Journal of Memory and Language xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

# Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation

Sümeyra Tosun *, Jyotsna Vaid, Lisa Geraci

*Department of Psychology, Texas A&M University, United States*

ARTICLE INFO

*Article history:*
Received 23 August 2012
revision received 19 March 2013
Available online xxxx

*Keywords:*
Evidentiality
Turkish
Source monitoring
Recognition memory
Linguistic relativity
Thinking for speaking

ABSTRACT

This study examined the influence of obligatory linguistic marking of the source of information on source memory. Turkish grammar requires speakers to indicate if an assertion is based on first hand knowledge or non-firsthand knowledge (hearsay or inference); English grammar does not require this distinction. We hypothesized that obligatory coding of source of evidence leads to a greater weighting of first hand relative to non-firsthand accounts of events (an "evidentiality effect"), resulting in better memory for first hand sources. In support of this hypothesis, across two experiments native Turkish speaking adults showed significantly better recognition and source memory for assertions coded with first hand than non-firsthand evidential markers. Further, among Turkish speakers who also knew English, those who learned English later had less accurate recognition and source memory for non-firsthand sources presented in English than those who learned English earlier, suggesting a carryover from the first language (Turkish). English monolingual speakers showed no difference in recognition or source memory as a function of source type, but showed better memory than Turkish speakers for non-firsthand sources. These findings provide the first empirical support for an evidentiality effect, suggesting that when marking the source of evidence is required by the grammar first hand sources are privileged in memory and non-firsthand sources are discounted.

Published by Elsevier Inc.

## Introduction

The notion, attributed to Whorf (1956), that language serves as a kind of filter for perceiving and internalizing the world has seen a recent resurgence of interest (Boroditsky, 2003; Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2005) and different formulations of the Whorfian hypothesis of linguistic relativity are being theorized (Wolff & Holmes, 2011), investigated, and contested (e.g., January & Kako, 2007). Whereas the literature on linguistic relativity has primarily sought to examine influences of particular languages on *nonlinguistic* cognition, that is, cognition examined under conditions in which language is not being used, overtly or covertly (Lucy, 1996), a different approach has been taken by investigators seeking to examine the thinking that precedes and surrounds language in use. This approach, termed "thinking for speaking" by Slobin (1996, 2003) argues that differences across languages in the semantic distinctions their users are required to make by virtue of their grammar can also affect the thinking required for *linguistic* cognition. Moreover, given that a large part of everyday cognition involves preparing, producing, understanding and remembering verbal messages, and that some events exist only in a verbal realm (e.g., reported events), a more complete investigation of linguistic relativity should also address cognitive processes that arise in the course of using language (Slobin, 2003).

According to the thinking for speaking view, structural differences in how languages codify events may affect how speakers come to attend to, talk about and represent events. Many studies conducted within this framework

* Corresponding author. Address: Department of Psychology, Texas A&M University, College Station, TX 77843-4235, United States. Fax: +1 979 845 4727.

*E-mail address:* sumeyra@neo.tamu.edu (S. Tosun).

0749-596X/$ - see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.jml.2013.03.004

have shown differences across languages in how speakers describe motion events or spatial relations (Slobin, 2003; Wolff & Holmes, 2011). Recently, crosslinguistic studies have also investigated the role of language in perceptions of causality or agency. For example, whereas English and Spanish have available similar linguistic means for constructing agency (using the active voice, a fully specified agent, and a transitive verb, (as in *Phia broke the vase*), Spanish speakers can explicitly de-emphasize agency through the use of a polysemous clitic (*se*), for which there is no equivalent in English. This clitic is used when the agent is either unknown or considered unimportant, or when reference is made to unplanned or accidental occurrences. Studies have found that English and Spanish speakers differ in their memory of whether an agent was present in a scene they viewed involving accidental occurrences (Fausey & Boroditsky, 2011) and in their frequency of invoking agents when summarizing passages depicting such occurrences (Cunningham, Vaid, & Chen, 2010).

A number of other studies have shown that subtle differences within a language in the linguistic framing of an event, e.g., whether it is referred to with a definite or an indefinite article (Strack & Bless, 1994), whether a particular verb is used (e.g., *hit* vs. *smashed*), or a particular verb aspect – imperfective vs. perfective (e.g., *was taking hush money* vs. *took hush money;* see Fausey & Matlock, 2011) may also affect users' interpretations of and judgments in response to an event. Studies conducted within a thinking for speaking framework as well as those examining the cognitive effects of different linguistic realizations share a common goal of demonstrating that properties related to the grammar or lexicon within or across languages may shape mental activity associated with language use. What distinguishes the two approaches is that the thinking for speaking approach (like the linguistic relativity approach more generally) rests on cross-language contrasts between habitual uses of a particular language structure (typically, a structure required by the grammar) and occasional/voluntary uses of a structure (reflecting an element of speaker choice).

The present research used a thinking for speaking approach to examine memory for directly vs. indirectly experienced events across languages as a function of evidential marking. Evidentiality is a semantic category present in a quarter of the world's languages and refers to the linguistic encoding of source of evidence, that is, whether the basis of an asserted proposition is first hand knowledge (e.g., knowledge based on seeing or witnessing an event) or non-firsthand knowledge (e.g., knowledge based on inference or hearsay). Evidentiality is closely related to but considered distinct from epistemic modality, the marking of a speaker's degree of confidence in the truth of an asserted statement (Aikhenvald, 2004; Palmer, 2001). In many Balkan, Turkic, East European, Middle Eastern, and Native American languages source of evidence is encoded at the level of the grammar (Aikhenvald, 2004; Johanson, 2000), whereas many western European languages do not require the marking of source of evidence. Whereas all languages allow users to indicate source of evidence they differ in whether source marking is required by the grammar, in how source may be conveyed when it *is* required, and in what types of sources are encoded linguistically (Chafe & Nichols, 1986; Lazard, 2001; Plungian, 2001). For example, among languages that require the coding of source in the grammar, a distinction is made between direct evidentials, used when the speaker has first hand, perceptual evidence for an action or event, and indirect evidentials, used when the speaker did not personally witness the event but learned of it after the fact, either on the basis of an inference from available physical evidence or on hearsay. Languages may differ in their number of direct evidentials (either containing separate markers for each sensory modality or a single marker for all sensory modalities) and in their number of indirect evidentials (containing a single indirect evidential marker for any kind of nonfirsthand source or separate indirect evidentials for hearsay, inference, or quoted sources). Similarly, among languages in which source information is conveyed in the lexicon there may be differences in the number and type of lexical markers of direct and indirect sources.

Although there is clearly considerable variability across languages in how evidentiality may be marked, one may nevertheless classify languages into two types: those in which evidentiality is marked at the level of the grammar (and thus is routinely attended to) and those in which it is marked optionally, at the level of the lexicon. This raises an interesting question from a language and thought perspective: might speakers whose language requires them to encode source of evidence routinely (in the grammar) become more sensitized to source information (in linguistic or nonlinguistic contexts) as compared to those whose language does not require them to code source of evidence? This question has been posed by various scholars in the past, including Whorf himself (1956, p. 85; see Gerrig & Banaji, 1994; Robinson, 2009; Slobin, 2003) but has not so far been tested in adults. In a review of studies of language and thought, Gerrig and Banaji (1994) highlighted the case of evidentiality marking in Turkish and suggested that "If the experience of language acquisition focuses obligatory attention on a distinction that might otherwise be only voluntarily visited, we might fruitfully explore the possibility of lingering effects on cognition" (p. 254). The present study sought to test this possibility. It is, to our knowledge, the first experimental study of the "lingering effects on cognition" of evidentiality marking. The study examined the impact of evidential marking on recognition and source memory for first hand and nonfirsthand sentences in Turkish and English adults. Before describing the study further we review some characteristics of the marking of source of evidence in Turkish and English and describe relevant source memory findings.

### *Evidentiality marking in English and Turkish*

In English, source of evidence is conveyed at the lexical level. For example, to convey that an event was directly experienced or that the directness or indirectness of evidence is not relevant, a speaker would simply use the past tense of a verb in describing the event, saying, for example, that *Mary came first in the race*. To convey that an event was not directly known a qualifier may be added, e.g., *Apparently*, *Mary came first in the race*. Indeed, to express

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

non-firsthand accounts, one can select from a range of modal adverbs (e.g., *supposedly*, *reportedly,* or *apparently*), modal adjectives, (*it is probable that...),* mental state predicates (*I think...*) or modal auxiliaries (*...must have...*) (see Nuyts, 2001). Importantly, for speakers of English the decision about whether to convey the source of information about some proposition in an utterance is a matter of speaker choice; it is not something that is required by the grammar. Once the decision is made to mark the source of evidence, semantic and/or pragmatic considerations may come into play to influence how to mark it (DeLancey, 2001).

