This is a contribution from *Social Environment and Cognition in Language Development. Studies in honor of Ayhan Aksu-Koç*.
Edited by F. Nihan Ketrez, Aylin C. Küntay, Şeyda Özçalışkan and Aslı Özyürek.

CHAPTER 8

Evidentiality, questions and the reflection principle in Tibetan

What do children learn when they learn about evidentiality?

Jill de Villiers and Jay L. Garfield
Smith College, U.S.A.

Evidentials fall in the borderland between traditional semantics and pragmatics. A situation semantics for evidentials helps to explain their puzzling developmental pathway in children. Drawing on our work in Tibetan, we argue that there is no necessity for a child to master Theory of Mind, that is, awareness of others' mental states, in order to make or to understand assertions that carry evidential force. The meaning of evidentials does not make reference to states of knowledge of persons, but rather encodes relations between discourse, evidence and evaluation situations. On the other hand, when a Tibetan speaker asks a question, the form of the evidential used in the question must anticipate the kind of knowledge the interlocutor can access in reply. Full mastery of questions in Tibetan-speaking children does require attention to and representation of others' states of knowledge and belief.

Keywords: evidentials, Tibetan, questions, Theory of Mind, semantics

It is a pleasure to contribute a paper on evidentials to the volume celebrating the work of Ayhan Aksu-Koç, whose work on this topic in Turkish acquisition was groundbreaking. We have shared our keen interest in the development of language about the mind when we have met at conferences across the world, in meetings that have never failed to stimulate.

1. Introduction

Evidential markings are generally taken to encode the type of evidence a speaker has for a statement. About a quarter of the world's languages mark evidentiality grammatically (Aikhenvald 2004). Evidentiality has attracted attention in developmental

DOI 10.1075/tilar.21.08vil

psycholinguistics because its developmental track has the potential to reveal when children can attend to the sources of belief in others. The mastery of evidentiality seems to require understanding how someone else knows something and therefore taking the epistemic perspective of others. This capacity in turn seems to require understanding others' states of mind.

Ayhan Aksu-Koç (Aksu 1978; Aksu-Koç 1988) pioneered the study of children's acquisition of evidentials. She studied the development of the markers in Turkish that differentiate direct and indirect knowledge, and how these interact with tense. Her insight that this development might connect in important ways to the child's developing Theory of Mind (ToM) interested the child development research community in evidentials. She recognized that analyzing spontaneous discourse is only one method for finding out what children know. Subsequent researchers owe her a great debt of gratitude.

The acquisition of mental state verbs has been tightly linked to the cognitive developments that support the acquisition of ToM in the preschool years (Astington & Baird 2005; Bartsch & Wellman 1995; de Villiers & de Villiers 2000; Astington 2000; Milligan, Astington & Dack 2007; Shatz, Wellman & Silber 1983). The acquisition of evidentials promises to be revealing for the same reasons: if these morphemes encode information about abstract mental states, then children must have complex ToM skills in order to understand and to use these morphemes correctly.

The development of the ability to appreciate another's point of view or beliefs is complex. Infant studies using eye gaze, preferential looking and offers of assistance suggest that very young children are differentially affected by whether a person they are watching acts in a way appropriate to that person's previous experience or not; but children seem not to respond correctly to questions about false beliefs until around age 4 (Baillargeon, Scott & He 2010; Perner & Ruffman 2005; Southgate, Senju & Csibra 2007). This raises questions regarding whether infants are sensitive to beliefs, to intentions, or to something even more basic (Fenici 2015a,b), and whether eye gaze and other such implicit measures recruit the same processes recruited in verbal report and other explicit responses (Apperly & Butterfill 2009; Low & Perner 2012; Low & Watts 2013).

Languages have an array of devices which require attention to someone else's point of view. Articles, pronouns, adjectives of personal taste, opinion adverbs and deictic locatives all shift in denotation depending on the speaker, and children must learn the use of these terms from observing how they are used by someone whose perspective they do not share (see de Villiers submitted). Children appear to master the spontaneous use of many of these deictic forms by age 4. It is therefore plausible that children's facility with implicit perspective taking – demonstrated in infancy through eye gaze, and through the gradual mastery of deixis – is in place well

before the capacities measured by the standard false belief tasks in ToM research. In a standard ToM task, an object may be moved or replaced by a different object, out of sight of a character who saw it originally. The child is then asked to predict either where the character will look for the object (Wimmer & Perner 1983) or what he thinks is in the container before he looks inside (Gopnik & Astington 1988).

Evidentials may be acquired in the same way that spatial deictic terms are acquired across languages (see discussion in de Villiers & Garfield 2009). Evidentiality is harder than spatial deixis because it is not simple position in space that governs the perspective shift, but the appropriate alignment of event situations, time, and information access. Evidentials are always egophoric: an evidential encodes the evidence the *speaker* has for a statement.[1] Nonetheless, to acquire the meanings of evidentials from the speech around them, children must figure out not only the point of view of the speaker, but also how other people got their knowledge, and they must map that information onto the appropriate morphemes. Thus, acquisition of the evidential system seems to occupy a special place in the inventory of terms that entail perspective, more complex than spatial deixis, but perhaps less demanding than understanding the contents of another's false belief.

