporting Others’ Sources). Since materials consisted of stories shown on a laptop screen, a particular story was embedded into a source monitoring task (Exp.4) for half of the participants and into the corresponding evidentiality task (Exp. 1) for the other half. Because of the nature of the stories in Experiments 3 and 6, it was not possible to swap stimuli for these two experiments. A list of all materials is provided in the Appendix.

The evidentiality experiments were conducted in a separate session from the source monitoring experiments in the order presented below. The experiments within each session were administered in a fixed order, so as to ensure that, within the linguistic session, the easiest task (Production) would come first and the more difficult (Pragmatic) task last. The two sessions were approximately 15 minutes apart. No warm-up trials were used because of the length of each session.
Linguistic Evidentiality Experiments

Experiment 1: Production of Evidential Morphology

This experiment attempted to elicit production of the evidential morphemes for direct evidence (-DI) and indirect evidence (–mIş).

Stimuli and Procedure

Stimuli were presented on the screen of a laptop computer and consisted of 8 animated scenarios. In the See vs. Hear condition, there were 4 See and 4 Hear trials. Similarly, in the See vs. Infer condition, there were 4 See and 4 Infer trials.

In the See trials, participants watched something happen (e.g., a butterfly flew towards the roof of a house). In the Hear trials, participants did not see an event but heard someone utter a sentence about something they had done (e.g., a man said: “I spoke on the phone today”). In the Infer trials, participants saw some hints indicating that something had happened (e.g., a child with a sad face holding an ice cream cone and looking at some ice cream on the floor). After each trial the experimenter encouraged participants to say what happened by beginning a sentence and letting participants finish it:

(3) See trial

Kelebek çatı-y-a doğru… [uç-tu].
Butterfly roof-Dat towards… [fly-past dir.3sg]
'A butterfly [flew] towards the roof'
(4) **Hear trial**

Bugün telefon-da… [konuş-muş].

Today phone-Loc [talk-past indir.3sg]

'[He spoke] on the phone today'

(5) **Infer trial**

Çocuk dondurma-y-ı… [düşür-müş].

Child ice cream-Acc [drop-past indir.3sg]

'The child [dropped] the ice cream'

Turkish has a Subject-Object-Verb word order; hence the unmarked position for verbs is at the end of a sentence. Since the evidential markers are verbal suffixes, by not providing the verb the experimenter avoided using the evidential marker and gave participants the chance to do so. If children were too shy to answer, the experimenter encouraged them by providing the root of the verb but not the tense/evidential marker. If participants witnessed the event (See trials), they were expected to employ the evidential morpheme for direct evidence –DI. If participants did not witness the event (Hear and Infer trials), the indirect evidence morpheme –mIş was required. Notice that, in Hear trials, participants heard a sentence from a speaker who had directly experienced a past event and hence used –DI; responding correctly thus involved converting a –DI sentence into one that contained –mIş instead. We used two orders of presentation of the 8 stories (one the reverse of the other) across participants.

**Results**
The proportion of children’s correct responses in each condition is presented in Figures 1a and 1b. Children only employed the past tense morphemes –DI and –mIş in their answers and no other morpheme. Hence, whenever children made a mistake, they replaced the correct evidential morpheme with the incorrect one (i.e., they used the direct marker instead of the indirect one or vice versa). Beginning with the See vs. Hear condition, an ANOVA with the proportion of correct responses as the dependent variable and Age (5, 6, 7) and Source (See, Hear) as factors revealed no main effect of Age (F (2, 45) = 2.116, partial $\eta^2 = .04, p = .132$), a main effect of Source (F (1, 45) = 44.36, partial $\eta^2 = .33, p = .000; M_{\text{See}} = 95, M_{\text{Hear}} = 60$), and an interaction between Age and Source (F (2, 45) = 4.24, partial $\eta^2 = .131, p = .04$). This interaction was due to the fact that performance on the See items did not differ across ages (5- vs. 6-year-olds, p = .11; 5- vs. 7-year-olds, p = .823; 6- vs. 7-year-olds, p = .064) while performance on the Hear items differed between the youngest and the oldest group (t(30) = -2.67, p = .014; $M_5 = 48.44, M_7 = 70.3$).

A similar ANOVA for the See vs. Infer condition returned only a main effect of Source (F (1, 45) = 56.13, partial $\eta^2 = .39, p = .000; M_{\text{See}} = 98, M_{\text{Infer}} = 52$) but no main effect of Age (F(2,45)= 2.130, partial $\eta^2 = .03, p=.131$) and no interaction between Source and Age (F(2,45)= 2.365, partial $\eta^2 = .07, p=.106$).

Figures 1a and 1b. Percentage of correct responses in Exp.1 (See vs. Hear and See vs. Infer condi-
tion). Error bars depict standard error.

**Discussion**

The results of the production experiment show that, when Turkish learners had direct access to information, they used the morpheme –DI most of the time. However, when they had indirect access to information, they did not consistently use the morpheme –mIs across ages. Specifically, 5-year-olds used both –DI and –mIs for information acquired through either hearsay or inference; 6- and 7-year-olds used the indirect evidential –mIs reliably for hearsay but not inference. Children’s responses always included either –DI or –mIs (and no other morpheme), a fact indicating that children acquire the past tense meaning of these two morphemes from distributional evidence before they fully learn their evidential dimension (see also Aksu-Koç, 1988; Ozturk & Papafragou, 2007, for similar observations). In sum, even though there appears to be some sensitivity to evidential meanings even in the group of 5-year-olds, these findings are broadly consistent with previous elicited production studies in Turkish (e.g., Aksu-Koç, 1988) that have pointed out persistent difficulties especially with the indirect evidential.

**Experiment 2: Comprehension of Evidential Morphology**
This experiment was conducted to see if children can attribute a sentence with an evidential morpheme to a speaker that has appropriate access to information. Given the findings from Experiment 1, Turkish-speaking children might show better evidence of comprehension of evidentials compared to their production levels.

**Stimuli and Procedure**

Stimuli for each condition consisted of 8 short stories, each developing over two videos presented sequentially. In the See vs. Hear condition, in the first video of each story, a person acted on some object(s) (e.g., a girl chose one doll from a set and played with it) while someone else watched. In the second video, a third character was shown alone in the same room looking at the objects (e.g., the dolls); when a new person came in, the character pointed to the objects and whispered something into this person’s ear and the person nodded. Participants were informed that “one character is now telling the other one what happened before.” Next pictures of the person who watched what happened and the person who was verbally informed appeared on the screen. The experimenter said that one of the people on the screen uttered a sentence. Sentences were pre-recorded using the voice of a native Turkish speaker. Half of the time the sentence used the direct evidence morpheme –DI (‘See’ trials) and the other half the indirect evidence morpheme –mIş (‘Hear’ trials):

(6) Kiz  bebek   oyna-di / –mIş.
   Girl   doll       play-past evid.

   ‘The girl played with a doll’ (direct / indirect)

The experimenter then asked participants: “Who said that?” Participants were expected to match
the sentence containing the evidential morpheme –DI with the person that had witnessed the event and the sentence containing the indirect evidence morpheme –mIş with the person that had heard what happened.

In the first video of each story for the See vs. Infer condition, one person again acted on some object(s) (e.g., a girl drank lemonade from a bottle) while someone else watched. In the second video, a new person came in the same room, found some evidence indicating what might have happened (e.g., the half-empty lemonade bottle on the table) and spent some time inspecting the evidence. No other characters were present. Next, pictures of the person who watched what happened and the person who saw a clue about what happened appeared on the screen. The experimenter said that one of the two people uttered a sentence. Half of the time the sentence contained the direct evidence morpheme –DI (‘See’ trials) and the other half the indirect evidence morpheme –mIş (‘Infer’ trials):

(7) Kız limonata ic-di / -mış.