In contrast, Turkish is a well-cited example of a language in which the source of information is obligatorily marked (Aksu-Koc & Slobin, 1986). That is, when referring to a past event, speakers of Turkish must differentiate between whether or not they witnessed the event first hand. First hand accounts require a particular suffix <-*dI*> attached to the root of the verb, and non-firsthand accounts require a different suffix <-*mIş*>. The first hand past tense form is used when a narrated event was directly and consciously experienced by the narrator. For example, *Bu sabah mükellef bir kahvaltı yap**tı***. 'This morning s/he had a substantial breakfast (I saw).' Use of the non-firsthand past tense form indicates that the event was not directly or consciously experienced. For example, *Bu sabah mükellef bir kahvaltı yap**mış*** 'This morning s/he had a substantial breakfast (I heard from someone else, or I saw by what was on the breakfast table, or I was surprised to discover)'. In fact, there are four different contexts identified by linguists in which the non-firsthand suffix in Turkish is used: when the event is inferred or assumed to have occurred, when the narrator hears of the event through someone else, when the narrative event or its effect is experienced as unexpected or surprising, and when the narrated event refers to a realm of discourse that is outside of the ordinary, as in the telling of jokes, myths, folktales, or dreams (Aikhenvald, 2004; Aksu-Koc & Slobin, 1986; Johanson, 2000).

### *Evidentiality marking and source monitoring in children*

Previous empirical studies of evidentiality have primarily taken a developmental approach, seeking to determine the course of acquisition of evidential markers and the relationship between children's use of evidential markers to express information source and their nonlinguistic source monitoring skills (see Matsui & Fitneva, 2009, for an overview). Based on naturalistic and elicited data, studies with Turkish-speaking children have established that appropriate use of evidential markers emerges by the age of three (with first hand use preceding the various non-firsthand uses by a few months, and with production of evidentials preceding stable comprehension by a year or more), nonlinguistic understanding of different knowledge sources emerges around age four, and a more explicit understanding of the relationship between the linguistic form and the source of knowledge it conveys (e.g., perceptual, inferred, assumed, or reported) emerges only around age six and may not be stable until age 7 or later (see Öztürk, 2008; Aksu-Koç, Ögel-Balaban, & Alp, 2009, for a review). This general developmental pattern has been found in other evidentially-marked languages as well, such as Korean (see Papafragou, Li, Choi, & Han, 2007).

In view of this early emergence of evidentiality markers in children's speech, Aksu-Koç et al. (2009) hypothesized that "children's ability to use evidentials would predict their ability to understand knowledge acquisition through different source modalities and to identify the source of knowledge held in memory" (p. 21). The hypothesis predicts that awareness of and memory for source of knowledge will be facilitated for Turkish-speaking children, sensitizing them early to the relationship between knowledge source and corresponding mental states (of belief, ignorance, or certainty).

In a test of this hypothesis, Ögel (2007, cited in Aksu-Koç et al. (2009)) studied the performance of 3–6 year old Turkish speaking children on an elicited production task and on two source monitoring tasks adapted from previous studies with English speaking children. In the production task the children had to describe an event that happened to them personally, an event they heard about from a friend, and an event in which they had to infer something that happened. Their appropriate use of first-hand and non-firsthand markers was assessed. The first source monitoring task used the hidden objects paradigm; the children were presented with six containers and asked to describe what was in them and how they knew that. For two of the containers the contents were plainly visible, for two others the contents could be guessed based on an observable clue, and for the remaining two the contents were told to the children. In the second source monitoring task, involving memory for source of information, the children were taught a set of novel facts by the experimenter or a puppet. A week later they were asked to remember the facts and how they knew them. Ögel (2007) found that performance on the two source monitoring tasks was better for older than younger children and for first hand than inferred or reported sources. No significant relationship was found between the use of evidential markers and performance on the first source monitoring task. However, on the second source monitoring task (assessing memory for source) there was a clear predictive relationship between the use of the reportative marker in the production task and source memory accuracy (see also Öztürk, 2008). Moreover, when comparing results on the source memory task to those reported for English-speaking children by Drummey and Newcombe (2002) it was noted that the Turkish 4 year olds outperformed their English speaking counterparts (40.3% vs. 24.1%). These findings were interpreted by Aksu-Koç et al. (2009) as support for their hypothesis that evidentials confer an advantage for Turkish speakers in distinguishing between different sources of knowledge.

A study with Korean speaking 3 and 5 year olds and English speaking counterparts used an adaptation of the hidden objects source monitoring task (Papafragou et al., 2007). In the 'self' condition children were allowed to see a hidden object in a dollhouse or were merely told where the object was hidden. Then the children were asked where the object was hidden and how they knew that. In the 'others' condition the children saw two characters; one looking into a container and discovering what was in

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

it while the other character just kicked the container. The children were then asked to state the characters' knowledge state based on the source of their knowledge. Papafragou et al. (2007) found that both age groups successfully monitored source of information but 5 year olds performed significantly better than the 3 year olds, and performance overall was better on the 'self' task than the 'others' task. As in the studies with Turkish speakers a positive relationship was observed between elicited production of non-firsthand reportative markers and accuracy of monitoring of reported information. However, no group differences were noted between Korean and English speakers in source monitoring accuracy, leading the authors to conclude that the presence of evidential markers in a language does not lead to enhanced source monitoring.

Taken together, the existing studies indicate that there is a developmental hierarchy in acquisition of evidential markers, with firsthand forms emerging earlier in acquisition than non-firsthand markers. There is also preliminary support for the notion that growing up in a linguistic community in which the source of knowledge is marked grammatically facilitates awareness of different sources of knowledge (Robinson, 2009) but the evidence is mixed: whereas Ögel (2007) showed an advantage for speakers of evidentially-marked languages compared to speakers of English in source memory, Papafragou et al. (2007) showed no difference.

Although important and informative, studies of source monitoring ability development in young children can be difficult to interpret because asking children to articulate how they know something places a computational burden on them by requiring them to maintain certain kinds of information at a conscious or explicit level for correct performance (Aksu-Koç et al., 2009). This may lead to lower performance, particularly in comprehension or memory tests, and to an underestimation of children's actual implicit understanding at a particular age, which may in turn complicate age-based or cross-language comparisons. Clearly, more crosslinguistic studies are needed in order to adequately test the hypothesis of an evidentiality effect in source monitoring ability. To this end, comparative studies with adults would be particularly valuable. To date there has been no crosslinguistic investigation of evidential marking and source monitoring ability in adult speakers. This was the aim of the present research, which examined source memory in adult speakers of Turkish and English. Whereas the developmental studies largely examined non-linguistic source monitoring, our study examined source memory for sentences varying in form (firsthand vs. non-firsthand).

### *Evidentiality marking and source monitoring in adults: the present study*

A source monitoring task was used in the present research, building on the extensive literature on source monitoring in adults (see Johnson, Hashtroudi, & Lindsay, 1993, for a review). In the typical source monitoring paradigm, people study material and are later presented with that material again as well as new material. Participants have to indicate not only whether they recognize the material, but they also have to identify the source of the material. Typically, the recognition decision is an easier decision—simply requiring the participant to judge whether the item is familiar. The source decision is more difficult, as it requires one to determine more specific information about the study episode. As Johnson et al. suggested in their seminal paper on source monitoring, "The term source refers to a variety of characteristics that, collectively, specify the conditions under which a memory is acquired (e.g., the spatial, temporal, and social context of the event; the media and modalities through which it was perceived)" (1993, p. 1). Thus, the source of a memory can refer to the physical context (the room, the particular study list, or the city) in which the information was first encountered. It can also refer to the way in which the information was encountered. For example, in some studies, participants are asked to indicate whether a word was heard or seen (e.g., Henkel, Franklin, & Johnson, 2000; Markham, Howie, & Hlavacek, 1999) or even imagined—a decision process called *reality monitoring*, Johnson & Raye, 1981). Source memory can be studied by examining the extent to which people can accurately recall the source information (the approach taken in the current study) and by examining the specific patterns of errors, which is useful for understanding how source errors can give rise to false memories (e.g., Johnson & Raye, 1981; Roediger & Geraci, 2007; see also Roediger & McDermott, 2000, for a review of the role of source monitoring in the creation of false memories). In the current study, we are interested in the effect of one linguistic feature, the obligatory marking of source (first hand vs. non-firsthand) on memory for those sources.