How can evidentiality be learned? One possibility would be negative feedback from caregivers if the child misused an evidential. Yet evidentials present a classic problem of negative evidence in language learning: since the morphemes cannot themselves be denied (see Section 4), caregivers cannot directly correct children's evidential use. Second, caregivers do not in general correct young children's grammar (Marcus 1993). Even if they did, if caregivers were to correct and produce their own evidential as a better example, that would not be a satisfactory model for the child since their two epistemic situations might not be identical: the child may see something that the adult has not seen. Imagine the perilous discourse of pronoun correction in the hypothetical example in (1):

(1) Child: "Pick you up!"
Mom: "No, pick YOU up" or "No, pick ME up"

On the other hand, it might be that the correct use of terms requiring perspective shift requires less of children than many have supposed. Enough perspectives may be shared between, say, mother and toddler, that full competence even with deictics is more apparent than real. The study of evidentials in both ordinary discourse and controlled experimental situations may shed light on this issue (see Uzundağ,

1. But see Schenner (2010) for an interesting discussion of shifts in egophoricity in embedded context and in questions. We return to questions below.

Taşçı, Küntay & Aksu-Koç 2015). Indeed as we will show below, Tibetan mothers often come to the rescue of young children by explicitly distinguishing the felicity conditions of the evidentials in conversation. Suitably scaffolded discourse may both make it difficult for children to make mistakes, and give them clues about how to use the evidentials felicitously even before they have mastered the meaning of the evidential system.

Evidentials seem to fall between the traditional domains of pragmatics and semantics. On the one hand, they are grammatical features and, like tense or modal operators, contribute to the meanings of the sentences in which they occur. On the other hand, they do not seem to contribute to the *truth conditions* of those sentences. Instead they seem to contribute to *felicity* conditions, which are typically considered to be pragmatics, as they concern the ideal conditions of usage, not truth. This is no accident, and the answers to the developmental and semantic questions are intertwined. A situation semantics for evidentials will help to explain why this is so, and our semantic analysis will also partially explain the otherwise puzzling developmental pathway of evidentials.

2. Overview of the Tibetan evidential system

Tibetan has a rich evidential system, representing a set of distinctions involving all of the known evidential types except hearsay, with two distinct types of inferential evidentials (Garrett 2001). In Tibetan, most sentences end in a form of the copula or a verb of existence. Tibetan evidentials are distinctive forms of the copula or the verb of existence. As a result, evidentiality is a feature of virtually every Tibetan assertion or question. Tibetan represents ego evidentials, direct perception evidentials, evidentials that mark inference from specific evidence, and evidentials that mark inference from general knowledge as well as non-inferential general knowledge evidentials, as illustrated in (1) through (5) respectively:

Ego evidentials are marked with the verbs *yin* or *yod*.[2, 3]

---

2. All Tibetan spellings are rendered in the standard Wylie transcription system. The Wylie system is not phonetic, and codes Tibetan letters in Roman script. Tibetan orthography separates syllables each of which will have a root letter, an optional vowel, as well as optional prefix letter, superscript, subscript and one or two optional suffixes. There are no breaks between words. Some letters are coded with a single Roman letter, some with two. (e.g, the first four letters of the Tibetan alphabet are coded *k, kh, g, ng*) For details and for pronunciation guides, see any standard textbook of Tibetan language.

3. The specific grammatical and semantic distinctions between the various Tibetan evidential morphemes are explored in detail in Kalsang, Garfield, Speas & de Villiers (2013).

(1) *yin*: Ego Copula
Nga skyid po yin
I happy am (ego).
'I am happy.'

(2) *yod*: Ego Possessive.
Nga la khyi zhig yod.
I LOC dog DET is (ego)
'I have a dog.'

These encode the fact that I know the truth of the sentence I am asserting just in virtue of *being me.* It is first person knowledge drawn neither from any particular perceptual evidence nor by inference.

Direct evidentials are marked by *'dug, song* and *shag*.

(3) '*dug*: Direct Witnessed State.
Khong tshos ja btungs gyi 'dug.
They tea drink (INSTR) IMP are (DIR).
'They are drinking tea.'

(4) *song*: Direct Witnessed Past Action.
Khong tsho Lha sar phebs song.
They Lhasa (LOC) go (DIR PAST).
'They went to Lhasa.'

(5) *shag*: Direct Resultative
Bum pa bchag shag.
Vase broke (DIR resultative)
'The vase broke.'

Each of the statements in (1) through (3) encodes the fact that I (the speaker) know the truth of the sentence because I witnessed the state or event being reported. To use these felicitously I must be seeing the tea-drinking in (3), have seen their departure for Lhasa or their arrival there in (4), or have seen the pieces of the shattered vase in (5).