Girl lemonade drink-past evid.

‘The girl drank the lemonade’ (direct / indirect)

The experimenter then asked participants: “Who said that?” It was expected that, if participants understood the difference in the source meanings associated with the two morphemes, they would pick the person that saw what happened when the sentence included the direct evidential morpheme –DI and the person that inferred what happened when the sentence included the indirect morpheme –mIş.

Within each between-subjects condition, the 8 stories were presented in two orders, one of which was the reverse of the other. Half of the participants saw one order and the other half the re-
verse order. The same events were used for the indirect (report or hearsay) access.

**Results**

An ANOVA with the proportion of correct responses as the dependent variable and Age and Source as factors was conducted separately for each between-subjects condition. Results are presented in Figures 2a and 2b. For the See vs. Hear condition, the analysis revealed a main effect of Age (F (2, 45) = 8.11, partial $\eta^2 = .17$, p = .001), a main effect of Source (F (1, 45) = 5.72, partial $\eta^2 = .06$, p=.02, MSee = 66.6, M Hear = 56.2) and a significant interaction between Age and Source (F (2, 45) = 3.49, partial $\eta^2 = .08$, p = .04, MSee Age 5 = 67.19, MSee Age 6 =54.69, MSee Age 7 = 78.12; M Hear Age 5 = 42.19 , M Hear Age 6 = 57.81, M Hear Age 7 = 68.75). Pairwise t-tests revealed that the interaction was due to the fact that children's performance on the Hear trials was significantly different between the ages of 5 and 6 ($t(30) = -2.698$, p = .03) but on the See trials it was not (p = .07). Additionally, children's performance on the See trials was significantly different between the ages of 6 and 7 ($t(30) = -2.838$, p = .01) whereas their performance on the Hear trials was not. Also, children's performance on the See trials was significantly different between the age groups 5 and 7 ($t(30) = -2.838$, p = .01) while their performance on the Hear trials was not. Children's performance on the See trials was significantly different from chance at ages 5 ($t(15) = 2.711$, p = .02) and 7 ($t(15) = 5.582$, p = .001) but not at the age of 6 ($t(15) = .716$, p = .49). Children's performance on the Hear items was significantly different from chance only for the oldest age group ($t(15) = 3.873$, p = .002).

For the See vs. Infer condition, the analysis revealed only a main effect of Age (F(2, 45) = 4.60, partial $\eta^2 = .08$, p = .02) but no significant effect of Source (F (1, 45) = 1.44, partial $\eta^2 = .02$, p=.24), and a marginal interaction between Age and Source (F (2, 45) = 3.09, partial $\eta^2 = .08$, p
One-sample t-tests revealed that children’s performance on the See trials was significantly different from chance at the age of 5 ($t(15) = 2.71$, $p = .03$) and 7 ($t(15) = 5.58$, $p = .00$) but not at the age of 6 ($t(15) = .71$, $p = .48$). Children’s performance on the Infer trials was significantly different from chance only for the older two groups ($t_{6}(15) = 2.61$, $p = .03$; $t_{7}(15) = 3.10$, $p = .00$) but not for the youngest children ($t_{5}(15) = -.69$, $p = .49$).

Figures 2a and 2b. Percentage of correct responses in Exp.2 (See vs. Hear and See vs. Infer condition). Error bars depict standard error.

**Discussion**

Our comprehension data show that young learners of Turkish appeared to have some success with the comprehension of the direct marker $–DI$ but only children in the oldest age group understood the indirect morpheme $–mIş$ and used evidential morphology to make consistently correct inferences about source knowledge. These findings are consistent with comprehension data in Aksu-Koç (1988) and confirm the conclusion (already supported by the production data) that the acquisition of evidentiality in Turkish is quite protracted.  

6 Inspection of the data in Figures 1 and 2 suggests that Turkish children’s productive command of evidentiality is better compared to their comprehension of evidentials. This was confirmed in an ANOVA comparing accuracy in Experiments 1 and 2: the analysis revealed an effect of Experiment ($F(1, 93) = 16.93$, partial $\eta^2 = .03$, $p=.00$), an effect of Age ($F(2, 93) = 4.66$, partial $\eta^2 = .01$, $p = .01$), and no significant interaction between Experiment and Age ($F(2, 93) = 1.036$, partial $\eta^2 = .002$, $p = .36$). The production-comprehension asymmetry runs contrary to the general observation that produc-
The current comprehension task involves tracking multiple characters over two segments of a story, comparing two of these characters’ access to events and attributing an evidentially marked sentence to one of them. Therefore it is an open possibility that a simpler task might reveal higher sensitivity to the meaning of evidential morphology in Turkish learners. We address this possibility in the next experiment.

**Experiment 3: Evaluating Others’ Reliability from Use of Evidentiality**

The primary goal of this experiment was to investigate whether children are aware of the discourse functions of the evidential markers – specifically, whether children know that a speaker who employed the direct evidence morpheme -DI should be trusted over a speaker that employed the indirect evidential morpheme –mIş (other things being equal; see Fitneva, 2008; Willett, 1988). This task involved comparing two conflicting evidential sentences to each other so as to decide which speaker to believe. A similar paradigm has been used to demonstrate that 5-year-olds distinguish between must and may when they have to find a hidden object and two speakers offer conflicting modal statements about the location of the object (Hirst & Weil, 1982; Noveck, Ho & Sera, 1996; Ozturk & Papafragou, 2012). This simpler design allowed us to address the concern over task complexity raised at the end of Experiment 2.

A secondary goal of the experiment was to explore whether any difficulties children might face with the discourse functions of evidential morphemes are due to the demands of reasoning about speaker reliability or the specific properties of the linguistic devices used to convey reliability. We therefore compared evidential morphemes to full lexical verbs with related meanings.
Stimuli and Procedure

Materials consisted of 8 animated stories, each involving one box and two animals. The experimenter informed participants that they were going to play a game to find out what was in a box. Participants were told that all of the boxes were going to be opened at the end of the game to see whether or not participants were right in their choices. 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 contents of the box. In half of the stories, the pair of statements included a direct and an indirect evidential morpheme:

(8) Bu kutu-da bir uçak var-dı.
This box-loc. an airplane to.be-evidential
Intended reading: ‘I saw that there is an airplane in this box.’

(9) Bu kutu-da bir helikopter var-mış.
This box-loc. a helicopter to.be-evidential
Intended reading: ‘I heard that there is a helicopter in this box.’

In the other half of the stories, the pair of statements included the full verb counterparts of evidential morphemes, i.e., the verbs görmek ‘to see’ and duymak ‘to hear’ - we did not include an inference verb (see Lucas et al., 2013, on Turkish learners’ ability to handle similar embedded structures):
(10) Bu kutu-da bir uçak ol-duğ-u-n-u gör-dü-m.

This box–loc. an airplane be-Nomin.-Poss.3sg.-Acc. see-past.-1sg.

‘I saw that there is an airplane in this box.’

(11) Bu kutu-da bir helikopter ol-duğ-u-n-u duy-du-m.

This box–loc. a helicopter be-Nomin.-Poss.3sg.-Acc. hear-past.-1sg.

‘I heard that there is a helicopter in this box.’

The participants’ task was to say what was in the box. We reasoned that the animal using the morpheme –DI (or the verb ‘see’) should be trusted more than the animal using the morpheme –mIş (or the verb ‘hear’).

Stories with a specific type of expression (Verb vs. Morpheme) were presented in blocks with block order counterbalanced across participants. Within each block, there were two orders of presentation, one the reverse of the other. The left-right position of the animals producing the correct answer was counterbalanced within participants.