Other studies that have examined the source monitoring of linguistic information have examined memory for the voice in which an item was presented (e.g., Johnson, De Leonardis, Hashtroudi, & Ferguson, 1995) and the language in it was presented (see Francis, 1999, for a review). In addition, researchers have examined the use of specific linguistic features in source monitoring decisions (Geraci & Franklin, 2004). For example, based on an examination of source errors, we know that people use semantic information to make source decisions (e.g., relative to control items, they are more likely to incorrectly judge that they saw an item that they only imagined, such as an image of a rabbit, if they had later seen its synonym, a picture of a bunny). People also use the physical characteristics of words to make source decisions (relative to control items, they are more likely to incorrectly judge that they saw a picture of an item they only imagined, such as a baseball bat, if they had seen a picture of that item's homograph, a flying bat).

In the present study, we are examining whether obligatory and non obligatory marking of first hand and non-firsthand source information influences source memory. Two experiments were conducted with native speakers of Turkish and English monolingual counterparts. Participants were first given sentences to read containing either first hand or non-firsthand source marking. Then they were then given a recognition memory test in which they were presented with previously studied sentences and nonstudied sentences. Participants were asked to judge whether each sentence was new or had been previously

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

encountered, and, in the latter case, what form (first hand or non-firsthand) the sentence had appeared in earlier. Experiment 1 compared English-speaking monolinguals with Turkish-speaking monolinguals and Turkish–English bilinguals (the latter two groups tested in Turkish) and Experiment 2 compared English monolinguals and Turkish–English bilinguals tested in Turkish and English.

Taken together, the design of the present experiments allows us to determine (1) if an evidentiality effect is present and (2) the nature of such an effect, that is, whether obligatory coding of source of evidence is associated with better source memory for first hand and non-first hand assertions alike, or with selectively better source monitoring of first hand sources only, or instead, selectively worse monitoring of non-firsthand sources. Finally, by including a bilingual sample, the present research allowed us to test whether the evidentiality effect is restricted to the language that marks evidentiality in the grammar or spills over to a language in which evidentiality is not marked in the grammar.

## Experiment 1

This experiment examined recognition and source memory accuracy for sentences containing first hand vs. non-firsthand assertions in speakers for whom coding of source of evidence is required by the grammar (Turkish) or is permitted but not required (English). If automatic linguistic coding of source affects cognition, we would expect Turkish but not English speakers to show better recognition and source memory for sentences containing first hand assertions than non-firsthand assertions.

By directly comparing Turkish speakers' performance with that of English speakers the study further allowed us to determine if Turkish speakers would also show superior memory than English speakers for first hand assertions and/or poorer memory for non-firsthand assertions. Finally, the study explored whether the size of any evidentiality effect in Turkish might be weaker if the speaker also knows a language that does not obligatorily mark source of evidence (English).

### Method

#### Participants

Twenty-six native Turkish speakers (all female), with a mean age of 26 ($SD$ = 2.05) years, 29 Turkish–English speakers (18 female), with a mean age of 29.3 years ($SD$ = 3.97), and 21 English monolingual speakers (15 female), with a mean age of 19.2 ($SD$ = 1.36) years, were recruited as participants. Turkish monolinguals were all tested in Istanbul. Although they had previously been students they were not enrolled in an educational institution at the time of testing. English speakers were recruited from the research participant pool at a large southwestern university in the US Turkish–English bilingual speakers were recruited from a Turkish students' association at the same institution. The bilinguals were all native speakers of Turkish and were pursuing graduate level study in the US. Although no systematic proficiency measures were obtained from them, they had lived in the US for a few months up to 4 years and reported feeling more comfortable speaking Turkish than English.

#### Stimuli

Per language stimuli consisted of 60 declarative, transitive sentences (30 studied and 30 nonstudied) each containing a verb in the past tense. For half of the Turkish sentences, the past tense suffix used was the first hand form (*-di*) and for the remainder the non-firsthand past tense suffix was used (*-mIş*). The particular set of sentences containing first hand vs. non-firsthand forms at initial presentation was counterbalanced across participants. For the English sentences, first hand form sentences were translation equivalents of the Turkish first hand sentences and simply contained the verb (e.g., *Mary missed her flight*); non-firsthand form sentences were translation equivalents of the Turkish non-firsthand sentences and contained an adverb such as "reportedly", "presumably", or "allegedly" before the verb. See Appendix A for sample sentences per language. The stimuli were presented in a fixed random order.

#### Procedure

Turkish speakers were tested in Turkish only, and English speakers were tested in English. In the study phase participants were presented with a list of thirty critical sentences one at a time on a computer screen. Six filler sentences were also presented (three filler sentences were presented at the beginning and the end of the study list to blunt potential serial position effects). To make the study phase more similar to a natural language situation in which participants would not be attempting to remember the sentences or the source of evidence, participants were told that the experiment was about their ability to comprehend sentences and their reading times would be measured. Participants were instructed to press a key as soon as they had read and understood each sentence. Regardless of the timing of the key press, each sentence remained on the screen for three seconds and participants were instructed to look at the sentence for the entire time it was on the screen.

After the study session, there was a 5 min rest period following which participants were given a surprise recognition and source monitoring test, which included the earlier list of 30 sentences and 30 new sentences. The test sentences' study status (studied vs. non-studied) and form (first hand vs. non-firsthand) were counterbalanced across participants. Participants were presented with two versions of the same sentence (in first hand form and in non-firsthand form) along with the word "new" and were asked to make a simultaneous recognition/source decision. That is, they had to decide if the sentence had previously been presented or not (old/new recognition and if it had whether it had been presented in first hand or non-first hand form.

#### Data analysis

We examined recognition accuracy and source monitoring accuracy. Recognition accuracy was computed by taking the difference between hit rates (correctly deciding

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

**Table 1**
Experiment 1. Mean source monitoring and recognition accuracy by group and source type.

| Group | N | Source monitoring | | Hits | | False alarms | | Mean accuracy | |
|---|---|---|---|---|---|---|---|---|---|
| | | First hand | Second hand | First hand | Second hand | First hand | Second hand | First hand | Second hand |
| English monolingual | 21 | 62.25 (22.51) | 65.96 (15.23) | 42.85 (15.92) | 44.92 (13.52) | 9.68 (9.82) | 14.76 (16.65) | 33.17 (18.3) | 30.15 (18.08) |
| Turkish monolingual | 26 | 64.6 (17.37) | 42.20 (21.560 | 44.87 (16.25) | 28.59 (14.05) | 11.02 (10.82) | 11.15 (14.38) | 33.84 (16.48) | 17.43 (15.5) |
| Turkish bilingual | 29 | 67.41 (24.4) | 39.56 (19.4) | 49.65 (16.09) | 27.81 (14.23) | 6.89 (6.6) | 7.58 (10.15) | 42.75 (15.76) | 20.22 (18.8) |

that an old item was 'old') and false alarms (incorrectly deciding that a new item was 'old'). Source monitoring accuracy was based on the ratio of correctly recognized items of a source to all given answers to the source. For each analysis (recognition and source monitoring accuracy) responses to first hand and non-firsthand items were analyzed separately. Thus, for example, hit rates for first hand items were based on the total number of "first hand" responses to the first hand items and "first hand" responses to the non-firsthand items divided by 30 (the total number of old items) multiplied by 100. False alarm rates for first hand items were based on the number of "first hand" responses to new items divided by 30 (the total number of new items) multiplied by 100. An analogous procedure was followed to obtain hit and false alarm rates for non-firsthand items. Recognition accuracy scores were obtained by taking the difference between percent hits and percent false alarms per item type. Similarly, source monitoring accuracy of first hand items was based on "first hand" responses to the first hand items divided by the total of "first hand" and "non-firsthand" responses to the first hand items, multiplied by 100. An analogous procedure was followed to obtain source monitoring accuracy of non-firsthand items.

*Results*

For all tests, significance was set at $p < .05$ and partial eta $\eta_p^2$ is reported as the measure of effect size. Mean hits and false alarms, recognition accuracy and source accuracy scores are summarized in Table 1 per group and condition.

Separate 2 (Condition: First hand vs. Non-Firsthand) × 3 (Language Group: Turkish monolingual vs. Turkish–English bilingual vs. English monolingual) ANOVAs with repeated measures on condition were conducted for recognition memory accuracy and source memory accuracy.

