

The Acquisition of Evidentiality and Source Monitoring

Ozge Ozturk and Anna Papafragou
University of Delaware

1. Introduction

Evidential markers encode the source of a speaker's knowledge. While some languages express evidentiality by lexical markers (e.g. *I saw that it was raining* vs. *I heard that it was raining*), about a quarter of world's languages grammaticalize evidentiality through specialized markers. For instance, Turkish obligatorily marks all instances of past reference with one of the following two suffixes: *-DI* (the neutral form, which denotes the past of direct experience and is realized as *-di, -di, -du, -dü, -ti, -ti, -tu, -tü* depending on the vowel harmony) and *-mİş* (which denotes the past of indirect experience and is realized as *-miş, -miş, -muş, -muş* depending on the vowel harmony). As part of their evidential function, the morpheme *-DI* is used to describe witnessed events and the morpheme *-mİş* is used to describe non-witnessed events, i.e. knowledge acquired from someone else's report (hearsay) or some clue (inference):

- (1) Çocuk gel -DI/ -miş
Child come -PAST direct evidence / -PAST indirect evidence
'The child came' (I saw it) / (I heard it or I inferred it)

This paper is concerned with the acquisition of evidential markers by Turkish-speaking children. In early pioneering work on Turkish, Aksu-Koc (1988) has shown that there is an asymmetry in the acquisition of the two evidential markers: the emergence of the direct marker *-DI* precedes that of *-mİş* in spontaneous speech. On the basis of an elicited production and a comprehension experiment, Aksu-Koc concluded that children's performance with evidential morphology between the ages 3;0 and 5;8 is around chance level; from 5;8 onwards a steady improvement is observed though performance is still far from perfect. These findings have been confirmed by more recent production studies that found non-adult performance with Turkish evidentials by 3- and 4-year-old children (Ozturk & Papafragou, 2007; see also Choi, 1995).

One question that is left open by prior work is the relation between the acquisition of evidentiality and the development of source monitoring abilities in young learners. A crucial part of the acquisition of evidential markers is mapping them onto different types of non-linguistic source categories. Since information sources are themselves abstract and unobservable, their identification as potential candidates for grammatical meaning may prove to be a lengthy and complex process for children.

Additionally, or alternatively, the development of source monitoring abilities themselves may hinder the acquisition of the corresponding linguistic morphology. Recent studies suggest that children develop source-monitoring abilities between the ages of 3 and 6. O'Neill and Gopnik (1991) found that 3-year-olds have difficulty identifying which source led to a belief regardless of source type. Even though young children were able to remember individual events, they failed to relate them to the beliefs that they led to. Four- and 5-year-olds' performance, on the other hand, was found to be considerably better. Children's communicative behavior was found to imply some awareness of another person's visual access to information and their resulting knowledge state (Perner & Leekam, 1986): specifically, 3-year-olds have been found to adjust their speech according to their listeners' prior experience. However, children of this age have difficulty identifying the source of their beliefs.

Verbal information might be more difficult to identify as a source of information than visual access. Three- and 4-year olds have been reported to have difficulty understanding that a person who has been verbally informed about the contents of a box has informational access and would know what is in the box (Wimmer & Hogrefe, 1988; cf. Pratt & Bryant, 1990). In a recent study, Whitcombe (2003) reported that there is an asymmetry in children's understanding of visual vs. verbal access to information. As early as 3-year-old, children have an implicit understanding that visual access to information is more reliable than hearsay evidence, i.e. verbal access.

Finally, inference has been claimed to be particularly challenging as an information source. Sodian & Wimmer (1987) suggest that inference is a less common and less easily observable source of knowledge than perception and communication. In their studies, 4- and 5-year-olds were able to perform a simple inference and thereby came to know a certain fact. However, they did not understand how their own knowledge was obtained. Moreover, they also neglected someone else's inferential access to information even when the other person shared their own perspective. An understanding that inference leads to knowledge may not be acquired until as late as age 6 (Sodian & Schneider, 1990; cf. also O'Neill & Gopnik, 1991).