Tibetan distinguishes two kinds of indirect or inferential evidentials, which we call *specific inference* and *general inference* evidentials respectively. Specific inference is marked by *yod sa red* or *yin sa red*:

(6) *yod sa red* (specific inference)
bKra[4] *shis las kung nang la yod sa red.*
Tashi office in is (SPEC).
'Tashi is in his office.'

4. *bKra shis* (pronounced *Tashi*) is a proper name, and so the root letter of the first syllable (*ka*, coded *k* in Wylie) is capitalized.

This utterance would be felicitous if, and only if, I have some specific piece of evidence that indicates Tashi's presence, such as his umbrella in the hall and the light on in the office.

General inference is marked by *yod kyi red,* as in (7).

(7) *yod kyi red (*general inference)
bKra shis las kung nang la yod kyi red.
Tashi office in is (GEN).
'Tashi is in his office.'

This sentence would be felicitous in a case where I can neither see Tashi in his office nor do I have specific evidence that he is there, but can deduce that he is there from more general knowledge. I might utter this when I know that it is his office hour, or when all staff are in their offices.

Finally, Tibetan allows a general knowledge (sometimes called neutral) evidential. This is marked by *red* (copula) or *yod red* (existence):

(8) Neutral *yod red*.
Bod la gyag mang po yod red.
Tibet (LOC) yak many are (NEUTRAL)
'There are many yaks in Tibet.'

(9) Neutral *red*.
gYag nag po red.
yak black are (NEUTRAL)
'Yaks are black.'

These sentences are felicitous because they report general knowledge, for which no specific evidence is cited.

3. Evidentiality and acquisition challenges

Studies of the acquisition of evidentials across languages reveal four noteworthy phenomena. First, evidentials appear in *spontaneous speech* around age 2 years in Turkish (Aksu 1978), in Korean (Choi 1991, 1995) and in Tibetan (de Villiers, Garfield, Gernet-Girard, Roeper & Speas 2009). Second, when production is elicited in *controlled conditions*, children do not demonstrate control until age 4 years or later, and they show earlier control of direct evidentials than of indirect evidentials (Aksu-Koç 1988; Öztürk & Papafragou 2008; de Villiers et al. 2009). Recent results by Ünal and Papafragou (2016) suggest that children at 3 years old speaking Turkish have some productive control, even for the indirect evidential.

Nevertheless, control in production may not be complete, as their 5-year-olds overgeneralized the direct past tense use to the inferred events 28% of the time. Third, productive competence in evidentials does not entail receptive competence. Comprehension of the meaning of evidentials in controlled circumstances is not reliable until after age 4 (see also Aksu-Koç & Alıcı 2000; Öztürk & Papafragou 2008; and Ünal & Papafragou 2016, for Turkish; Papafragou, Li, Choi & Han 2007, for Korean; Kyuchukov & de Villiers 2009, for Bulgarian and Romani). Fourth, children seem to understand that if a person has directly seen something one has knowledge about it (the information conveyed by direct evidentials) at around the same age across languages, that is, about age three or four years (Pillow 1989; Pratt & Bryant 1990). Ünal and Papafragou (2016) used a conceptual task testing whether 4- and 5-year-old children understood how evidence and knowledge were linked, and compared it to a matched evidential comprehension task. When the children had to understand the other's perspective, they had more difficulty than when they were asked from the perspective of self. This conceptual understanding about others' knowledge seems to arise later than the correct use of direct evidentials in production.

One might argue that unnatural controlled experimental tasks do not reflect the demands of ordinary conversation. Most studies use a methodology that requires a judgment about the aptness of the evidential. For example two characters describe a scene using different evidentials, and the child is asked "who said it best?" (Ünal & Papafragou 2016). Or, two voices are heard describing a scene, where one puppet can see it and one can not see it, and the child is asked "which puppet said that?" (Kyuchukov & de Villiers 2009). The meta-linguistic demands of such tasks, and not a failure to master evidentials, may be responsible for errors. Consequently, children might know the meanings of the evidentials at an earlier age than that suggested by experimental data, but their competence might be masked by the sophisticated task demands of the experiments. We address this possible objection in our own experimental work reported below, but Ünal and Papafragou (2016) get similar results from a range of tasks and argue against the methodological explanation for the asymmetry between production and comprehension.

4. Tibetan evidentials and maternal-child speech

We have collected a small database of elicited and spontaneous conversations between eleven Tibetan mothers and their young children aged 3 to 5 years (de Villiers & Garfield 2009). The sample of dialogues contained 587 utterances from the children and 896 utterances from their mothers. In the course of these conversations,

children produced 296 evidentials, and mothers produced 680. We were able to compare the relative frequencies of particular evidential use by mothers and children to those of Tibetan adults speaking to each other (Edward Garrett personal communication), but to do so we had to consider the total numbers of words in each database. Table 1 reveals that the frequencies of evidential use are massively weighted towards the direct evidentials, with inferential forms being very rare in the mothers' speech. Mothers are actually producing speech containing evidentials – even the rare ones – at much greater rates than reported in conversations between Tibetan adults. Their use of direct evidentials is well matched to that of their children, but the children produced no indirect evidentials.