*Overall recognition accuracy*

A main effect of condition showed that sentences in first hand form were recognized more accurately than non-firsthand sentences [$F(1,73) = 24.38$, $p < .001$, $\eta_p^2 = .25$]. The effect of group did not reach significance [$F(2,73) = 2.66$, $p = .076$, $\eta_p^2 = .07$] but there was a significant interaction of condition by group, $F(2,73) = 3.93$, $p < .03$, $\eta_p^2 = .1$. Further analysis of the interaction showed that significantly better recognition for first hand than non-firsthand sentences characterized only the Turkish speaking groups: Turkish monolinguals, $t(25) = 3.92$, $p < .001$ (33.8% vs. 17.4%, respectively) and Turkish–English bilinguals, $t(28) = 4.98$, $p < .001$ (42.8% vs. 20.2%, respectively). English monolinguals showed no significant difference in recognition accuracy for first hand vs. non-firsthand sentences $t(20) = .49$, $p = .63$ (33.2% vs. 30.2%, respectively). Furthermore, for first hand sentences Turkish bilinguals showed significantly higher recognition accuracy than both Turkish monolinguals (42.8% vs. 33.8%, respectively), $t(53) = 2.05$, $p < .05$, and English monolinguals (42.8% vs. 33.2%, respectively), $t(48) = 1.98$, $p = .053$. For non-firsthand sentences, English monolinguals showed significantly higher recognition accuracy than Turkish monolinguals, $t(45) = -2.6$, $p < .05$ (30.2% vs. 17.4%, respectively), and Turkish bilinguals, $t(48) = -2.2$, $p < .05$ (30.2% vs. 20.2%, respectively). See Fig. 1.

*Recognition accuracy in terms of hit rates and false alarms*

To determine what was driving the observed group differences in overall recognition accuracy, analyses of variance were conducted for hit rates and false alarm rates separately by language group and condition. The analysis of false alarms showed no significant main effects, either for condition, [$F(1,73) = .02$, $p = .89$, language group, [$F(2,73) = .52$, $p = .6$,] or language by condition, [$F(2,73) = .66$, $p = .52$]. The analysis of hits showed a main effect of condition [$F(1,73) = 15.08$, $p < .001$, $\eta_p^2 = .17$], indicating higher performance on first hand than non-firsthand forms (46.14% vs. 32.81%). a main effect of language group [$F(2,73) = 6.39$, $p < .01$, $\eta_p^2 = .12$], indicating higher performance by both groups of Turkish speakers relative to English speakers, and a condition by language interaction [$F(2,73) = 5.11$, $p < .01$, $\eta_p^2 = .12$]. The interaction showed that both Turkish groups showed better hit rate accuracy for first hand than non-firsthand forms: Turkish monolinguals $t(25) = 3.3$, $p < .01$ (44.87% vs. 28.58%, respectively) and bilinguals, $t(28) = 4.33$, $p < .001$ (49.65% vs. 27.82%, respectively), but no difference between the two was found in English speakers, $t(20) = -.34$, $p = .74$ (42.85% vs. 44.92%, respectively). Furthermore, hit rates for recognition of first hand forms did not differ across the three groups, $t(45) = .43$, $p = .67$ (Turkish vs. English monolinguals), $t(48) = 1.48$, $p = .15$ (Turkish bilinguals vs. English monolinguals), $t(53) = 1.1$, $p = .28$ (both Turkish groups). However, English speakers showed an advantage over both Turkish groups in percent hit rates for non-firsthand forms: $t(45) = -4.03$, $p < .001$, for Turkish monolinguals vs. English (28.58% vs. 44.92%), $t(48) = -4.28$, $p < .001$, for Turkish bilinguals (27.81%) and English speakers); the two Turkish groups did not differ from each other, $t(53) = -.2$, $p = .84$.

To summarize, the nature of the group by condition interaction observed in recognition memory was mainly driven by group differences in hit rates for non-firsthand forms. That is, English speakers were significantly better at detecting non-firsthand studied forms than were either

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

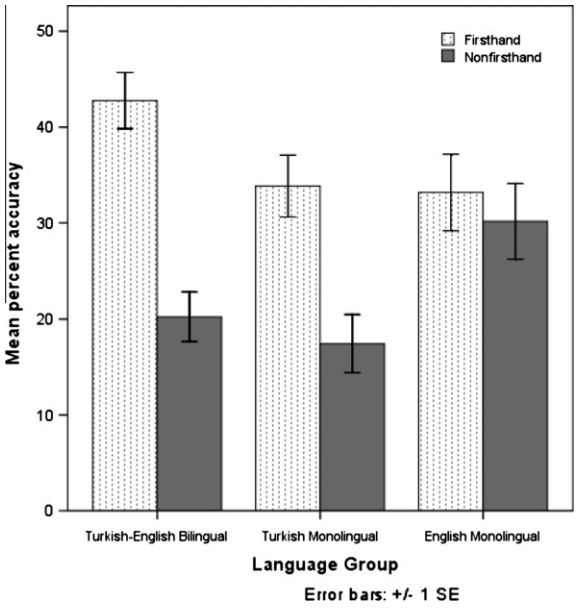


**Fig. 1.** Mean recognition accuracy as a function of type of source and group (Experiment 1).

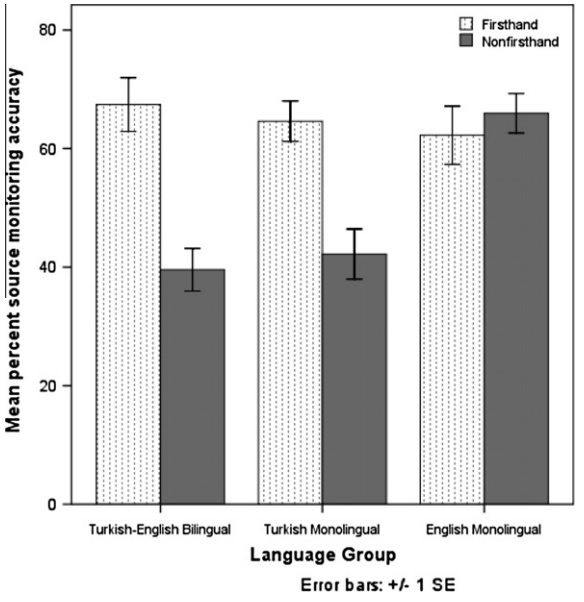


**Fig. 2.** Mean percent source monitoring accuracy as a function of type of source and group (Experiment 1).

of the Turkish groups. Moreover, both Turkish groups were better at detecting first hand than non-firsthand studied forms, whereas the performance of English speakers did not differ across the two types.

*Source monitoring accuracy*

Analysis of source memory data revealed a main effect of condition [$F(1,73) = 14.56$, $p < .001$, $\eta_p^2 = .17$] showing better source memory for first hand sentences than non-firsthand sentences. In addition, there was a main effect of group [$F(2,73) = 7.81$, $p < .001$, $\eta_p^2 = .18$], showing better overall source memory for English monolinguals compared to both of the Turkish groups, and an interaction of group by condition [$F(2,73) = 5.34$, $p < .01$, $\eta_p^2 = .13$]. Further analysis of the interaction revealed that superior memory for first hand vs. non-firsthand sources characterized the performance of Turkish monolinguals (64.6% vs. 42.2%, respectively), $t(25) = 3.25$, $p < .01$, and Turkish bilinguals, (67.4% vs. 39.6%, respectively), $t(28) = 4$, $p < .001$, but not English monolinguals, who showed no difference between the two types (62.2% vs. 66%, respectively). Furthermore, although the groups did not differ in their source memory for first hand forms, English monolinguals showed significantly better source memory on non-firsthand forms as compared to that of Turkish monolinguals, $t(45) = -4.26$, $p < .001$ (66% vs. 42.2%, respectively), or Turkish bilinguals, $t(48) = -5.18$, $p < .001$ (66% vs. 39.6%, respectively). See Fig. 2.

*Discussion*

The primary aim of this experiment was to test whether recognition and/or source memory would be better for first hand than non-firsthand sources of assertions in users of languages in which source is obligatorily coded. Our findings showed clear support for this outcome: Turkish speakers showed better recognition memory and source memory for first hand than non-firsthand sources whereas English speakers were equally good at recognizing and monitoring the source of information regardless of the source. As such, our findings support the presence of an evidentiality effect and suggest that obligatory coding of source of evidence leads to first hand sources being prioritized over second hand sources in memory. Consistent with this interpretation is the additional finding that Turkish speakers tended to misremember non-firsthand sources as first hand ones. This was not the case for English speakers, who showed no difference in source memory for first hand and non-firsthand sources.

A second outcome of interest concerned the relative memory performance of Turkish and English speakers. Turkish speakers were no different than English speakers in their accuracy of source memory for first hand sentence forms but were significantly worse than English speakers in recognition of non-firsthand sources. Moreover, Turkish speakers were more likely to misremember non-first-hand forms as first hand forms. This finding was previously noted in a developmental context by Aksu-Koc et al. (2009; see also Aydın, 2012). Taken together, the results suggest that first hand information may become the default form of memory representation for Turkish speakers.