In this paper, we explore the relation between linguistic evidentiality and source monitoring. We report findings from two experiments conducted with Turkish learners between the ages of 5 and 7. Experiment 1 systematically examines the production of the full range of evidential meanings (direct vs. indirect: hearsay/inference) encoded in the Turkish past tense system and attempts to chart their developmental timetable. Experiment 2 is designed to be closely parallel to the production experiment and it investigates children's non-linguistic source monitoring abilities building on some of the methods available in the literature for tracking early source monitoring. The goal of this investigation is to clarify the scope of children's early difficulties with linguistic evidentiality and the nature of these difficulties.

One possible outcome is that the acquisition of the linguistic evidentiality may rely on the development of non-linguistic source monitoring abilities. In

that case, we would expect an asymmetry between our linguistic and non-linguistic findings such that non-linguistic evidential concepts should precede and form the basis for subsequent linguistic development.

Another possibility is of the exact opposite nature: children's linguistic development might precede their non-linguistic (conceptual) development. It has been suggested in previous literature (Whorf, 1956) that the use of linguistic representations might enhance the development of non-linguistic categories. In other words, systematic encoding of a certain conceptual category in grammar may encourage the young learner to develop a previously non-existing (or not as complex) non-linguistic concept. If this view is correct, we may find an asymmetry between our linguistic and non-linguistic findings in favor of the knowledge of linguistic evidentiality, which in turn is expected to help the acquisition of the non-linguistic source-monitoring concepts.

Yet a third possibility is that conceptual and linguistic development might proceed in parallel. If linguistic evidentiality and non-linguistic source monitoring abilities develop simultaneously and exert their mutual influences, there should be no difference between our linguistic and non-linguistic findings.

2. Experiments

2.1 Participants

A total of 96 monolingual Turkish-speaking children participated in this study.¹ Children were assigned to one of the three groups on the basis of their age: Group 1 (mean: 5;10, range: 64-72 mo); Group 2 (mean: 6;6, range: 73-84mo); Group 3 (mean: 7;8, range: 85-96mo). Each group included 32 children. All children came from upper-middle-class families and they were recruited either from a preschool or a grade school in Istanbul, Turkey.

2.2 Experiment 1: Elicited Production

2.2.1 Stimuli and Procedure

In this experiment we attempted to elicit children's production of the evidential morphemes for direct evidence (*-DI*) and indirect evidence (*-mİş*). Stimuli were presented on laptop screen and they consisted of animated scenarios. The digitized audio for the animations was recorded from the voice of a native Turkish speaker. There were two between-subjects conditions: **See vs. Infer** and **See vs. Hear**. At the beginning of the experiment half of the participants were randomly assigned to the **See vs. Infer** condition and the other half were assigned to the **See vs. Hear** condition. In each condition we had a total of 8 trials. 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 the participant watched something happen (e.g. a girl

¹ Ten monolingual Turkish-speaking adults also participated as a control group. Adult results were at ceiling and are not reported here.

jumped over the stone). In the **Hear** trials, the participant heard the character in the animation utter a sentence (e.g. a woman said: “I went shopping today.”). In the **Infer** trials, the participant saw some hints indicating something had happened. After each trial the experimenter encouraged the participant to say what happened by beginning to utter a sentence. However, she did not finish the sentence and let the participant finish it:

- (2) Kiz tas -in ust -u (n)-den...
 Girl stone-GEN above-3sg. -abl...
 ‘The girl over the stone...’

Turkish is an SOV word order language; hence the verb’s unmarked position is at the end of a sentence. The evidential markers are verbal suffixes. By not finishing the sentence the experimenter avoided using the evidential marker and gave the participant the chance to do so. If the participant witnessed the event (**See** trials), the participant was expected to employ the evidential morpheme for direct evidence *-DI*. In the **Hear** and **Infer** trials, however, the participant did not witness the event, hence the indirect evidence morpheme *-mİş* was required.

Two pseudo-random orders of presentation were employed for a total of 8 trials. Materials for the **See** trials were identical in the two between-subjects conditions.

2.2.2 Results

Results are presented in Figures 1 and 2. Children only employed the past tense morphemes *-DI* and *-mİş* in their answers and no other morpheme. So, whenever children made a mistake, they replaced the correct evidential morpheme with the incorrect one (usually by overgeneralizing the direct marker).