Table 1. Incidence of particular evidentials in Tibetan speech

Evidential:	**Mother-to-child**	**Mother%**	**Adult-to-Adult**	**Adult %**	**Child-to-mother**	**Child %**
dug (direct)	219	2.44	9002	.59	79	0
yod sa red (indirect)	8	.09	28	.0019	0	0
yod kyi red (indirect)	1	.01	76	.0050	0	2.48
Total words	**8,968**		**1,513,462**		**3,181**	

In the development of direct evidentials, the demonstrative *'dug ga* is used almost always with a demonstrative gesture to elicit shared attention on a focal object (like the English *look!*). This draws a child's attention over time not only to the object of shared attention, but to the fact that *'dug* is being used to reflect the fact that something can be seen by the speaker. This demonstrative construction probably plays a crucial role in scaffolding the direct evidential meaning.

The harder distinction is that between *yod sa red* and *yod kyi red,* the *indirect evidentials.* But the fact that *'dug* is established helps here, too. Whenever *yod sa red* is felicitous, there is always some *other* state of affairs – the relevant evidence – for which a *'dug* statement is felicitous. If we consider the spontaneous dialogues between Tibetan mothers and their children that we recorded (see Table 2 for illustrations), we can see that the mother uses a specific inferential evidential for a broad claim, and backs it up with a statement marked by the direct evidential about visible signs justifying the inference. Hearing conjunctions of claims like these provides good information for the child about how an inference is warranted. These pedagogical dialogues reveal the adult's sensitivity to clarifying the reasoning for the child, and in the process reveal the particular felicity conditions for specific indirect evidentials versus direct evidentials.

Table 2. Examples of *Tibetan* mother's use of indirect (inferential) evidentials in natural samples [5]

Example 1						
Kyod rang	*gyi*	*choc ho*	*ku li*	*rgyug*	*ga phyin*	*yod sa red.*
You	gen	brother	coolie	resembles	become (past)	indirect evidential
Gzugs la		*nag po*	*god*	*'dug*		
Body	locative	black	dirt	direct evidential		
'Your brother looks like a laborer; he has black dirt on his body'.						
Example 2						
phun tsok	*yang*	*so*	*rus*	*'dug*		
Phunstogs	possessive	tooth	rotten	direct evidential		
kho	*yang*	*so*	*rus*	*'dug*		
He	possessive	tooth	rotten	direct evidential		
Youngling	*slob*	*gra la*	*cong*	*tso*	*mngar mo*	*kyang kyang*
Youngling	school	locative	kid	plural	sweets	repeatedly
bza'	*sdad*	*kyi*	*yod sa red*			
eat	present	cont.	indirect (specific inference) evidential.			
Phuntsog's teeth are rotten. Youngling school kids are always eating sweets.						

We have argued (de Villiers, et al 2009) that the understanding of evidentials is established through a series of contrasts, just as phonological distinctions are mastered. The child represents increasingly subtle and abstract distinctions between epistemic situations. The earliest evidentials to be mastered (both in comprehension and production) are the ego and direct evidentials in virtue of their frequency in mother-child speech and the concreteness of their meanings. The distinction between these evidentials emerges first (around ages 3 to 4 years). The distinction between direct and indirect evidentials appears next (at about age 6), but involves a conflation of the two indirect evidentials. The more difficult distinction between the two kinds of indirect evidentials does not emerge until approximately 9–10 years of age. We will show that this developmental track reflects the relative complexity of the meanings of the different kinds of evidentials.

5. The database was collected as part of NSF grant # HSD 0527509. It contains conversations between 11 Tibetan-speaking children and their mothers.

5. The semantics of evidentials

There are many ways to say that there is a mouse over there in Tibetan: A speaker who says *rtsi rtsi pha gir* ***'dug*** is asserting that there is a mouse over there, and indicating by the use of *'dug* that she saw it. Or she could say *rtsi rtsi pha gir* ***yod sa red*** when she directly sees mouse footprints in the dust, but not the mouse itself. On the other hand, by uttering *rtsi rtsi pha gir* ***yod kyi red***, she indicates that she knows the presence of the mouse by inference of a more general sort, for example, the mouse is there at this time every day. Finally, she could say this expressing general knowledge by saying *rtsi rtsi mang khrul di la yod red,* in the condition where everyone in the community just knows that there are mice in that area . The truth-conditions of these statements are the same, but their implicatures are very different. Evidentials, as we have pointed out, encode pragmatic or illocutionary information; they do not contribute to truth-conditions (unlike epistemic modals or propositional attitude verbs).