A third question of interest in this study was whether recognition and/or source memory in native users of a language in which source is obligatorily coded is affected by additional knowledge of a language in which it is not so coded. To test this we examined whether knowledge of English reduces the size of the source memory effect in Turkish–English speakers relative to that in Turkish monolinguals. Our results showed that knowing English as a second language did *not* alter the pattern observed for the speakers' native language, Turkish. Bilingual Turkish–

English users behaved no differently than monolingual Turkish speakers; both groups showed poorer source memory for non-firsthand than first hand accounts in Turkish. These findings suggest that, as long as one of the languages has obligatory coding of source of evidence, knowing another language that does not grammatically mark source does not make a difference in the relative weighting of first hand and second hand information in memory.

In fact, the only group difference that emerged in the comparison of Turkish–English bilinguals and Turkish monolinguals was in overall recognition accuracy: Turkish–English bilinguals recognized more "old" sentences than Turkish monolinguals (a similar effect approached significance in the comparison with English monolinguals). We are inclined to interpret this as a sampling effect, given that the bilinguals were drawn from a pool of mainly graduate students whereas the Turkish monolinguals, although generally close in age to the bilinguals, were not currently enrolled in higher education, and the English monolinguals were predominantly undergraduate students. What is important to remember is that there was no difference between the bilinguals and Turkish monolinguals in source monitoring accuracy. Thus, knowing English did not improve the relative accuracy of Turkish non-firsthand source information in bilinguals. In order to eliminate the sampling effect, the best way would be to conduct a within subject design experiment which was the next step we took.

Although no support was found for a possible influence of L2 (English) on L1 (Turkish), it remains to be determined if there is an influence of L1 on L2, that is, whether the evidentiality effect found in Turkish may carry over to affect source memory in English. Since the bilinguals in Experiment 1 were tested only in Turkish the results from this study do not speak to this issue. To examine this potential effect of speaking Turkish for memory for source in English, Experiment 2 tested Turkish–English bilingual participants in both Turkish and English. Secondly, Experiment 2 compared the performance of different subgroups of Turkish–English bilinguals to determine if the pattern observed for Turkish extends to English the more recently English was learned. Finally, the experiment tested whether the poorer memory for non-firsthand information noted previously for Turkish stimuli presented under incidental memory conditions (Experiment 1) may be altered if participants are given the opportunity to prepare for the memory test. To address this issue we compared the performance of bilinguals by test block, since in the first block the memory test was unexpected but by the second block it was anticipated.

## Experiment 2

### Method

#### Participants

Twenty-seven Turkish–English bilingual (25 female) and 23 English monolingual speakers (12 female) participated in the study. None had participated in the previous experiment. The bilinguals ranged in age from 19 to 29 years, with a mean age of 21.93 ($SD$ = 5.95); the mean age of the monolinguals was 18.39 years ($SD$ = 0.58). English monolingual speakers were recruited from the subject pool of a large southwestern university in the US. Bilingual speakers were recruited from a Turkish students' association of the same institution and from Turkish cultural centers in two metropolitan cities in the US.

A detailed language questionnaire administered to the bilinguals indicated that all bilingual participants were native speakers of Turkish and had learned English as a second language either before the age of 12 ($n$ = 16 early bilinguals) or after the age of 12 ($n$ = 11 late bilinguals). There was no difference in self-reported mean L2 reading or writing proficiency between the early and late bilinguals (5.97 vs. 5.18, respectively, on a 7-point scale); however, early bilinguals reported higher speaking proficiency in English relative to late bilinguals (6.12 vs. 4.64). Moreover, 51.8% of the bilinguals (including 10 early and 4 late bilinguals) indicated that their overall knowledge of English was as good as or better than their Turkish whereas the remainder (6 early and 7 late) reported that their English was worse than their Turkish.

#### Stimuli and procedure

A new set of 24 transitive, declarative sentences containing a past tense verb (and 24 unstudied sentences, similarly constructed, presented at test) were selected in Turkish with 24 translation equivalent counterparts in English (and 24 unstudied test sentences in English). The design of the sentences was similar to that in the previous experiment, but the number of studied sentences was reduced from 30 to 24 to avoid fatigue given that bilingual participants in the present experiment were to perform the task in both languages. Per participant, half of the Turkish sentences were presented in the first hand past tense form (the *-di* suffix) and half were presented in the non-firsthand past tense form (the *-mIş* suffix). As in the previous experiment, the English counterparts of the latter form contained an adverb preceding the verb (e.g., *reportedly, presumably, apparently*, etc.). Stimuli were blocked by language with language order counterbalanced. Per language, first hand and non-firsthand sentences were presented in a fixed random order. Also for each language the particular sentences chosen in the study phase to be in first hand vs. non-firsthand form were counterbalanced across participants.

Participants were tested individually in a single session divided into two language blocks. In the initial phase of block 1, participants studied twenty-four sentences in a given language (Turkish for half the participants, English for the other half), containing an equal number of sentences in first hand and first hand form. Each sentence was presented for 3 s on a computer screen. As in Experiment 1, participants studied the sentences under incidental study instructions. After reading the sentences, a 5 min rest period was given. Next, participants were given a surprise recognition test in which the 24 studied sentences were presented again, in both first hand and non-firsthand form, together with 24 new sentences also shown in each form. Participants were to decide whether or not each sentence

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

**Table 2**
Experiment 2: mean percent source monitoring and recognition accuracy by group and source type.

| Language | N | Source monitoring | | Hits | | False alarms | | Mean accuracy | |
|---|---|---|---|---|---|---|---|---|---|
| | | First hand | Second hand | First hand | Second hand | First hand | Second hand | First hand | Second hand |
| English | 27 | 63.61 (16.4) | 67.53 (21.68) | 39.35 (11.57) | 43.67 (18.75) | 8.17 (9.42) | 12.19 (16.38) | 31.17 (13.04) | 31.48 (21.72) |
| Turkish | | 64.7 (15.97) | 42.35 (22.04) | 48.30 (14.35) | 32.25 (14.13) | 10.03 (11.21) | 8.95 (13.05) | 38.27 (14.48) | 23.3 (16.87) |
| English Monolinguals | 23 | 60.27 (16.97) | 69.39 (17.43) | 38.95 (13.63) | 46.01 (12.84) | 12.68 (14.68) | 18.11 (13.85) | 26.27 (17.29) | 27.9 (13.61) |

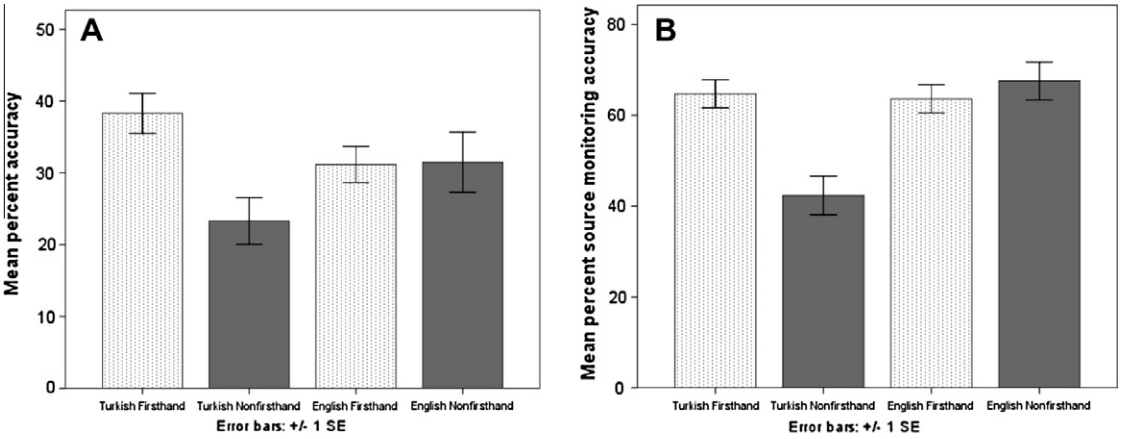


**Fig. 3.** Mean recognition accuracy and source monitoring accuracy of Turkish–English bilinguals by language and condition in Experiment 2. Fig. 3A shows source monitoring accuracy by condition and language. Fig. 3B shows mean recognition accuracy by condition and language.

had been presented earlier (“old” vs. “new” judgments) and if it had whether it had been shown in first hand or non-firsthand form (source judgment).

Upon completing the memory test participants filled out a language background questionnaire and then were given the second block. Stimuli were now presented in the other language (i.e., if Turkish was presented in the previous session, English was presented in the second session). Although the recognition tests presented at the end of the study phase of the first block had been unexpected, participants were presumably no longer surprised on getting the recognition test at the end of the second block.