A 3 (Age: 5, 6, 7) x 2 (Condition: **See vs. Hear**, **See vs. Infer**) ANOVA with the number of correct responses as the dependent variable revealed only a significant main effect of Age ($F(2, 90) = 4.182, p < .05$): pairwise comparisons revealed that the oldest children performed significantly better than each of the other two age groups ($M_5 = 68.75, M_6 = 70.25, M_7 = 78.75; p < .05$), but performance did not differ between the two youngest groups.

Next, we had a closer look at children’s performance within each condition. Beginning with the **See vs. Hear** group, a 3 (Age: Age 5, Age 6, Age 7) x 2 (Source: **See**, **Hear**) ANOVA with the number of correct responses as the dependent variable revealed a significant main effect of Source ($F(1, 189) = 49.570, p < .001; M_{\text{see}} = 95, M_{\text{hear}} = 43$), no significant main effect of Age and a significant interaction between Source and Age ($F(2, 189) = 4.743, p < .01$). This interaction was due to the fact that performance on the **See** items did not differ across ages ($M_5 = 94, M_6 = 80, M_7 = 92$) while in the **Hear** items there are differences among age groups ($M_5 = 48, M_6 = 63, M_7 = 70$; the difference between 5- and 7-year-olds significant, $t(126) = -2.564, p < .05$).

The same analysis for the **See vs. Infer** condition returned again a significant main effect of Source ($F(1, 189) = 79.976, p < .001; M_{\text{see}} = 98, M_{\text{infer}} = 62$), no significant main effect of Age and a significant interaction between Source and Age ($F(2, 189) = 3.370, p < .05$). The interaction was due to the fact that there was a difference between the oldest and youngest group in the **Infer** items only ($t(126) = -2.329, p < .05$) but there were no significant age differences for the **See** items.

Separate one-sample t-tests for each age group per condition and source revealed that children's performance on the **See** items was consistently different from chance ($p < .05$) in both conditions. For the **Hear** items, performance was significantly different from chance only in the two older age groups (for 6-year-olds: $M = 63, t(63) = 2.049, p < .05$; for 7-year-olds: $M = 70, t(63) = 3.529, p < .005$). Finally, for the **Infer** items, performance was not different from chance even for the oldest children in our sample.

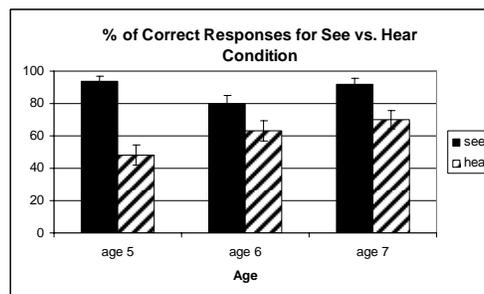


Figure 1. Percentage of correct responses for the **See vs. Hear** condition

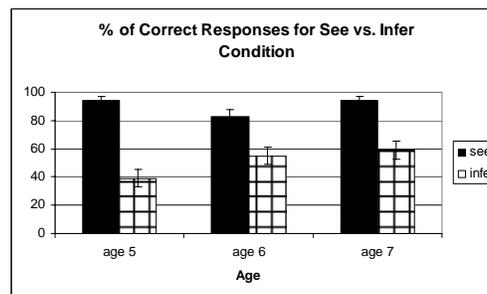


Figure 2. Percentage of correct responses for the **See vs. Infer** condition

2.2.3 Discussion

The results of the production experiment support the existence of an acquisition hierarchy of the direct and indirect evidential morphemes: children - even in the youngest age group - produced the morpheme *-DI* for direct perception cases with almost no difficulty. However, their performance for the indirect

experience morpheme *-mİş* (encoding hearsay and inference) was shown to improve with age. The error pattern we observe in the data suggests that whenever the child has direct access to information she is able to use the direct evidential morpheme (most of the time); however, whenever she has indirect access to information she mostly fails to use the indirect morpheme *-mİş* and instead uses the direct evidential morpheme *-DI*. These results confirm earlier data suggesting that Turkish learners have acquired the past tense function of the two morphemes before they begin to differentiate them in terms of their evidential meanings (Aksu-Koc, 1988; Ozturk & Papafragou, 2007).