Tibetan evidentials are grammatically unlike propositional attitude verbs (e.g., think, believe, know) in that they do not assign case and are always egophoric, i.e. reflecting the speaker's perspective. Secondly, they are unlike epistemic modals in that they are felicitous (indeed mandatory) in conditions of known truth or falsehood and do not weaken the force of a claim. Thirdly, evidential force cannot be denied. To deny a sentence asserted with evidential force is to deny its asserted content, not to deny the felicity of the evidential.[6] These three properties of evidentials locate them semantically squarely among illocutionary operators, operators that have felicity-conditions, but not truth-conditions, that contribute implicature and force to an assertion, but do not directly assert content. They are hence *pragmatic* operators. Nonetheless, evidentials are highly restricted in meaning. As Speas (2010) has argued, the kinds of evidentiality encoded across the world's languages is tightly constrained, and those constraints appear to be systematic, determining a set comprising only direct, inferential, hearsay and ego meanings. Evidentials are hence, unlike most other pragmatic operators, syntactically mandatory and semantically constrained. Is part of the difficulty in learning the evidential system that Theory of Mind is necessary in order to understand them? It is to this question that we will turn shortly. First, however, it is useful to explore the semantics of evidentials.

Kalsang et al. (2013), following Speas (2010), present a situation semantics (see Barwise & Perry 1981, 1983) for evidentials (see also Schenner 2010). This semantic

6. When a statement governed by an evidential is denied, the denial must be read as a denial of the truth of the content of the assertion. When a statement governed by an epistemic modal of a propositional attitude verb is denied, on the other hand, one can deny the modal force of the fact that the relevant attitude obtains.

framework explains the restricted set of permissible evidentials and demonstrates their systematic relation to one another. They show that the relative complexity of the Tibetan evidential system consists simply in its representing more of the permissible relations between situations that evidentials can encode than do many other languages. Kalsang et al. (2013) show that evidentials encode inclusion and accessibility relations between situations, not a primitive category of evidence. That is, while it appears that the phenomenon of evidentiality entails that *kinds of evidence* are semantic primitives, this is not the case.[7]

As Barwise and Perry (1983) and those who have followed them suggest, discourse itself forces an account of meaning in which situations are essential elements of a semantics. Three types of situations emerge as central to discourse semantics: the discourse situation in which the speaker and hearer find themselves; the evaluation situation; and the information situation. The first two are familiar from earlier situation semantic models and fall out naturally from the demands of evaluation. It is necessary to distinguish the situation in which an utterance occurs from that which determines its truth or falsity. Evidentiality calls our attention to the information situation, the situation from which evidence is drawn on the basis of which a sentence is evaluated. The set of possible relations between these situations determines the class of evidentials represented in natural languages. (For details, see Kalsang et al. 2013).

The two relations between situations that determine the meaning of evidentials are inclusion and accessibility. A situation S includes a situation S′ iff S′ is a part of S. So, for instance, a situation in which you are I are talking includes a situation in which I am talking. So inclusion is analogous to the familiar subset relation. When we assert – or implicate – that S includes S′, we are conveying the view that S′ is a part of S. If I say something that encodes that information, and you believe me, you come to believe that S includes S′.[8]

7. Indeed, it would be surprising if it were. Given the enormous range of types of evidence one could have for a proposition, why, if evidence type is a primitive, would the range of evidentials be so restricted in the world's languages?

8. Inclusion is different from accessibility. A situation S is accessible from another S' if, and only if, S' includes the information available in S. S' need not include S itself, but only the information present in S. Consider the situation in which you and I are talking and my diary is open to a page that tells me that I have a doctor's appointment tomorrow afternoon. Now consider the situation tomorrow afternoon when I am at the doctor's office. The present situation does not contain the future situation, but it contains information about it. So, tomorrow's situation is accessible from today's, even though not included in it. Note that neither the inclusion nor the accessibility relation requires reference to any inner mental or epistemic states: one could understand these relations (implicitly or explicitly) even if one had not mastered ToM.

The structures that determine the meanings of evidentials are important as they set the learning task for the child acquiring an evidential language. Instead of basic facts about epistemology, the child needs to learn about inclusion and accessibility relations between situations. So, while it might appear that the evidential system requires the child to understand the contents of others' minds, it does not. With this apparatus in hand, we can explain the difference between direct, indirect and ego evidentials in Tibetan. A direct evidential encodes the fact that the information situation includes the evaluation situation and is accessible from the discourse situation. That is, when I say

(10) *Nga'i skyi sha za gi* ***'dug***.
My dog meat eat GEN is (direct)
'My dog is eating meat.'

I convey the following information: (a) The evaluation situation – that is, the situation that makes my sentence true (if it is true), the one in which the dog is eating beef – is a part of the situation in which I gain the information that he is doing so – the one in which I see him eating it; and (b) the information that he is eating meat is present in the situation in which we are talking, even if he and the meat are not.

An *indirect evidential*, on the other hand, encodes the information that the information situation is accessible from the evaluation situation, and that the information situation includes the discourse situation. That is, when I say

(11) *Nga'i skyi sha za* ***yin sa 'red***.
My dog meat eat is (indirect/specific)
'My dog is eating meat.'

When I assert (11), I convey the following information (beyond asserting that my dog is eating meat): (a) In the evaluation situation – the situation in which the dog is eating beef – we find the information on the basis of which I make the claim (say, the sound of his chomping), and so the information situation is accessible from the evaluation situation; (b) the discourse situation is a part of the information situation. That is, the situation in which we are now talking is part of the larger situation in which I have the information (the sound of the chomping) on the basis of which I assert that my dog is eating meat.