**Results**

Results from Experiment 3 are shown in Figure 3. There were no block order effects so we collapse across orders in what follows. An ANOVA with the proportion of correct responses as the dependent variable and Age and Expression (Verb, Morpheme) as factors revealed only a main effect of Age (F (2, 93) = 10.93, partial η² = .47, p = .00, M5 = 57.8, M6 = 65.6, M7 = 90.6) but no effect of Expression (F(1, 93) = 2.32, partial η² = .03, p = .13) and no interaction of Expression by Age.
(F (1, 93) = 2.32, partial $\eta^2 = .22, p = .36$). A series of t-tests revealed that children's performance differed significantly between the ages of 5 and 7 ($t(62) = -8.06, p = .00$) and 6 and 7 ($t(62) = 34.49, p = .00$) but not between the ages of 5 and 6 ($t(62) = -1.801, p = .08$).

Further analyses revealed that the youngest children's performance was not significantly different from chance for either type of expression (Morpheme: $t(31) = 1.06, p = .3$; Verb: $t(31) = .70, p = .49$). Six-year-olds' performance was significantly different from chance for morphemes ($t(31) = 2.70, p = .01$) but not verbs ($p = .3$). Seven-year-olds’ performance was significantly different from chance for both types of expression (Morpheme: $t(31) = 10.06, p = .00$; Verb: $t(31) = 6.31, p = .00$).

Figure 3. Percentage of correct responses in Exp. 3 (Morpheme and Verb conditions). Error bars depict standard error.

**Discussion**

The results of this experiment showed that Turkish children’s pragmatic understanding of the effects associated with the use of evidential markers begin around the age of 6 and their performance improves dramatically at the age of 7. Reliability effects from Turkish source verbs such as *see* and
hear are also acquired around the same age. This finding fits with results from the two previous experiments according to which Turkish children’s production and comprehension of evidential morphology becomes stable around the age of 6 or 7. Furthermore, this task shows that children’s difficulties persist even if one removes the specific demands of our comprehension task (Exp.2).

The present findings are consistent with Papafragou et al. (2007) who found difficulties with the pragmatic functions of Korean evidentials in 3- and 4-year-olds using a very similar task (cf. also Fiteeva, 2008, on Bulgarian; Matsui et al., 2006, on Japanese; and Aksu-Koç & Alici, 2000, on Turkish for similar patterns). Notice that these results remove the possibility that Turkish learners’ early uses of evidential morphology encode relative speaker certainty as opposed to source meanings: if that were the case, even 5-year-old children should perform very well on this task (e.g., picking the direct over the indirect morpheme for semantic, as opposed to pragmatic, reasons).

Source Monitoring Experiments

Experiment 4: Reporting One’s Own Sources

This experiment asked whether children were able to report their own sources of information.

Stimuli and Procedure

Materials consisted of 8 animated stories (see also Exp.1). In the See vs. Hear condition, participants either saw something happen on the screen (e.g., a frog jump along a river) or heard a character on the screen utter a sentence (e.g., a woman say: “I rode a bike.”). In the See vs. Infer condition...
tion, participants again either saw something happen on the screen or saw some hints (e.g., a slice next to a birthday cake) and were expected to infer what happened (here, that somebody cut the cake). After each trial, the experimenter asked participants to report what happened (notice that the question did not include an evidential morpheme):

(12) Ne ol-duğ-un-u biyl-iyor mu-sun?

What happen-Nom.-3.sg-Acc. know-Pr.Prog. Q.-2.sg

‘Do you know what happened?’

After participants answered, the experimenter asked them how they had found out:

(13) Ner-den biyl-iyor-sun?

Where-Abl. know-Pr.Prog-2sg

‘How do you know?’

(14) **See vs. Hear condition**

Gör-du-n-mü? Duy-du-n-mu?

See-Past-2sg.-Q. hear-Past-2sg.-Q.

‘Did you see? Did you hear?’

(15) **See vs. Infer condition**

Gör-du-n-mü? Ya da bir sey-ler görüp mü anla-di-n?8

---

8 Since children would probably not understand the verb “infer” in Turkish, we used a complex phrase including the verb anlamak ‘understand’. We decided to embed the verb in a phrase because our prior pilot work (Ozturk & Papafragou, 2007) showed that stand-alone use of the verb was not helpful to children.
See-Past-2sg.-Q. Or one thing-Pl. see-Q. understand-Past-2sg.

‘Did you see? Or did you see something and infer (what happened)?’

In each between-subjects condition there were 4 See trials and 4 trials of either the Hear or the Infer type. Trials were administered in two orders (one the reverse of the other) across participants.

**Results**

For each between-subjects condition, an ANOVA was conducted with the proportion of correct responses as the dependent variable and Age and Source as factors (see Figures 4a and 4b). For the See vs. Hear group, the ANOVA yielded a main effect of Source ($F (1, 45) = 156.35$, partial $\eta^2 = .78$, $p = .00$), with See trials eliciting much higher accuracy compared to Hear trials ($M_{\text{see}} = 94.5$, $M_{\text{hear}} = 39.58$). However, the ANOVA yielded no main effect of Age ($F (2, 45) = 1.78$, partial $\eta^2 = .07$, $p = .18$) and no interaction between Age and Source ($F (2, 45) = .85$, partial $\eta^2 = .04$, $p = .44$). For the See vs. Infer group, a similar ANOVA again yielded a main effect of Source ($F (1, 45) = 66.93$, partial $\eta^2 = .60$, $p = .00$), with See trials eliciting higher accuracy compared to Infer trials ($M_{\text{see}} = 97.91$, $M_{\text{infer}} = 61.98$). There was no main effect of Age ($F (2, 45) = .23$, partial $\eta^2 = .09$, $p = .12$) and no interaction between Age and Source ($F (2, 45) = .38$, partial $\eta^2 = .02$, $p = .69$). In all cases in which children made an error in this task, they incorrectly chose the alternative in the forced-choice question. This meant that all errors in the Hear and Infer trials were cases in which children reported having seen what they had only heard or inferred.

Children’s performance for the See trials was consistently different from chance in all age groups (all $p$s < .05). In the Hear trials, performance was significantly different from chance in the two older age groups ($t_5(15) = -.96.50$, $p = .35$; $t_6(15) = -3.22$, $p = .01$; $t_7(15) = -.64$, $p = .05$) – but in
6-year-olds this was because children were systematically incorrect. In the Infer trials, only the oldest children succeeded above chance \( t_5(15) = .46, p = .65; t_6(15) = 1.59, p = .13; t_7(15) = 2.82, p = .01 \).

Figures 4a and 4b. Percentage of correct responses in Exp. 4 (See vs. Hear and See vs. Infer condition). Error bars depict standard error.

**Discussion**

The results of this experiment show that Turkish children are able to successfully report their own sources of information for the direct perception (See) items; performance on such items was at ceiling even in the youngest age group. Performance on the indirect items was lower and not consistently different from chance until age 6 (Infer) or later (Hear). This finding is in line with previous literature claiming that the ability to make reference to visual perception as a source of one’s own knowledge appears early unlike the ability to refer to indirect sources such as inference (Gopnik & Graf, 1988; O’Neill & Astington, 1990; O’Neill & Chong, 2001; O’Neill & Gopnik, 1991; Woolley & Bruell, 1996; Wimmer et al., 1988; Sodian, & Wimmer, 1987; Aksu-Koç et al., 2009; among others).