Additionally, the production results reflect an asymmetry within the indirect experience morpheme between the two indirect sources: hearsay and inference. Six-year-olds in our study employed this morpheme solely with the hearsay function but not with the inference function. Moreover, our results indicate that even the 7-year-olds' performance with inference items is around chance level. This finding suggests that the two types of indirect sources, hearsay and inference have different levels of complexity for the young child. This complexity may be conceptual (cf. Sodian & Wimmer, 1987), or it may have to do with the fact that inference is a less common and less easily observable source of information than perception and communication. The sources of this difficulty are examined in Experiment 2.

2.3 Experiment 2: Reporting One's Own Sources

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

2.3.1 Stimuli and Procedure

Stimuli and design were as in Experiment 1. As before, in the beginning of the experiment, half of the participants were assigned to the **See vs. Infer** condition and the other half were assigned to the **See vs. Hear** condition. Assignment to conditions followed assignment in the previous experiment.

In the **See vs. Hear** condition, the participant either saw something happen on the screen (e.g. a fairy fly from the flower to the tree) or she heard a character on the screen utter a sentence (e.g. "I went shopping today."). In the **See vs. Infer** condition, the participant again either saw something happen on the screen or she saw some hints indicating that something had happened (e.g. a cake that was partially eaten) and was expected to infer what happened (here, that somebody must have eaten the missing pieces of the cake). After each trial, the experimenter asked the participant to report what happened. Immediately after the participant's answer, the experimenter asked about the participant's source of information:

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

(4a) **See vs. Hear** condition

Gor-du -n -mu? Dinle -di -n -mi?
see-Past-2sg.-Q. Listen -Past-2sg.-Q.
'Did you see? Did you listen/hear?'

(4b) **See vs. Infer** condition

Gor-du -n -mu? Ya da bir sey-ler gorup mu anla -di -n?
see-Past-2sg.-Q. Or one thing-Pl. see -Q.understand-Past-2sg.-Q
'Did you see or did you see something and understood?'

2.3.2 Results

Results are presented in Figures 3 and 4. A 3 (Age: 5, 6, 7) x 2 (Condition: **See vs. Hear**, **See vs. Infer**) ANOVA with the number of correct responses as the dependent variable revealed a significant main effect of Condition ($F(1, 90) = 2.101, p < .001$): overall, children performed better in the **See vs. Infer** condition than they did in the **See vs. Hear** condition (see Figures 3 and 4). However, there was no significant main effect of Age or interaction between Condition and Age. We suspect that one of the reasons performance is lower in the **See vs. Hear** condition is the choice of lexical verb for the **Hear** items. We chose the Turkish counterpart of the verb "listen" (see (4a) above) which, for some reason proved confusing for the children in the **Hear** items. In previous research with younger children, a different Turkish verb (the equivalent of 'hear') led to higher success rates in similar trials (Ozturk & Papafragou, 2007).

Next, we had a closer look at each condition. For the **See vs. Hear** condition, a 3 (Age: 5, 6, 7) x 2 (Source: See, Hear) ANOVA with the number of correct responses as the dependent variable was conducted. A significant main effect of Source was revealed ($F(1, 190) = 148.89, p < .01; M_{\text{see}}=95, M_{\text{hear}}=43$). Moreover, a significant main effect of Age was revealed ($F(2, 189) = 3.263, p < .05; M_5=72, M_6=62.5, M_7=72.5$): independent samples t-tests revealed a significant difference of performance between ages 5 and 6 ($t(127) = -3.42, p < .05$) and 6 and 7 ($t(127) = -5.39, p < .05$) but not between ages 5 and 7. There was no significant interaction between Source and Age.

For the **See vs. Infer** condition, a similar ANOVA yielded a significant main effect of Source ($F(1, 190) = 4.2, p < .016; M_{\text{see}}=97, M_{\text{infer}}=62$). Moreover, a significant main effect of Age was revealed ($F(2, 189) = 6.264, p < .013$), $M_5=73, M_6=82, M_7=84$). Independent samples t-tests revealed a significant difference of performance between the ages 5 and 6 ($t(127) = -2.049, p < .05$) and 5 and 7 ($t(127) = -3.167, p < .05$) but not between the ages 6 and 7. There was no significant interaction between Source and Age.