Finally, an *ego evidential*, encodes the assertion that the information situation includes both the evaluation and the discourse situations. So, when I say:

(12) *Ngas mogs mogs mang po za song.*
I (INSTR) momos many eat past (ego)
'I ate many momos.'

I convey the following information: (a) the situation on the basis of which I say that I stuffed myself with momos (Tibetan dumplings) (my happy, sated state with the memory of momos) includes the situation that makes this true – that is, the present state is part of a broader temporal period an early stage of which had me gorging on momos; (b) that broader situation includes the situation in which we are now talking. It is bigger than both of them. After all, it is the situation in which I ate momos in the past and now remember them as I speak to you. (This is a very quick tour through the situation semantics of evidentials. For details, see Kalsang et al. 2013).

The general knowledge evidential *red* encodes the fact that the discourse situation is as broad as can be: it includes both the evaluation and the information situation. So, when I assert:

(13) *Bod la gyag mang po yod red.*
Tibet LOC yak many copula (neutral)
'There are many yaks in Tibet.'

I convey the information that in the situation in which we are now talking, it is simply true that there are lots of yaks there, and that we have this information. That is what common knowledge claims are like: we claim that we are in the situation where the claim is true and that anyone can come to know it.

So much for assertion. The important point here is that when we understand the structure of the meanings of evidentials, we see that there is no reason to expect that a child must master ToM in order competently to make or to understand assertions that carry evidential force. Nothing in their meaning makes reference to internal states or states of knowledge of particular persons. Instead, evidentials encode relations between discourse, evidence and evaluation situations. The difficulties facing the child are not those involved in learning about internal mental states.[9] Moreover, the complexity of these distinctions can explain the specific developmental arc of evidential acquisition in Tibetan. But, as we shall now see, the grammar of questions in Tibetan adds a wrinkle.

9. This is not, of course, to say that competent users of evidentials are reflectively aware of the details of their semantics. Of course they are not. But that is not surprising. Most competent speakers of any language are incapable of providing a compelling formal semantics of their native language. But once we appreciate what evidentials mean, we see that there is no need in order to master them to have any knowledge – implicit or explicit – about the mental states of others.

6. The reflection principle for questions

When asking a question in Tibetan, one uses the evidential that one expects the interlocutor to use in his reply, that is, the evidential one presumes to be felicitous for the interlocutor. Suppose I ask you:

(14) *bSod nams khong gi las kung nang la 'dug gas*?
Sonam he GEN office in is (DIR) QUEST
'Is Sonam in his office?'

The fact that I use *'dug* in my question reflects my anticipation of the felicity of *'dug* in *your* answer, even though *I* could not report Sonam's absence or presence using a direct evidential. Or, when I ask how you are doing today, as in

(15) *Khed rang* *bde po* *yin* *pas?*
You comfortable copula (ego) question
'Are you well?'

I anticipate that you will reply with the ego evidential *yin*, one I could never use when reporting how you feel.

We have argued that the child may control evidentiality with no real awareness of or attention to the mental states of others. Instead, they learn to identify the sets of situations and their relations that call for different evidentials. Do *questions* change this picture? In asking questions in Tibetan, one must suspend one's own perspective on situations and attend to the evidential one supposes felicitous for the interlocutor. This therefore requires more than what is necessary to learn to use the evidential in assertion. For that reason, asking questions in Tibetan seems more likely to require the kind of competencies in ToM that children achieve by four or five years. Therefore, even though mastering the situation semantics of some evidentials even for ordinary use is a protracted process, we expected that mastering appropriate use of questions might take even longer.

To address this question, we designed a game involving ordinary discourse, using puppets to simulate real life demands. The protocol is just a game of discourse with puppets, so it should come closer to ordinary linguistic demands than a meta-linguistic task. The events were acted out with puppets and boxes that represented locations such as a bus, a shop, a kitchen and so on. These props showed whether the puppets who were talking were in different positions with respect to access to crucial information about the events in the story. For example, in one story – see excerpt in Table 2 – one puppet is outside the cafe and does not have access to visual information about the event happening in the cafe. A second puppet did get to see what happened in the cafe. In that way we contrasted a character who was a direct witness, with someone who had to rely on inference. In other stories we had

characters who reported on their own tastes (e.g. for tea) using ego-evidentials. Importantly, all of this was acted out with puppets and props, so the burden on memory and understanding was not so high.

We asked several questions. First, we wanted to adjudicate between two possibilities: (1) since this testing procedure mimics ordinary discourse, it might show that indirect evidentials are easier to understand, or acquired earlier than was apparent from our more metalinguistic tasks; or (2), it might be that children's productive mastery of appropriate evidentials might in fact be later, and more consistent with comprehension than was earlier apparent. Most importantly, we wondered whether, if the perspective-taking burden is in fact heavier for evidential use in questions, competence with all evidentials might be more difficult in this protocol, despite its being more naturalistic.