*Data analysis*

As in the previous experiment separate analyses of variance were conducted for recognition accuracy and source monitoring accuracy. However, to examine the effect, if any, of knowing or not knowing that a memory test was forthcoming, recognition and source memory performance were also examined as a function of language order. This variable served as a proxy variable for expecting or not expecting a memory test. Thus, those who began the study phase in Turkish were not expecting the subsequent memory test in Turkish but were presumably expecting it later when they got it in English; similarly, those who got the English block first were not expecting a memory test in English but were presumably expecting it by the time they got it in Turkish. There were about equal numbers of early and late bilinguals assigned to each language order condition but due to the small numbers per condition this variable was not entered into the language order analyses. Finally, to ascertain how bilinguals fared relative to English monolinguals (who were only tested once), the monolingual/bilingual comparison was based only on bilinguals who got English first, so that both groups were comparable in not expecting a memory test.

*Results and discussion*

Recognition accuracy (including mean hits and false alarms) and source accuracy scores per group by language and condition are summarized in Table 2.

*Performance of Turkish–English bilinguals on Turkish vs. English*

*Recognition accuracy*

Overall recognition accuracy was first analyzed without regard to language testing order. There was a condition main effect [$F(1,26) = 5.54$, $p < .05$, $\eta_p^2 = .18$] and a condition by language interaction [$F(1,26) = 4.31$, $p < .05$, $\eta_p^2 = .14$]. See Fig. 3. In Turkish, participants recognized first hand sentences better than non-firsthand sentences, $t(26) = 3.32$, $p < .01$ (38.27% vs. 23.3%, respectively); however, in English they showed no difference between the two, $t(26) = -.06$, $p = .95$ (31.17% vs. 31.48%, respectively). Moreover, bilinguals recognized non-firsthand sentences significantly better in English than in Turkish, $t(26) = 1.96$, $p = .06$ (31.5% vs. 23.3%, respectively).

*Source monitoring accuracy*

Analysis of source monitoring accuracy for bilinguals considered as a whole revealed a main effect of condition

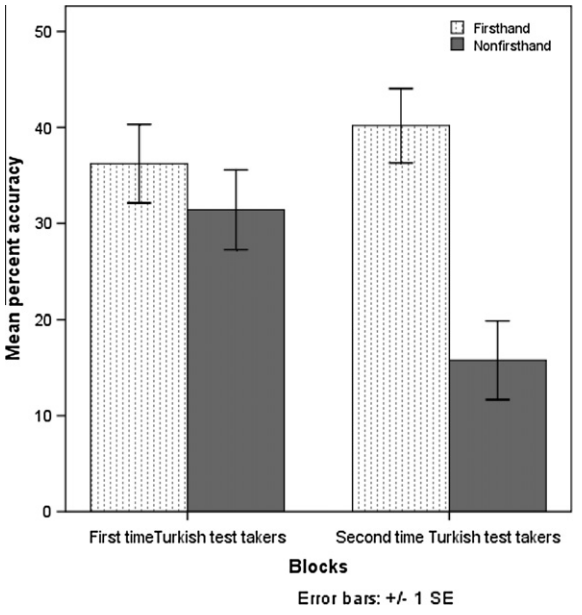


**Fig. 4.** Mean recognition accuracy of Turkish–English bilinguals on Turkish by source type and language block (Experiment 2).

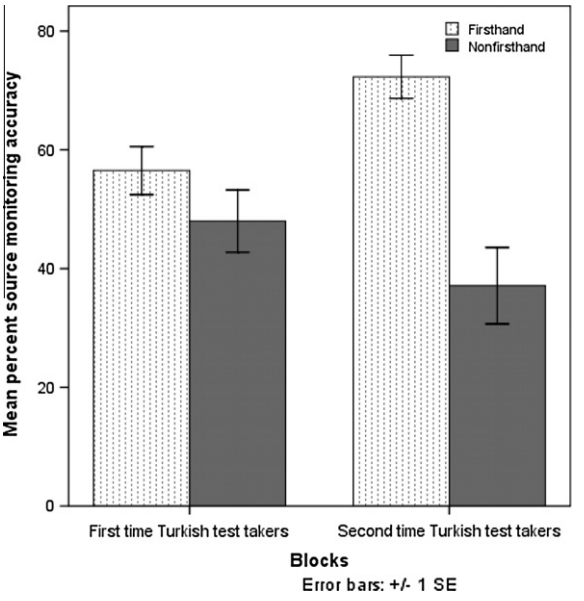


**Fig. 5.** Mean source monitoring accuracy in Turkish–English bilinguals on Turkish by source type and language block.

[$F(1,26) = 5.8$, $p < .05$, $\eta_p^2 = .18$], a main effect of language [$F(1,26) = 15.19$, $p < .01$, $\eta_p^2 = .40$] and a condition by language interaction [$F(1,26) = 8.03$, $p < .01$, $\eta_p^2 = .24$]. Bilinguals were significantly better at monitoring first hand than non-firsthand sources in Turkish only, $t(26) = 3.93$, $p < .001$ (64.7% vs. 42.35%). Moreover, non-firsthand sources were monitored better in English than Turkish, $t(26) = 4.22$, $p < .001$ (67.53% vs. 42.35%); there was no difference across languages in monitoring first hand sources, $t(26) = .21$, $p = .83$ (63.61% vs. 64.7% for English vs. Turkish, respectively).

The above pattern of findings using a within-subjects design essentially replicates what had been observed in the previous experiment in a between-subjects comparison, showing that recognition and source memory performance were significantly better for first hand than non-firsthand sentences in Turkish only. However, given that the present design allowed for a separate examination of performance under incidental memory vs. non-incidental memory testing conditions, we conducted a set of additional analyses to explore the effects of this manipulation.

### *Performance of Turkish–English bilinguals on Turkish by test block*

#### *Recognition accuracy*

When performance on Turkish was compared across participants who received the Turkish test in the first block and those who received it in the second block (and were presumably no longer surprised by getting a memory test) there was a main effect of condition [$F(1,25) = 12.26$, $p = .002$, $\eta_p^2 = .33$] and a condition by block interaction, [$F(1,25) = 5.52$, $p = .03$, $\eta_p^2 = .18$] in overall recognition accuracy. Although first hand sentences were recognized better than non-firsthand sentences, the condition effect was actually larger in block 2, $t(13) = 4.73$, $p < .001$ (40.2% vs. 15.8%, respectively). Specifically, recognition of non-firsthand sentences worsened for those who received the Turkish test in the second block than those who got it in the first block, $t(25) = -2.67$, $p < .05$ (15.8% vs. 31.4%) even though the memory test was not a surprise for block 2 participants. See Fig. 4.

#### *Source monitoring accuracy*

When a similar analysis by block was conducted for Turkish source monitoring accuracy, there was again a main effect of condition, [$F(1,25) = 18.02$, $p = .001$, $\eta_p^2 = .42$], and a condition by block interaction, [$F(1,25) = 6.71$, $p = .016$, $\eta_p^2 = .21$]. The difference in source memory for first hand vs. non-firsthand sentences was actually significantly larger for second time than first time Turkish test takers, $t(12) = 1.05$, $p = .31$ for Block 1 and $t(13) = 5.42$, $p < .001$ for Block 2 (Turkish Block 1: 56.5% vs. 48%; Turkish Block 2: 72.3% vs. 37.1%). Analysis of this interaction further showed that memory for first hand sources was significantly boosted when the memory test was expected than when it was not expected, $t(25) = 2.91$, $p < .01$ (72.3% vs. 56.5% for Block 2 vs. Block 1); there was no comparable elevation of source memory for non-firsthand sources across test blocks. See Fig. 5.

### *Performance of Turkish–English bilinguals in English by test block*

Unlike the findings for Turkish, there was no change in recognition or source memory in English as a function of whether the English block was presented first or second. The block effect (31.2% vs. 31.5% for first hand vs. non-firsthand), $F(1,25) = .31$, $p = .58$, condition effect (30.1% vs. 32.6% for Block 1 vs. Block 2), $F(1,25) = .002$, $p = .96$, and

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

condition by block interaction, $F(1,25) = .17$, $p = .68$, were not significant in the recognition memory analysis. For the source analysis also the effects of block (67.6% vs. 63.3% for Block 1 and Block 2), $F(1,25) = 1.28$, $p = 27$, condition (63.6% vs. 67.5% for first hand vs. non-firsthand), $F(1,25) = .33$, $p = .56$, and condition by block, $F(1,25) = 2.65$, $p = .12$, were not significant. Thus, unlike the case for Turkish, performance on English was largely unaffected by knowing that there would be a memory test.

### *Performance on English by Turkish–English bilinguals and English monolinguals*

To compare the performance of Turkish–English bilingual speakers with that of native English speakers, a 2 (Group: Bilingual vs. Monolingual) × 2 (Condition: First hand vs. Non-firsthand) ANOVA with repeated measures on condition was conducted separately for mean recognition accuracy and source monitoring accuracy. Only bilinguals who got the English block first were included in this analysis.