Finally, we examined the performance of each age group per condition and source. For the **See** items, performance was consistently different from chance for even the youngest children in both between-subjects conditions (one-sample t-tests, p 's < .05). For the **Hear** items, the youngest and oldest children did not perform at a level different from chance; only the middle age group did so, but its performance was consistently incorrect ($M = 38, t(63) = -2.05, p < .05$, two-

tailed). For the **Infer** items, the youngest children's performance did not differ from chance but the other two groups' did (6-year-olds: $M = 64$, $t(63) = 2.326$, $p < .05$, two-tailed; 7-year-olds: $M = 69$, $t(63) = 3.211$, $p < .005$, two-tailed).

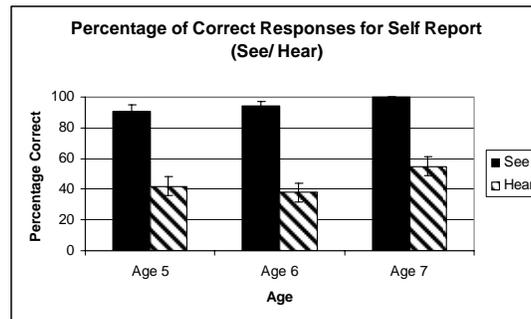


Figure 3. Percentage of correct responses for the **See vs. Hear** condition

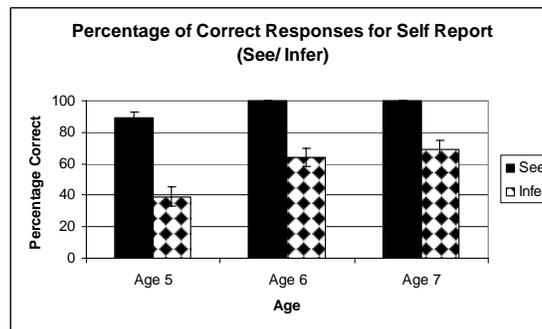


Figure 4. Percentage of correct responses for the **See vs. Infer** condition

2.3.3 Discussion

The findings of this experiment support the claim that there is a hierarchy in the development of non-linguistic conceptual source monitoring abilities. Children were first able to successfully report their own sources of information for the direct access (**See**) type of items, even in our youngest age group. Performance on the indirect access (**Hear** and **Infer**) type of the items was shown to improve over time, a finding in line with previous literature claiming an asymmetry in the development of different types of sources of information. According to this literature, visual perception is the most readily source accessible to young children, with verbal information following visual perception and inference being the most complex and least accessible information source for the young child (Wimmer & Hogrefe, 1988).

2.4 Comparison of Production and Source Report Experiments

A combined score per subject for each experiment was calculated. A repeated-measures ANOVA with the percentage of correct responses of each participant as the dependent variable, Experiment (1, 2) as a within-subjects variable and Age (5, 6, 7) as a between-subjects variable revealed a significant main effect of Age ($F(2,381) = 7.748, p < .005$): this effect was due to the fact that the oldest children's performance differed from that of the other two groups ($p < .05$) even though the two youngest groups did not differ from each other ($M_5 = 70.71, M_6 = 71.1, M_7 = 78.7$). The analysis revealed no significant main effect of Experiment ($M_{\text{PROD}} = 72.53, M_{\text{SOURCE-REPORT}} = 74.48$) but it yielded a significant interaction between Age and Experiment ($F(1,383) = 2397.6, p < .001$). Separate ANOVAs revealed a main effect of Age in the Production ($F(2,381) = 4.784, p < .01$) but not the Source Report Experiment, which explains the interaction.