We tested 11 native Tibetan-speaking children between the ages of 4 and 8 years, in Tibetan-speaking communities in India. All were tested by a native Tibetan-speaking research assistant who also had conducted other studies for us and was very well practiced in working with children. Each child was tested using the two protocols in Figure 1. All together the protocols called for a total use of 30 evidentials in questions: 10 ego evidentials, 9 direct, 7 indirect, and 4 neutral evidentials. All responses were transcribed carefully by native speakers and coded for whether the child asked the question of the protagonist using the right evidential type.

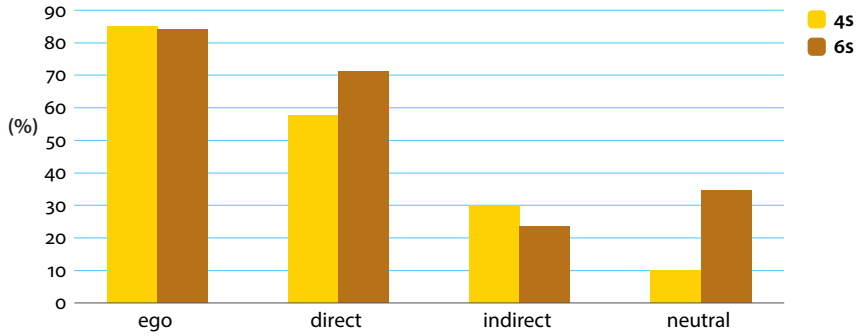

Figure 1. Children's percentage correct on different evidentials

A repeated-measures ANOVA comparing the percentage correct across the four types of evidential confirms that these are significantly different ($F = 27.5$, [df 3,10] $p < .001$). Furthermore, across the children they are consistently ranked in the difficulty ordering: ego>direct>indirect>neutral (Kendall's W, $p < .001$).

Ego-evidentials were used correctly most often, as they are in ordinary speech. Direct evidentials were the next easiest, but they do seem more difficult in questions than in elicited production protocols, only reaching 70% for the older group of

children. Indirect evidentials were quite difficult in this age range as expected, with very low use, and directs were used inappropriately instead. Neutrals were rare and were also most often replaced by direct evidentials.

Though this was a small study, these data are at least consistent with the claim that the requirements of the reflection principle in questions place demands on the speaker beyond the requirements of ordinary egophoric use. Hence these forms may indeed have ToM skills as prerequisite, but a good deal more investigation is necessary to confirm this.

7. Conclusion

Mastery of the evidential system is staged. Ego and direct evidentials are mastered relatively early. Spontaneous speech is heavily loaded towards these evidentials, and enough circumstances are shared between conversation partners that performance can look errorless even when comprehension is not fully in place. This may be similar to the acquisition pattern of spatial and pronominal deictic terms. Indirect evidentials, however, which require an additional understanding of inference, are only acquired later. Understanding the entire evidential system thus entails more than simply syntactic and semantic development; an understanding of inference is essential as well. This is reflected by the fact that mothers using indirect evidentials in conversation with their children clarify their conditions of use by pairing them with statements with direct evidentials regarding the evidence that justifies the inference. Though the process of evidential acquisition is protracted, we do not see any need to invoke ToM skills – even implicitly represented- in the development of the production or comprehension of evidentials in assertion.

The reflection principle involved in Tibetan questions may be another matter. Anticipating others' epistemic states in order to ask questions is more difficult. The degree of difficulty of using evidentials in this context mirrors their order of difficulty in production and comprehension. This, in turn, reflects their frequency in maternal conversation. But there appears to be an additional lag in using the right forms in questions. This may be a sign that the reflection principle does require some representation of the other's epistemic state, necessary to resist the evidential appropriate for the self and to compute the right form for the listener to use in reply.

If this is correct, the late and complex acquisition of the ordinary use of evidentials in assertions reflects the need to understand a complicated situation semantics, not the need for ToM. In this respect our position is close to that of Ünal and Papafragou (2016), who argue that in comprehension, a child needs to attend to the information access in another's mind. This requires perspective taking but not necessarily full blown attribution of mental states to others. Our situation

semantics shows how one can represent the information available to participants in a discourse without representing their mental states. On the other hand, mastery of evidentials in questions in Tibetan does require attention to and representation of others' epistemic states.

Acknowledgement

We thank the National Science Foundation Human Dynamics for support under grant BCS#0527509. We also thank Smith College, the University of Massachusetts and the Central University of Tibetan Studies for institutional support. Thanks to Kalsang, Tsering Toepgyal, Tashi Dolma, Namgyal Norbu, Caroline Sluyter, Harper Genet-Girard, Megan Kravitz, Tom Roeper and Peggy Speas for research assistance and collaboration. Thanks to audiences for helpful comments at the International Association for Child Language in July 2011, the U.Mass/ U.Conn/ Smith Language Acquisition workshop (UUSLAW) in May 2010 and the Central University of Tibetan Studies in January 2010.