Despite this broad similarity, our data differ from prior literature, according to which the ability
to explicitly refer to indirect sources is present in children around the age of 5 or 6. Notice that, even though our task was loosely based on earlier studies (e.g., Gopnik & Graf, 1988; O’Neill & Gopnik, 1991), it departed in several notable ways from prior work. First, children in our experiment were asked to report on how they found out about an event, not on the identity of a hidden object. Second, the act of gaining access to information was deemphasized: unlike prior studies in which children gained access to the contents of an opaque container (such as a tunnel), and later had to report how they found out about the contents, here children simply watched short clips and gained access to events in different ways. Relatedly, unlike prior studies, there was no training phase directing the children’s attention to different information sources and giving children practice with the language to refer to them (“Did you see? Did I tell you? Did you figure it out from a clue?”)\(^9\). In both these respects, our study was closer to the everyday situations in which children might have to reconstruct the sources of their memories. Third, and relatedly, the way of granting access to information was not as sharply differentiated in our stimuli as in prior work: since children always watched a clip, there was an element of seeing even in the Hear and Infer trials (much like in many real-life instances where knowledge is acquired). It is known that the more similar the contents of experiences that lead to knowledge formation, the less accurate children are in separating out the sources of these experiences (Lindsay, Johnson & Kwon, 1991). Finally, the types of hearsay and inference we used were different from what has been studied in the literature. In our Hear trials, children heard a verbal report from a character in a clip (in previous studies verbal information came from the experimenter – a potentially more salient source). Furthermore, in our Infer trials, children saw the endpoint of an event and had to inferentially reconstruct the event. In previous studies such as O’Neill and Gopnik’s (1991), children were presented with a thematically

\(^9\) Children’s low success with the Hear items may be at least partly due to the choice of verb for reporting information acquired via conversation (dinlemek ‘listen, make a conscious effort to hear’; see Ozturk & Papafragou, 2007, for higher-accuracy data using a different verb in an earlier version of this task).
related cue such as an egg carton and were explicitly told that the hidden object “belonged in there” (in this context, the presence of an overt cue makes tracking and reporting inference easier). In all these respects, the present task (especially the Hear and Infer portion) is more demanding than previous source monitoring tasks that served as its inspiration (see also Introduction, and Gopnik & Astington, 1991).

**Experiment 5: Reporting Others’ Sources**

This experiment asked whether children were able to report other people’s sources of information.

**Stimuli and Procedure**

Materials consisted of 8 stories, each developing over 2 videos presented sequentially (see also Exp.2). In the See vs. Hear condition, in the first video of each story, a person acted on some object(s) (e.g., a boy took off his jacket) while someone else watched. In the second video, a third character was shown alone in the same room looking at the object(s) (e.g., the jacket on the table); when a new person came in, the character pointed to the object(s) and whispered something into this person’s ear and the person nodded. Participants were informed that “one character is now telling the other one what happened before.” At the end of each story, participants were presented with pictures of the character that witnessed the event and the character who was verbally informed about it. Participants were next asked the questions in (16) and (17) and had to match the right character with the right action. Half of the participants in this condition were first asked (16) followed by (17) and the other half was asked (17) first followed by (16):
(16) Ne ol-duğ-u-(n)-u kim gör-dü?

What happen-Nom.-Poss.-Acc. who see-past

‘Who saw what happened?’

(17) Ne ol-duğ-u-(n)-u kim duy-du?

What happen-Nom.-Poss.-Acc. who hear-past

‘Who heard (what happened)?’

In the See vs. Infer condition, in the first video of each story, one person again acted on some object(s) (e.g., a boy peeled an orange) while someone else watched. In the second video, a new person came in the same room, found some evidence indicating what might have happened (e.g., the skin of the orange on a table) and spent some time inspecting the evidence. After each story, the participants were presented with pictures of the character who witnessed the event and the character who had been given clues about it. Participants were next asked the questions in (16) and (18) (order counterbalanced across participants) and had to match the right action with the right character:

(18) Kim bir şey-ler görüp anla-di?

Who one thing-Pl. see understand-past

‘Who saw something and inferred (what happened)?’

Since success on the first test question predicted success on the second one (no child picked the same character twice as the person who both saw and heard/inferred), we only focus on answers to the first question. Therefore See trials are defined as trials in which the ‘Who saw’ question in (16)
was presented first and Hear or Infer trials as trials in which the ‘Who heard/inferred’ question in (17)/(18) was presented first. In each between-subjects condition there were 4 See and 4 Hear or Infer trials. The trials were administered in two orders, one of which was the reverse of the other. The same events were used for the indirect (report or hearsay) access.

**Results**

Results from Experiment 5 are shown in Figures 5a and 5b. For the See vs. Hear condition, an ANOVA with the proportion of correct responses as the dependent variable and Age and Source as factors revealed only a main effect of Source (F(1, 45) = 29.53, partial $\eta^2 = .08$, p = .00), with See trials eliciting higher accuracy rates (M = 88.54) than Hear trials (M = 63.54) but no main effect of Age (F(2, 45) = 1.97, partial $\eta^2 = .08$, p = .15), and no interaction of Age by Source (F(2, 45) = 1.50, partial $\eta^2 = .06$, p = .23). For the See vs. Infer condition, a similar ANOVA again yielded only a main effect of Source (F (1, 45) = 56.13, partial $\eta^2 = .56$, p = .001), with See trials eliciting higher accuracy (M = 90.1) compared to Infer trials (M = 51.04) but no main effect of Age (F(2, 45) = 2.13, partial $\eta^2 = .09$, p = .13), and no interaction of Age by Source (F(2, 45) = 2.37, partial $\eta^2 = .10$, p = .11).

A closer look at each age group’s performance within the two conditions using one-sample t-tests revealed that performance of all age groups on the See trials in both the See vs. Hear condition and in See vs. Infer condition was significantly different from chance (all ps < .05). Children’s performance on the Hear trials was significantly different from chance only for the two oldest age groups ($t_5(15)=1.10$, p = .29, $t_6(15)=2.24$, p = .04, $t_7(15)=5.98$, p = .00). Finally, children’s performance on the Infer trials was not significantly different from chance for any of the age groups ($t_5(15)=1.24$, p = .24, $t_6(15)=.82$, p = .42, $t_7(15)=1.70$, p = .11).
Figures 5a and 5b. Percentage of correct responses in Exp.5 (See vs. Hear and See vs. Infer condition). Error bars depict standard error.

**Discussion**

This experiment asked whether Turkish-speaking children could track how two different characters found out about the same event. Results show that children were able to successfully report whether someone else had seen an event already at age 5. Children’s ability to report that someone had heard about an event showed developmental change, and performance became significantly different from chance at age 6. The ability to report that someone had figured out something developed later: even 7-year-olds were not reliably successful at this task. Finally, our data reveal an asymmetry between direct and indirect sources that was also observed in Experiment 3. We return to the significance of this pattern in the General Discussion below.

**Experiment 6: Evaluating Others’ Reliability from Their Access to Information**
The goal of this experiment was to test whether children could recognize that an agent with direct evidence for an event generally knows more about the event compared to an agent with only indirect evidence about the event.

**Stimuli and Procedure**

Materials consisted of 4 stories per condition, each developing over two videos presented sequentially. The videos were presented on a single screen in windows next to each other. In the See vs. Hear condition, in the first video of each story, a character performed an action (e.g., a girl drew a happy face on a piece of paper) while another character watched. In the second video, a third person (without clear access to the event) reported to a fourth person what had happened. After each video played, the last frame froze on the screen. At the end of each story, two arrows appeared pointing to the character who witnessed the event in the first video and to the character who heard about it in the second video. Participants were then asked:

(19) Ne ol-duğ-unu kim daha iyi bil-iyor? Kim-e sor-a-lim?