#### *Recognition accuracy*

There was no significant effect of group or condition nor was there an interaction. Recognition accuracy by Turkish–English bilinguals and English monolinguals for first hand English sentence forms was no different from that for non-firsthand English forms (bilinguals: 31.7% vs. 31.4%, respectively; monolinguals: 26.3% vs. 27.9%, respectively).

#### *Source monitoring accuracy*

The source memory analysis yielded a main effect of condition [$F(1,35) = 6.66$, $p = .014$, $\eta_p^2 = .16$]. The direction of this effect was different from that previously observed for Turkish: source memory for non-firsthand sentences was actually *better* than that for first hand sentences (71.9% vs. 60.6%). This held equally for monolinguals and bilinguals.

In summary, the results show that performance of Turkish–English bilinguals on English recognition accuracy and source memory was very different from their performance in Turkish but no different from that of English monolinguals. This would suggest that there was no carryover effect of the Turkish pattern. However, this analysis was based only on a subset of the participants, those who got English first in the experiment, and did not subdivide bilinguals on the basis of age of acquisition of English.

The contribution of age of L2 (English) acquisition was examined in additional analyses. For these analyses all early and late bilinguals were included (not just those who got English first in the experiment).

#### *Relationship between age of $L_2$ acquisition and source monitoring accuracy*

A correlational analysis revealed a significant negative relationship between age of acquisition of English and accuracy of monitoring the source of non-firsthand English sentences [$r(25) = -.38$, $p < .05$]. See Fig. 6. That is, the later participants had started learning English the less accurate they were in recognizing English non-firsthand sources. Put differently, the performance profiles of bilinguals on English more closely resembled those on Turkish the more recently the bilinguals had learned English.

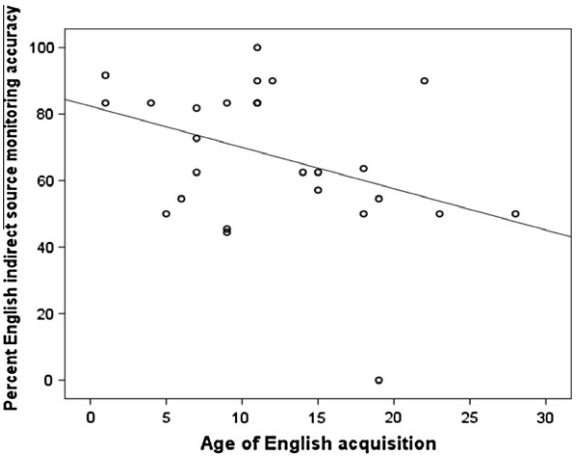


**Fig. 6.** Correlation between age of English acquisition and source monitoring accuracy of English non-firsthand source sentences (Experiment 2).

This effect was confirmed in a subsequent independent samples *t*-test that compared early bilinguals ($n = 16$) with late bilinguals ($n = 11$) on English vs. Turkish non-firsthand source memory. Whereas no group differences were found on Turkish source memory, on English early bilinguals monitored non-firsthand sources better than late bilinguals (74% vs. 57%, respectively); [$t(25) = 2.17$, $p < .05$]. Thus, even when responding to stimuli in English, late bilinguals behaved as they did in responding to Turkish stimuli, showing poorer source memory for non-firsthand information relative to early bilinguals. This finding lends support to the notion that processing strategies developed in the first language may carry over to a later acquired language.

## General discussion

Previous authors have suggested that obligatory marking of evidentiality may attune speakers to source information (e.g., Boroditsky, 2003; Gerrig & Banaji, 1994; Robinson, 2009; Slobin, 2003) but there had been no previous test of this notion in adults. Our findings thus constitute the first empirical investigation of an evidentiality effect in memory for source as a function of the status of evidential marking in the language.

Across two experiments our findings demonstrate support for an evidentiality effect in that within-language comparisons consistently showed that memory for first hand sources was better than that for non-firsthand sources in Turkish speakers. On the other hand English speakers showed equally accurate memory for both first and non-firsthand sources. Secondly, cross-language comparisons conducted between-subjects (Experiment 1) and within-subjects (Experiment 2), showed that whereas there was no difference across English and Turkish in memory for firsthand sources, memory for non-firsthand sources was poorer in Turkish than in English. Whereas we had initially expected that obligatory marking of information source may be associated with better overall mem-

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

ory for source, our findings suggest that this was not the case. Rather, what our results suggest is that habitually having to mark whether an assertion is based on firsthand or non-firsthand knowledge leads to a privileging in memory of first hand sources and a discounting of non-firsthand sources. Non-firsthand sources may be discounted because one cannot be as certain of their truth value. As such they may be perceived as less reliable in epistemic value compared to first hand sources. Such an interpretation is consistent with developmental research that Turkish speaking children are more likely to trust information presented in first hand form than non-firsthand form (Öztürk & Papafragou, 2005; see also Fitneva, 2001, for a similar finding with Bulgarian-speaking children). If Turkish adults similarly implicitly believe first hand accounts to be more trustworthy, they may pay less attention to second hand accounts and thus not encode them to the same degree as first hand sources, and thereby show poorer retrieval of non-firsthand sources. An "attentional neglect" account for non-firsthand information would explain why even when participants anticipated a source memory test and thus could prepare for it (Experiment 2) did not show improvement for such sources, but did improve their recall of first hand sources. More generally, our findings with Turkish fit with a number of findings noted in the memory literature in English that memory accuracy may be affected by the linguistic framing of an event, such as a subtle way of creating uncertainty about whether the event occurred at all by use of a definite vs. an indefinite article, e.g., 'Did you see *the* broken headlight? vs. 'Did you see *a* broken headlight?' (Strack & Bless, 1994).

Taken together, these results show a robust effect of obligatory encoding of evidential marking on source monitoring accuracy. Further work is needed to establish whether the pattern of primacy of coding of first hand source information we observed for Turkish generalizes to other languages in which evidentiality is also marked in the grammar, or whether it reflects the particular type of evidential distinctions made in Turkish. Further work would also benefit from extending investigation of source memory to related phenomena such as the misinformation effect. By examining how use of the first hand vs. non-firsthand marker may interact with other experimental manipulations of source credibility to affect memory accuracy would provide useful insights into the relative weighting of grammatical and contextual influences on memory accuracy.

Another important finding from our research relates to the impact of obligatory marking on performance in a second language in which source is not obligatorily marked. Our results showed that the pattern carries over to a language in which evidentiality is not marked in the grammar, particularly if the second language was acquired relatively late. Late bilinguals in English showed worse source memory for non-firsthand information relative to that shown by early bilinguals. This pattern is similar to that observed in the domain of perceptions of agency (see Cunningham et al., 2010) where the "Spanish" pattern of perceiving less agency in accidental occurrences extended to English in bilinguals (see Bassetti & Cook, 2010, for further discussion). Again, it remains to be determined if this pattern is specific to Turkish–English speakers or would be found in other language pairs as well.

There were certain limitations of the present research. One relates to how we operationalized evidential marking in English, i.e., through use of modal adverbs. It could be argued that our choice of this type of marking of non-firsthand assertions made the English non-firsthand sentences more perceptually salient than the Turkish ones and that had we chosen some more subtle form of non-firsthand marking in English (e.g., modal auxiliaries, such as "must have") the superior performance of English over Turkish may not have emerged. However, if salience of marking *per se* was responsible for the memory performance in the present study, we should not have gotten a difference in response to first hand and non-firsthand markers in Turkish as these are arguably equally salient perceptually (both are monosyllabic suffixes). Moreover, based on a salience argument we should have obtained better memory for non-firsthand than first hand sentence forms in English. But, instead, there was no effect of evidential marking on memory for English sentences. Nonetheless, to test the generalizability of the effect we observed in English future research should explore more systematically the full range of lexical marking of indirectness in English by also using modal auxiliaries or mental state predicates such as "it seems" or "I think that…".

A second issue bears on how to interpret our findings of poorer memory for non-firsthand vs. first hand sources given the fact that the Turkish non-firsthand source marker does not have a single meaning but can reflect hearsay, inference, or surprise depending on the context. The poorer memory for non-firsthand sources observed in our study may in part reflect a difference arising from the one to one mapping of first hand information in Turkish vs. the one to many mapping of non-firsthand information in addition to a greater perceived reliability of first hand information. To test this interpretation of the nature of the difference between first hand and non-firsthand information, future research should be done with speakers of a language that uses different markers for hearsay, inference and expectancy violation, rather than a single marker. If the pattern observed here is replicated in speakers of such a language (e.g., Korean or Japanese) then one can more confidently conclude that it is the greater salience and/or conceptual prominence of first hand assertions rather than the greater ambiguity of non-firsthand assertions (due to their sharing a single marker for several types of non-firsthand sources) that underlies the effect observed.