Next, we analyzed performance on the two experiments separately for each of the two conditions. Beginning with the **See vs. Hear** condition, a repeated measures ANOVA with Age and Experiment as factors revealed a significant main effect of Age ($F(2, 189) = 4.447, p < .05$) but no significant main effect of Experiment ($M_{\text{PROD}} = 74.48, M_{\text{SOURCE-REPORT}} = 69.01$), or interaction between Age and Experiment. The same analysis for the **See vs. Infer** condition revealed a significant main effect of Age ($F(2, 189) = 6.099, p < .005$), a significant main effect of Experiment ($F(1, 189) = 11.127, p < .005; M_{\text{PROD}} = 70.57, M_{\text{SOURCE-REPORT}} = 79.95$) and no interaction.

The comparison of Experiments 1 and 2 revealed an asymmetry between children's linguistic command of evidential morphology and the ability to report non-linguistic knowledge sources, with non-linguistic source monitoring preceding the acquisition of linguistic evidentiality. However, this asymmetry emerged only in the **See vs. Infer** condition: in the **See vs. Hear** condition, children performed equally well in the two experiments. We hypothesize that the reason the asymmetry does not generalize to the **See vs. Hear** condition is the choice of verb ('listen') in the sentence used to report the child's source of information (see Results section in 2.3.2). Confusion about the interpretation of the verb may have led to lower success rates in the non-linguistic source monitoring task and may have masked an asymmetry between the use of the hearsay morpheme and the ability to report communication as a source of information.

3. Conclusion

In this paper, we have tested 5-7-year-old Turkish learners' productive command of evidential morphology to mark direct vs. indirect access to information; we have also compared this knowledge to the ability of the very same children to report how they acquired new knowledge. We have observed a protracted developmental sequence for children's acquisition of both linguistic evidentiality and source monitoring. Children's performance for direct

experience is almost adult-like starting with the age of 5. However, children's performance for indirect experience is not adult-like even by the age of 7. Hence, we conclude that there is an asymmetry between Direct vs. Indirect Access to information both on the conceptual and on the linguistic level.

Moreover, for the **See vs. Infer** group we have observed that the non-linguistic source reasoning abilities emerge prior to linguistic evidentiality in line with previous literature (Papafragou, Li, Choi, & Han, 2007 on Korean). This finding leads us to conclude that conceptual development cannot be (solely) responsible for the difficulty associated with the acquisition of linguistic evidentiality. Even after children acquire source monitoring concepts, evidential markers still need to be mapped onto these concepts. Our results can thus be interpreted as pointing to a role for Mapping factors in the acquisition of evidential markers (cf. Gleitman, 1990; Snedeker & Gleitman, 2004 on mental verbs). This mapping process for evidentiality may prove to be challenging since there are no direct observational correlates for informational access, especially for indirect experience. We conclude that both conceptual and mapping factors jointly contribute to the acquisition of linguistic evidentiality.

References

- Aksu-Koc, A. (1988). *The Acquisition of Aspect and Modality: The case of past reference in Turkish*. Cambridge: Cambridge University Press.
- Choi, S. (1995). The development of epistemic sentence-ending modal forms and functions in Korean children. In J. Bybee & S. Fleischman (Eds.), *Modality in Grammar and Discourse* (pp. 165–204). Amsterdam: Benjamin.
- Gleitman, L. (1990). The structural sources of word meaning. *Language Acquisition*, 1, 3-55.
- O'Neill, D. & Gopnik, A. (1991). Young children's ability to identify the sources of their beliefs. *Developmental Psychology*, 27, 390-397.
- Ozturk, O. & Papafragou, A. (2006). Evidentiality and theory of mind. *Proceedings of Boston University Conference on Language Development*, 31. Somerville, MA: Cascadilla Press.
- Papafragou, A., Li, P., Choi, Y. & Han, C. (2007). Evidentiality in language and cognition. *Cognition*, 103, 253-299.
- Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. In Hall & Waxman (Eds.), *Weaving a lexicon* (pp. 255-293). Cambridge, MA: MIT Press.
- Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development*, 58, 424–433.
- Whitcombe, E., & Robinson, E. (2000). Children's decisions about what to believe and their ability to report the source of their belief. *Cognitive Development*, 15, 329-346.
- Whorf, B. L. (1956). *Language thought and reality*. Cambridge, MA: MIT Press.
- Wimmer, H., Hogrefe, G.-J., & Perner, J. (1988). Children's understanding of informational access as a source of knowledge. *Child Development*, 59, 386–396.