Who happen-Nom who more good know-Pres. Prog who-Dat ask-Imp.-1st.pl.

‘Who knows better what happened? Who should we ask?’

In the See vs. Infer condition, in the first video of each story, one character performed an action (e.g., a girl took books out of a bag) while another character watched. In the second video, a third person came into the same room and saw some hints (e.g., the books and the empty bag) indicating what had happened. At the end of each story, after both of the videos had played and the arrows appeared, participants were asked the question in (19). As before, we expected children to
attribute better knowledge to the character who had witnessed the event. Stories were administered in two orders, one of which was the reverse of the other. The same events were used for indirect (hearsay or inferential) access.

**Results**

Results are shown in Figure 6. Children overall performed at ceiling on this experiment. An ANOVA with percentage correct as the dependent variable and Age and Condition as factors revealed no main or interaction effects (Age: F(2,93)= 2.07, partial $\eta^2 = .12$, p=.13; Source: F(1,95)=.00, partial $\eta^2 = .12$, p= 1.0; Age by Source Interaction: F(2,90)= .00, p=1.00).

![Figure 6. Percentage of correct responses in Exp.6 (See vs. Hear and See vs. Infer conditions). Error bars depict standard error.](image)

**Discussion**
The results show that children as early as 5 years can choose the more reliable of two sources of information (seeing over hearing or inferring) when it comes to gaining access to specific events in the world. Prior tasks showing that young children flexibly appeal to source information (e.g., by adopting others’ suggestions over their own judgments if others had more informative perceptual access; Robinson & Whitcombe, 2003) have used implicit tasks in which children were not explicitly asked questions about a cognitive agent’s mental state. Here we show that, even on an explicit task that asks children to say “who knows better”, children as young as 5 were able to choose the right character based on that character’s access to information.

Could it be that children equate seeing and knowing and take the question “Who knows better?” to mean “Who saw (what happened) better?” On this alternative interpretation of our findings, seeing would still have a special status in children's thinking and would be differentiated from other types of informational access, but in a more superficial way. We find this possibility unlikely: Several studies (at least in English) show good understanding of the verb “know” at age 5, including the ability to differentiate “know” from “think” (Moore et al., 1989). Other work has shown that even 3-year-olds can correctly pick the character who “knows” from among two characters, one of which lifts a box and another one who looks inside a box; furthermore, 3-year-olds appear to behave identically whether the question contains “know” or “could tell”, a fact that suggests that even 3-year-olds do not misinterpret the verb “know” as “see” (Pratt & Bryant, 1990).

Two further observations are in order about the finding that children prioritize visual perception over both hearsay and inference. First, our study leaves open the question whether children actually attribute some knowledge to the character who had hearsay/inferential access to information, or treat hearsay and inference as completely uninformative. Second, the privileged status of perceptual access in our data may not reflect a general property of source monitoring but might be specific to processes that lead to episodic or event-related information. This idea is consistent with
evidence that young children ignore reliability information and exclusively rely on an informant’s relevant perceptual access when learning episodic facts such as which of two objects is hidden in a given location (Brosseau-Liard & Birch, 2011). And at least older children do not always prioritize visual access over other knowledge sources: when learning about external, visible characteristics of unfamiliar animals and individuals, 6-year-olds rely on visual inspection to acquire information - but, when learning about internal, invisible qualities, they choose to direct questions to individuals with relevant established expertise (Fitneva, Lam, & Dunfield, 2013).

**Comparison of Linguistic Evidentiality and Source Monitoring Experiments**

A combined score per participant for all the Linguistic Evidentiality and all the Source Monitoring experiments was calculated by averaging children’s scores in Experiments 1-3 and 4-6. A mixed-design ANOVA with the percentage of correct responses of each participant as the dependent variable, Experiment Type (Linguistic Evidentiality, Source Monitoring) as a within-subjects variable and Age (5, 6, 7) as a between-subjects variable was conducted. Results are shown in Figure 7. The analysis revealed a significant main effect of Experiment Type (F(1,93) = 352.01, partial $\eta^2 = .79$, p = .00): overall, children performed better in the Source Monitoring experiments than they did in the Linguistic Evidentiality experiments ($M_{SMonitoring} = 81.67$, $M_{Ling} = 68.71$). Moreover, the analysis revealed a significant main effect of Age (F(2,93) = 6.13, p = .00, partial $\eta^2 = .12$, $M_5 = 71.67$, $M_6 = 74.31$, $M_7 = 79.58$). Additionally, the analysis revealed a significant interaction between Experiment Type and Age (F(2,93) = 6.51, partial $\eta^2 = .12$, p = .00): in the Linguistic Evidentiality tasks, the performance of the oldest children was significantly different from the performance of the youngest children (t(62) = -2.61, p=.01) as well as the middle group (t(62) = -4.26, p = .00). However, in the Source Monitoring tasks, performance did not differ significantly between chil-
dren of different age groups (all p’s > .05). In sum, children performed better on the Source Monitoring experiments than they did on the Linguistic Evidentiality ones. Furthermore, performance on the Evidentiality experiments showed developmental change while performance on the Source Monitoring experiments seemed overall stable between ages 5 and 7.

We repeated the same type of analysis focusing on more detailed comparisons between experiments that were theoretically and methodologically paired with each other. The first of these analyses compared children production of evidentials (Exp. 1) to their ability to report their own information sources (Exp. 4). A repeated-measures ANOVA with Experiment and Age as factors revealed a main effect of Experiment in the predicted direction with the Production task being worse than the Source Monitoring task (F(1, 93) = 26.113, p = .00, MProd = 62.65, MSelf-Report = 73.57). The analysis revealed no main effect of Age and no interaction between Experiment and Age. A second analysis compared children’s comprehension of evidentials in the speech of others (Exp. 2) to the very same children’s ability to recognize others’ sources of information (Exp. 5). The analysis revealed a main effect of Experiment (F(1, 93) = 23.754, partial $\eta^2 = .12$, p = .00), no main effect of Age (F(1, 93) = 23.754, partial $\eta^2 = .12$, p = .00), and no interaction between Experiment and Age (F(1, 93) = 23.754, partial $\eta^2 = .12$, p = .00): performance on the Comprehension of Evidentiality task lagged behind performance on the Source Attribution task (MComp = 62.65, MOthers-Report = 73.31). A third analysis compared children’s ability to evaluate an agent’s reliability on linguistic grounds (i.e., evidential use, Exp. 3) vs. source-monitoring grounds (i.e., access to information, Exp. 6). The analysis returned a main effect of Experiment (F(1, 93) = 150.06, partial $\eta^2 = .12$, p = .00), a main effect of Age (F(2, 93) = 29.23, partial $\eta^2 = .12$, p = .00), and no interaction between Experiment and Age (F(2, 93) = 29.23, partial $\eta^2 = .12$, p = .00): crucially, there was again an advantage of the non-linguistic/source monitoring over the linguistic evidentiality score (MNonLing-Reliability = 97.92, MLing-Reliability = 71.29).
Taken together, these findings show that the conceptual hypothesis cannot explain the delay in the acquisition of linguistic evidentiality: existing source concepts have not been mapped successfully onto the linguistic-evidential morphemes. These findings cohere with the position that conceptual development precedes and establishes a basis for subsequent linguistic development. The close relationship between source monitoring and linguistic performance was confirmed by the fact that there is a positive correlation between children’s overall linguistic and source monitoring score (Pearson's r = .37, p = .01).

Figure 7. Percentage of correct responses for Linguistic Evidentiality and Source Monitoring Tasks. Error bars depict standard error.