Finally, a third potential issue that limits the scope of the present findings arises from the fact that the stimuli were presented visually rather than auditorily. By presenting the sentences visually we were not able to study prosodic effects on the interpretation and representation of narrated events. In further investigations it would be useful to compare performance under conditions of auditory vs. visual presentation, as well as add discourse cues to broaden the scope of the investigation. In addition, a potentially fruitful follow-up to the present research would be to explore the time course of processing utterances with first hand vs. non-firsthand marking. This approach would allow a way of testing whether first hand

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004

assertions (particularly for speakers of languages such as Turkish) have a processing advantage and whether non-firsthand representations are computed on an as needed basis relying on contextual cues.

If it is indeed the case that Turkish speakers are worse than English speakers in memory for non-firsthand information one might expect to find differences in performance between these groups in various other contexts, for example, in the incidence of false memories of an event or the attribution of post-event information. In eyewitness testimony cases, post-event information is more likely to have non-firsthand sources, such as hearing from other witnesses, or from television or newspaper accounts. Turkish speaking eyewitnesses could conceivably be more likely to misattribute post-event information as first hand information and thus be more susceptible to misinformation, relative to English speakers. If that is the case one practical application of the present research might be to develop interventions to make Turkish speakers more attentive to non-firsthand sources of information (cf. Aydin & Ceci, 2009).

To conclude, all languages allow their users to mark source of knowledge. The way source of evidence is conveyed, however, differs from language to language. Some languages encode source in their grammar, which makes source coding obligatory, whereas other languages encode source of knowledge in the lexicon, which makes it optional. The present research demonstrated that obligatory coding source of knowledge affects retention. In either system first hand sources have robust representations. However, obligatorily coded non-firsthand sources, at least in the case of Turkish, appear to be more fragile than optionally coded ones.

## Acknowledgments

This research was supported by a Collaborative Seed Grant from the Psychology Department at Texas A&M University, by a College of Liberal Arts Dissertation Enhancement Award to the first author, and by a National Institute on Aging Grant RO1 AG039502 awarded to the last author. Portions of this work were presented at the 2010 annual meeting of the Cognitive Science Society, Portland.

## A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jml.2013.03.004.

## References

Aikhenvald, A. Y. (2004). *Evidentiality*. New York, NY: Oxford University Press.

Aksu-Koç, A., Ögel-Balaban, H., & Alp, I. E. (2009). Evidentials and source knowledge in Turkish. In S. A. Fitneva, & T. Matsui (Eds.), *Evidentiality: A window into cognitive development. New directions for child and adolescent development* (Vol. 125, pp. 13–28).

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

Aydin, C. (2012, May). Does grammatical marking of source affect memory for sentence content? In *Poster presented at annual meeting of the Association for Psychological Science*, Chicago.

Aydin, C., & Ceci, S. J. (2009). Evidentiality and suggestibility: A new research venue. In S. A. Fitneva & T. Matsui (Eds.). *Evidentiality: A window into language and cognitive development, new directions for child and adolescent development* (Vol. 125, pp. 79–93). San Francisco: Jossey-Bass.

Aydin, C., & Ceci, S. J. (2009). Evidentiality and suggestibility: A new research venue. In S. A. Fitneva & T. Matsui (Eds.). *Evidentiality: A window into language and cognitive development, new directions for child and adolescent development* (Vol. 125, pp. 79–93). San Francisco: Jossey-Bass.

Bassetti, B., & Cook, V. (2010). Language and cognition: The second language user. In V. Cook & B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 143–190). Oxford: Psychology Press.

Boroditsky, L. (2003). Linguistic relativity. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 917–921). London: MacMillan.

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

Cunningham, D., Vaid, J., & Chen, H. C. (2010). Yo no lo tiré, se cayó solito. 'I did not throw it, it just fell down': Interpreting and recounting accidental events in Spanish and English. In V. Cook & B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 407–429). New York, NY: Psychology Press.

DeLancey, S. (2001). The mirative and evidentiality. *Journal of Pragmatics, 33*, 369–382.

Drummey, A., & Newcombe, N. (2002). Developmental changes in source memory. *Developmental Science, 5*, 502–513.

Fausey, C. M., & Boroditsky, L. (2011). Who dunnit? Cross-linguistic differences in eyewitness memory. *Psychonomic Bulletin & Review, 18*, 150–157.

Fausey, C. M., & Matlock, T. (2011). Can grammar win elections? *Political Psychology, 32*, 563–574.

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

Francis, W. W. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin, 125*, 193–222.

Gentner, D., & Goldin-Meadow, S. (2003). Whither Whorf? In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind. Advances in the study of language and cognition* (pp. 3–14). Cambridge, MA: M.I.T. Press.

Geraci, L., & Franklin, N. (2004). The influence of linguistic labels on source monitoring decisions. *Memory, 12*, 571–585.

Gerrig, R. J., & Banaji, M. R. (1994). Language and thought. In R. J. Sternberg (Ed.), *Thinking and problem solving: Handbook of perception and cognition* (pp. 233–261). San Diego, CA: Academic.

Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge: Cambridge University Press.

Henkel, L. A., Franklin, N., & Johnson, M. K. (2000). Cross-modal source monitoring confusions between perceived and imagined events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 321–335.

January, D., & Kako, E. (2007). Re-evaluating evidence for linguistic relativity: Reply to Boroditsky (2001). *Cognition, 104*, 417–426.

Johanson, L. (2000). Turkic indirectives. In L. Johanson & B. Utas (Eds.), *Evidentials: Turkic, Iranian and neighboring languages* (pp. 61–87). Berlin: Mouton de Gruyter.

Johnson, M. K., De Leonardis, D. M., Hashtroudi, S., & Ferguson, S. A. (1995). Aging and single versus multiple cues in source monitoring. *Psychology and Aging, 10*, 507–517.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 2–28.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.

Lazard, G. (2001). On the grammaticalization of evidentiality. *Journal of Pragmatics, 33*, 359–367.

Lucy, J. (1996). The scope of linguistic relativity: An analysis and review of empirical research. In J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 137–169). Cambridge University Press.

Markham, R., Howie, P., & Hlavacek, S. (1999). Reality monitoring in auditory and visual modalities: Developmental trends and effects of cross-modal imagery. *Journal of Experimental Child Psychology, 72*, 51–70.

Nuyts, J. (2001). Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics, 33*, 383–400.

Ögel, H. (2007). *Developments in source monitoring and linguistic encoding of source*. Unpublished M.A. thesis. Istanbul, Turkey: Bogazici University.

Öztürk, Ö. (2008). *Acquisition of evidentiality and source monitoring*. Ph.D. dissertation. University of Delaware.

Öztürk, Ö., & Papafragou, A. (2005). The acquisition of evidentiality in Turkish. *University of Pennsylvania Working Papers in Linguistics, 11*.

Palmer, F. R. (2001). *Mood and modality*. Cambridge: Cambridge University Press.

Papafragou, A., Li, P., Choi, Y., & Han, C. (2007). Evidentiality in language and cognition. *Cognition, 103*, 253–299.

Plungian, V. A. (2001). The place of evidentiality within the universal grammatical space. *Journal of Pragmatics, 33*, 349–357.

Robinson, E. J. (2009). Commentary: What we can learn from research on evidentials. In S. A. Fitneva, T. Matsui (Eds.), *Evidentiality: A window into language and cognitive development. New Directions for Child and Adolescent Development* 125, 95–103.

Roediger, H. L., III, & Geraci, L. (2007). Aging and the misinformation effect: A neuropsychological analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 321–334.

Roediger, H. L., IIII, & McDermott, K. B. (2000). Distortions in memory. In F. I. M. Craik & E. Tulving (Eds.), *The Oxford handbook of memory* (pp. 149–164). Oxford, England: Oxford University Press.

Slobin, D. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 157–192). Cambridge: M.I.T. Press.

Slobin, D. I. (1996). From "thought and language" to thinking for speaking". In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). New York: Cambridge University Press.

Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language, 33*, 203–217.

Whorf, B. (1956). *Language, thought, and reality*. Cambridge, MA: The M.I.T. Press.

Wolff, P., & Holmes, K. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*, 253–265. http://dx.doi.org/10.1002/wcs.104.

Please cite this article in press as: Tosun, S., et al. Does obligatory linguistic marking of source of evidence affect source memory? A Turkish/English investigation. *Journal of Memory and Language* (2013), http://dx.doi.org/10.1016/j.jml.2013.03.004