**General Discussion**

A major goal of this paper was to explore the link between the acquisition of linguistic evidentiality and the development of source reasoning abilities in children as a way of probing the relations be-
between linguistic and conceptual development. We began with the observation that learners face difficulties with the acquisition of evidential morphology cross-linguistically and considered two hypotheses that can be used to explain this observation. According to the first of these hypotheses, the acquisition of evidential morphology is complicated by the abstractness and difficulty of the underlying concepts. According to an alternative hypothesis, the acquisition of evidentiality is delayed because of the complexities of mapping existing source concepts onto the corresponding morphemes. We focused on Turkish, a language with grammaticalized evidentiality, and attempted to evaluate these two hypotheses by systematically comparing linguistic and source monitoring performance on evidential tasks by the same group of Turkish-speaking children.

One set of experiments tested 5-, 6- and 7-year-old-children’s production (Exp.1) and comprehension (Exp.2) of evidential morphology, as well as their pragmatic ability to link linguistic evidentials to speaker reliability (Exp.3). An additional set of three experiments measured the same children’s source monitoring abilities, specifically, the ability to report their own (Exp.4) and others’ (Exp.5) information sources, as well as the ability to relate information source to reliable knowledge (Exp.6). Below we summarize the main findings of the studies in terms of the acquisition of linguistic evidentiality, as well as the contributions of mapping and conceptual/source monitoring factors to the acquisition of evidentiality.

The acquisition of evidentiality in Turkish

A major finding from this set of studies is that the acquisition of the Turkish evidential system develops along a protracted timetable. Specifically, as Experiments 1-3 show, 5-year-old Turkish-speaking children produce and comprehend the morpheme -DI in contexts that involve visual perception; nevertheless, the morpheme is overextended to cases where no direct evidence is involved.
Furthermore, it is not until the age of 6 that children reliably produce the morpheme \(-ml\) in contexts that involve hearsay; production of \(-ml\) in contexts that involve inference is not reliable even in 7-year-olds. Comprehension data show that the evidential morphology cannot be used as a cue to information access until after the age of 6 or 7. Pragmatic aspects of evidential morphology are similarly delayed: the ability to infer that a speaker who used the indirect evidential had access to less reliable knowledge compared to a speaker who used the direct evidential emerges at the age of 6. These data are broadly consistent with prior experimental demonstrations on Turkish that have concluded that the abilities to modify one’s speech in accordance with informational status are “developmental phenomena which approach adult language standards only gradually” (Aksu-Koç, 1988, p.194; cf. Aksu-Koç & Alici, 2000; Ögel, 2007; Ozturk & Papafragou, 2007; Aksu-Koç et al., 2009). Furthermore, these results contribute to a growing literature demonstrating challenges in the acquisition of evidential systems across different languages (de Villiers et al., 2007; Matsui et al., 2006; Papafragou et al., 2007; inter alia).

On the basis of these results, it appears that young Turkish learners use \(-DI\) and \(-ml\) in distributionally appropriate ways to mark past tense before they fully discover the evidential functions of these morphemes. The evidential functions of the two morphemes are fully discovered after the age of 6 or 7 – even though our data do not resolve the question of when knowledge of the Turkish evidential system becomes completely adult-like.

**Prerequisites for acquiring linguistic evidentiality**

A second major finding from the present experiments is that Turkish learners’ source monitoring abilities are better than their linguistic knowledge of evidentiality. Turkish-speaking children who
have difficulty with evidential morphology are better able to report knowledge sources for themselves and to link sources such as perception, communication and inference to knowledge in others in tasks which do not involve evidential morphology. Most strikingly, children who cannot link the use of the direct evidential with higher speaker reliability compared to the use of the indirect evidential can recognize that perceptual access is generally a more reliable information source than either hearsay or inference (other things being equal).

These asymmetries point to an important conclusion: the development of source monitoring abilities emerges prior to aspects of linguistic systems encoding source information. Our results are consistent with a rather classical view on the relationship between language acquisition and conceptual development, according to which language acquisition presupposes (and builds on) the development of the relevant concepts in the child. Our findings are markedly at odds with more recent proposals according to which conceptual development might be driven by exposure to lexical and grammatical learning – in other words, conceptual categories could be constructed during the process of acquiring one’s native tongue (see Bowerman & Levinson, 2001; Gentner & Boroditsky, 2001; cf. Lucas et al., 2013).

Our data further suggest that a critical part of the difficulty with the acquisition of Turkish evidentials – at least for 5-7-year-old children – seems to be discovering the correspondence between evidential morphology and antecedently available source concepts. This makes sense considering that few clues in the extra-linguistic environment point to source meanings: even in the cases of direct perception, where an act of observing an event has occurred, the learner needs to link a grammatical morpheme to a propositional attitude (‘direct access/perception’) which itself is unobservable. Beyond this general difficulty, we hypothesize that there are several, more specific factors that complicate the mapping problem for Turkish learners. First, the semantics of Turkish past-tense evidentials (direct vs. indirect) does not map in a one-to-one fashion onto individual types of
evidence. For instance, the ‘indirect’ marker can be used for both hearsay and inference. Furthermore, even though inference from observables includes a perceptual component (e.g., one can infer the presence of an animal by seeing its footsteps), inference, unlike perception, is considered indirect access in the Turkish past-tense system (cf. our Infer trials above). Somewhat similar points can be made about verbal reports that also need to be treated as indirect sources. Turkish-speaking children need to learn how the information encoded in the direct and indirect past-tense morpheme corresponds to different types of information access – that is, they have to learn what counts as direct vs. indirect evidence in their language. They also have to learn how these broad types of access bear on their encounters with specific events in the world.

Second, and relatedly, there is some evidence showing that adults’ use of the Turkish past-tense evidential morphemes depends on subtle considerations of how to draw the boundary between types of information access. In a recent study, Turkish speakers used –DI to refer to clearly witnessed events; however, their use of the indirect marker –mIş for non-witnessed, inferred events depended on the ‘distance’ between the available visual clues for the event and the full perceptual experience of the event (Ünal, Pinto, Bunger & Papafragou, subm.; see also Papafragou et al., 2007, for variation in adults’ use of Korean evidentials). Similarly, as mentioned in the Introduction, linguistic studies have shown that the Turkish indirect marker –mIş can be used to communicate the speaker’s dissociation from the event, and has several extended uses include surprise, distance, irony (Johanson, 2003) and ‘lack of preparedness’ for new information (Slobin & Aksu, 1982). In some cases, the appropriateness of the two Turkish past-tense morphemes is subject to convention-driven constraints on what counts as direct vs. indirect experience (e.g., story-telling canonically elicits the indirect past and historical descriptions of events the direct past tense; Slobin & Aksu, 1982). Together, these variable and highly context-dependent distributional patterns conspire to produce challenges for learners, especially in the case of the indirect past tense.
There are further complications from the fact that evidential meanings are encoded in other Turkish morphemes beyond the past-tense system. For instance, the marker –dIR encodes inference from general expectations (e.g., one can use it to convey that Susan must have come, based on knowledge that she usually comes home at 7pm; see Aksu-Koç et al., 2009, for discussion). The acquisition of such markers may interact with the acquisition of past-tense morphemes in complex ways. For example, –dIR could be interpreted to have a generic function and thus be high in reliability, despite being an inferential marker, unlike the past-tense indirect marker –mIş that can be used for inference from (potentially underinformative or misleading) visual premises (see Exp.3).

**Conceptual factors in the acquisition of evidentials**

Despite the clear role of mapping factors, our data strongly suggest that conceptual factors also contribute to the acquisition of evidentiality. Specifically, a third major finding from our studies is that source monitoring is characterized by an asymmetry between direct and indirect evidence, with direct evidence being understood better compared to indirect evidence (see Exps. 4 and 5). For instance, in our studies, Turkish-speaking children often mistakenly attribute indirectly acquired knowledge to direct perception: 5-year-olds report having seen what they only heard or inferred about 60% of the time and older children do so about 40% of the time (Exp.4). Critically, the opposite almost never happens: misattributions of visually perceived information occur in less than 10% of the child data (see Figs. 4a and 4b). This finding is consistent with prior literature on source misattribution in (English-speaking) children and adults which also appears to be uni-directional (Dallett & Wilcox, 1968; Hyman & Billings, 1998; Loftus & Pickrell, 1995; Intraub & Hoffman, 1992; cf. Ceci, Huffman, Smith & Loftus, 1994; Johnson, 1991; Johnson, Raye, Wang & Taylor,
The present data are also consistent with work demonstrating that children are generally able to identify perception as a source of knowledge early on but the understanding of how inference leads to knowledge appears later (cf. Gopnik & Graf, 1988; O’Neill & Astington, 1990; O’Neill & Chong, 2001; O’Neill & Gopnik, 1991; Woolley & Bruell, 1996; Wimmer et al., 1988; Sodian, & Wimmer, 1987; Aksu-Koç et al., 2009; among others). While some of these studies have shown that 5-year-olds can justify their own knowledge in terms of different sources (e.g., perception, hearsay, inference; Gopnik & Graf, 1988; O’Neill & Gopnik, 1991), in our experiments the gap between perception and hearsay/inference persists in older children and extends to reasoning about others’ sources. We attribute this difference from prior studies to the more demanding nature of our tasks. As mentioned earlier, our overall experimental set-up did not involve training in answering source questions or emphasize getting access to information in different ways (see Discussion of Exp.3, and also Exp.4). Thus access to information may have been less salient and/or harder to track for children. Furthermore, the way of granting access to information was not as sharply differentiated in our stimuli as in prior work: since children always watched a clip, there was an element of seeing even in the Hear and Infer trials, just as in many real-life instances where knowledge is acquired. This was particularly challenging in the Infer trials, which involved inference from observable cues (unlike past work that focused on associative or logical inference; O’Neill & Gopnik, 1991; Sodian & Wimmer, 1987). In all these respects, the present task (especially the Hear and Infer component) was more demanding than previous versions of source monitoring tasks. Our data show that, in these more complex circumstances, 5- to 7-year-old children can reliably invoke perception to report the source of their own newly acquired knowledge about events but have difficulties with hearsay/inference (Experiment 4). Similarly, when having to report two characters’ access to an event, children can report perceptual access correctly at age 5,
hearsay access at age 6 and inferential access after age 7 (Experiment 5).

The direct-indirect source asymmetry in our source monitoring tasks has a parallel in the use of the direct and indirect marker in the linguistic evidentiality tasks (see Exps. 1 and 2 – with the exception of the See vs. Infer condition of Exp.2, where there is no difference between correct uses of the direct vs. indirect morpheme). This parallel seems to suggest that part of the difficulty with the acquisition of the –DI/-mIş distinction in Turkish might lie with the difficulty of identifying and remembering indirect sources of information such as hearsay and inference. More concretely, if children tend to mistakenly think they saw something that they only heard or inferred, it is not surprising that they should tend to fail to use the indirect marker when it would have been appropriate. This argument is consistent with the presence of a correlation between linguistic and source monitoring scores in our studies which can be taken to mean that linguistic knowledge about evidentiality builds on and closely follows conceptual knowledge about sources of information. Such a tight relationship between linguistic and conceptual aspects of evidentiality is in accord with previous proposals according to which the order of appearance of individual morphemes within the Turkish evidential system reflects at least in part children’s developing abilities to handle various information sources (e.g., Aksu & Slobin, 1986).

One caveat in interpreting our own and previous tests of source monitoring is that such tests rely on children’s explicit reasoning about mental states using verbs and other linguistic labels (e.g., I saw, I heard). Even though this method seems rather straightforward, occasionally the linguistic material in the experimental scenarios may present challenges to learners that are unrelated to the process of source monitoring per se. For instance, the difficulty of talking about one’s access to hearsay or inference in Experiment 4 might have to do with having acquired the verbs that refer to the act of hearing or inferring; such verbs are likely to be less transparent to children compared to the verb that refers to acts of seeing. It is possible that tasks that do not involve explicit reference
to verbal report or inference but measure source monitoring more implicitly might reveal earlier success with some of the indirect sources that now pose the greatest difficulties for learners. Suggestive evidence in favor of this possibility comes from studies showing that 3-year-olds show some implicit understanding of source reliability despite the fact that they cannot overtly justify the evidential basis of their beliefs (Robinson & Whitcombe, 2003; see also the recent literature on false-belief reasoning in children and infants for a demonstration of the role of implicit vs. explicit tasks on performance: Baillargeon, Scott & He, 2010; Surian, Caldi & Sperber, 2007, among others). Notice that, if more implicit tasks were to reveal earlier understanding of indirect sources of evidence in children, the gap between conceptual source reasoning and knowledge of linguistic evidentiality in Turkish learners would turn out to be even wider than the present studies suggest – thereby adding further support to the role of mapping factors to the acquisition of evidentials.10

Setting aside these observations, our present findings, if taken together, point to a complex characterization of the learning process for evidentiality. Specifically, our data show evidence of a cognitive asymmetry in handling different types of sources – with hearsay and especially inference being understood later compared to direct perception. This asymmetry seems to place limits on the types of semantic information that learners are ready to acquire. Nevertheless, even when source concepts become available, discovering the correspondence between these concepts and linguistic input is far from trivial (as shown by the gap between linguistic and source monitoring scores in our studies). This finding is important because it has been widely assumed that the main source of difficulty in the acquisition of mental-state vocabulary comes from mastering the relevant concepts (see Bartsch & Wellman, 1995; Gopnik & Meltzoff, 1997; but see Gillette et al., 1999; Papafragou et al., 2007 for a different perspective). To the extent that both conceptual and mapping factors

---

10 It is an open question whether more implicit tasks might also reveal earlier sensitivity to the meaning of linguistic evidentiality, even though the issue arises only for the present comprehension task. We are pursuing this possibility in ongoing work.
contribute to Turkish learners’ successes and failures in acquiring evidentials, it is reasonable to expect their relative contributions to change over time. Specifically, as children become older/conceptually more mature, the explanatory burden is likely to shift almost exclusively to mapping factors: for instance, 7-year-olds in our studies can handle both direct and indirect sources in source monitoring tasks but still exhibit non-adult knowledge of evidential morphology.

Our findings have both theoretical and methodological implications for the interface between linguistic and conceptual development. On the theoretical level, our data argue against an often cited assumption in the developmental literature, according to which linguistic development is a more or less straightforward index of cognitive development (e.g., Smiley & Huttenlocher, 1995). Our evidence shows that what children say (or do not say) is determined not only by what they can (and cannot) think about but also by what sorts of linguistic meanings are easy to discover and link to specific lexical or grammatical devices. On the methodological level, our findings underscore the usefulness of non-linguistic, conceptual tasks as an independent way of testing hypotheses about the cognitive basis of language acquisition.
References


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.


Appendix: Stimulus Lists

(For half of the participants, materials were switched between Exps.1 and 4, and 2 and 5).

LINGUISTIC EVIDENTIALITY EXPERIMENTS

Experiment 1: Production of Evidential Morphology

See vs. Hear condition

See trials (animated action clips)
A bird flies to a tree.
A fish swims behind a rock.
A butterfly flies towards the roof of a house.
A ball rolls down the hill.

Hear trials (static clips with one character speaking)
A child says: “I just came back from a school trip.”
A woman says: “I just came back from shopping.”
A woman says: “I went to work today.”
A man says: “I spoke on the phone today.”

See vs. Infer condition

See trials (animated action clips)
A bird flies to a tree.
A fish swims behind a rock.
A butterfly flies towards the roof of a house.
A ball rolls down the hill.

Infer trials (static clips that lead to inference)
A child next to a festively wrapped box. (Inference: Somebody bought a gift for the child.)
A snowman. (Inference: Somebody built a snowman.)
A child with a sad face holding an empty cone and looking at some ice cream on the floor. (Inference: The child dropped the ice cream.)
Fried eggs in a pan. (Inference: Somebody fried eggs.)

Experiment 2: Comprehension of Evidential Morphology

See vs. Hear condition

Video 1: A character plays with a doll while another character is watching.
Video 2: A different character is in the same room looking at the dolls. When a new person comes in, the character points to the dolls and whispers something into this person’s ear.

Video 1: A character reads a book while another character is watching.
Video 2: A different character is in the same room looking at the open book on the table. When a new person comes in, the character points to the book and whispers something into this person’s ear.

Video 1: A character drinks lemonade while another character is watching.
Video 2: A different character is in the same room looking at the half-empty bottle of lemonade. When a new person comes in, the character points to the bottle and whispers something into this person’s ear.

Video 1: A character takes several pens out of a box while another character is watching.
Video 2: A different character is in the same room looking at the empty box and pens. When a new person comes in, the character points to the objects and whispers something into this person’s ear.
Video 1: A character hangs a picture on the wall while another character is watching. Video 2: A different character is in the same room looking at the picture. When a new person comes in, the character points to the picture and whispers something into this person’s ear.

Video 1: A character puts toys in a box while another character is watching. Video 2: A different character is in the same room looking at the toys and box. When a new person comes in, the character points to the objects and whispers something into this person’s ear.

Video 1: A character takes off her sweater while another character is watching. Video 2: A different character is in the same room looking at the sweater. When a new person comes in, the character points to the sweater and whispers something into this person’s ear.

Video 1: A character writes a note on a piece of paper while another character is watching. Video 2: A different character is in the same room looking at the note. When a new person comes in, the character points to the note and whispers something into this person’s ear.

See vs. Infer condition

Video 1: A character plays with a doll while another character is watching. Video 2: A different character comes into the room and inspects the dolls on the table.

Video 1: A character reads a book while another character is watching. Video 2: A different character comes into the room and inspects the open book on the table.

Video 1: A character drinks lemonade while another character is watching. Video 2: A different character comes into the room and inspects the half-empty bottle of lemonade on the table.

Video 1: A character takes several pens out of a box while another character is watching. Video 2: A different character comes into the room and inspects the box and the pens on the table.

Experiment 3: Evaluating Others’ Reliability from Use of Evidentiality

We do not list stimuli for this experiment. In each trial children were presented with an animated scene with two boxes that were identical in shape and size but differed only in color. Two puppets offered contrastive evidential statements about the contents of the boxes.

SOURCE MONITORING EXPERIMENTS

Experiment 4: Reporting One’s Own Sources

See vs. Hear condition

See trials (animated action clips)
A balloon escapes from the hands of a little girl.
A girl goes down a slide.
A frog jumps along a river. 
A car drives down the road.

**Hear trials (static clips with one character speaking)**
A woman says: “I rode a bike.”
A girl says: “I exercised.”
A boy says: “I played basketball.”
A girl says: “I read a book.”

**See vs. Infer condition**

**See trials (animated action clips)**
A balloon escapes from the hands of a little girl.
A girl goes down a slide.
A frog jumps along a river.
A car drives down the road.

**Infer trials (static clips that lead to inference)**
A slice of birthday cake on a plate. (Inference: Somebody cut the cake.)
A Christmas tree decorated with ornaments. (Inference: Somebody decorated the tree.)
An open box with a present on the floor. (Inference: Somebody opened the present.)
A painting on an easel. (Inference: Somebody painted a picture.)

**Experiment 5: Reporting Others’ Sources**

**See vs. Hear condition**

Video 1: A character takes a picture with a camera while another character is watching.
Video 2: A different character is in the same room looking at the camera and picture. When a new person comes in, the character points to the objects and whispers something into this person’s ear.

Video 1: A character knocks down several wooden toy animals while another character is watching.
Video 2: A different character is in the same room looking at the animals. When a new person comes in, the character points to the toys and whispers something into this person’s ear.

Video 1: A character turns on a lamp while another character is watching.
Video 2: A different character is in the same room looking at the lamp that is turned on. When a new person comes in, the character points to the lamp and whispers something into this person’s ear.

Video 1: A character opens the door while another character is watching.
Video 2: A different character is in the same room looking at the open door. When a new person comes in, the character points to the door and whispers something into this person’s ear.

Video 1: A character breaks a pencil while another character is watching.
Video 2: A different character is in the same room looking at the broken pencil. When a new person comes in, the character points to the pencil and whispers something into this person’s ear.

Video 1: A character peels an orange while another character is watching.
Video 2: A different character is in the same room looking at the peeled orange. When a new person comes in, the character points to the orange and whispers something into this person’s ear.

Video 1: A character takes off her jacket while another character is watching.
Video 2: A different character is in the same room looking at the jacket on the table. When a new person comes in, the character points to the jacket and whispers something into this person’s ear.

Video 1: A character cuts an apple in pieces while another character is watching.
Video 2: A different character is in the same room looking at the apple pieces. When a new person comes in, the character points to the pieces and whispers something into this person’s ear.

**See vs. Infer condition**

Video 1: A character takes a picture with a camera while another character is watching.
Video 2: A different character comes into the room and inspects the camera and picture on the table.

Video 1: A character knocks down several wooden toy animals while another character is watching.
Video 2: A different character comes into the room and inspects the toy animals lying around.

Video 1: A character turns on a lamp while another character is watching.
Video 2: A different character comes into the room and inspects the lamp that is on.

Video 1: A character opens the door while another character is watching.
Video 2: A different character comes into the room and inspects the open door.

Video 1: A character breaks a pencil while another character is watching.
Video 2: A different character comes into the room and inspects the broken pencil on the table.

Video 1: A character peels an orange while another character is watching.
Video 2: A different character comes into the room and inspects the orange skin on the table.

Video 1: A character takes off her jacket while another character is watching.
Video 2: A different character comes into the room and inspects the jacket on the table.

Video 1: A character cuts an apple in pieces while another character is watching.
Video 2: A different character comes into the room and inspects the pieces of apple.

**Experiment 6: Evaluating Others’ Reliability from Their Access to Information**

**See vs. Hear condition**

Video 1: A character puts on a scarf while another character is watching.
Video 2: A third character is told about the event by a fourth character.

Video 1: A character takes out books from a backpack while another character is watching.
Video 2: A third character is told about the event by a fourth character.

Video 1: A character hangs a coat on a hanger while another character is watching.
Video 2: A third character is told about the event by a fourth character.

Video 1: A character peels a banana while another character is watching.
Video 2: A third character is told about the event by a fourth character.

**See vs. Infer condition**

Video 1: A character puts on a scarf while another character is watching.
Video 2: A third character comes into the room and inspects the first character wearing a scarf.

Video 1: A character takes out books from a backpack while another character is watching.
Video 2: A third character comes into the room and inspects the books and backpack on the table.

Video 1: A character hangs a coat on a hanger while another character is watching.
Video 2: A third character comes into the room and inspects the coat on the hanger.

Video 1: A character peels a banana while another character is watching.
Video 2: A third character comes into the room and inspects the banana skin on the table.