References

Aikhenvald, A. 2004. *Evidentiality*. Oxford: OUP.

Aksu, A. A. 1978. *Aspect and Modality in the Child's Acquisition of the Turkish Past Tense*. PhD dissertation, University of California, Berkeley.

Aksu- Koç, A. 1988. *The Acquisition of Aspect and Modality: The Case of Past Reference in Turkish*. Cambridge: CUP. doi: 10.1017/CBO9780511554353

Aksu-Koc, A. & Alıcı, D. M. 2000. Understanding sources of beliefs and marking of uncertainty: the child's theory of evidentiality. In *Proceedings of the 30th Annual Child Language Research Forum*, E. V. Clark (ed.), 123–30. Stanford CA: Stanford University Press.

Apperly, I. A. & Butterfill, S. A. 2009. Do humans have two systems to track beliefs and belief-like states? *Psychological Review* 116(4): 953–970. doi: 10.1037/a0016923

Astington, J. W. 2000. Language and metalanguage in children's understanding of mind. In *Minds in the Making: Essays in Honor of David R. Olson*, J. W. Astington (ed.), 267–284. Malden: Blackwell.

Astington, J. W. & Baird, J. A. 2005. Representational development and false-belief understanding. In *Why Language Matters for Theory of Mind*, J. W. Astington & J. A. Baird (eds), 163–185. Oxford: OUP. doi: 10.1093/acprof:oso/9780195159912.003.0009

Baillargeon, R., Scott, R. M. & He, Z. 2010. False-belief understanding in infants. *Trends in Cognitive Sciences* 14(3): 110–118. doi: 10.1016/j.tics.2009.12.006

Bartsch, K. & Wellman, H. M. 1995. *Children Talk about the Mind*. Oxford: OUP.

Barwise, J. & Perry, J. 1981. Situations and attitudes. *The Journal of Philosophy* 78(11): 668–691. doi: 10.2307/2026578

Barwise, J. & Perry, J. 1983. *Situations and Attitudes*. Cambridge MA: The MIT Press.

Choi, S. 1991. Early acquisition of epistemic meanings in Korean: A study of sentence-ending suffixes in the spontaneous speech of three children. *First Language* 11(3): 93–119. doi: 10.1177/014272379101103105

Choi, S. 1995. The development of epistemic sentence-ending modal forms and functions in Korean children. In *Modality in Grammar and Discourse* [Typological Studies in Language 32], J. Bybee & S. Fleischman (eds), 165–204. Amsterdam: John Benjamins. doi: 10.1075/tsl.32.09cho

de Villiers, J. G. In preparation. Perspectives on truth: The case of language and false belief reasoning. In *Semantics in Language Acquisition* [Trends in Language Acquisition Research], K. Syrett & S. Arunachalam (eds). Amsterdam: John Benjamins.

de Villiers, J. G. & de Villiers, P. A. 2000. Linguistic determinism and the understanding of false beliefs. In *Children's Reasoning and the Mind*, P. Mitchell & K. Riggs (eds), 191–228. Hove: Psychology Press.

de Villiers, J. G., Garfield, J., Gernet Girard, H., Roeper, T. & Speas, M. 2009. Evidentials in Tibetan: Acquisition, semantics and cognitive development. In *Evidentiality: A Window into Language and Cognitive Development [Special Issue]*, S. Fitneva & T. Matsui (eds), *New Directions for Adolescent and Child Development* 125: 29–48.

Fenici, M. 2015a. Social cognitive abilities in infancy: Is mindreading the best explanation? *Philosophical Psychology* 28(3): 387–411. doi: 10.1080/09515089.2013.865096

Fenici, M. 2015b. A simple explanation of apparent early mindreading: infants' sensitivity to goals and gaze direction. *Phenomenology and the Cognitive Sciences* 14(3): 497–515. doi: 10.1007/s11097-014-9345-3

Garrett, E. 2001. Evidentiality and Assertion in Tibetan. PhD dissertation, UCLA.

Gopnik, A. & Astington, J. W. 1988. Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development* 59: 26–37. doi: 10.2307/1130386

Kalsang, T., Garfield, J., Speas, M. & de Villiers, J. G. 2013. Direct evidentials, case, tense and aspect in Tibetan. *Natural Language and Linguistic Theory* 31: 517–561. doi: 10.1007/s11049-013-9193-9

Kyuchukov, H. & de Villiers, J. G. 2009. Theory of Mind and evidentiality in Romani-Bulgarian bilingual children. *Psychology of Language and Communication* 13(2): 21–34. doi: 10.2478/v10057-009-0007-4

Low, J. & Perner, J. 2012. Implicit and explicit theory of mind: State of the art. *British Journal of Developmental Psychology* 30: 1–13. doi: 10.1111/j.2044-835X.2011.02074.x

Low, J. & Watts, J. 2013. Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science* 24(3): 305–311. doi: 10.1177/0956797612451469

Marcus, G. F. 1993. Negative evidence in language acquisition. *Cognition* 46(1): 53–85. doi: 10.1016/0010-0277(93)90022-N

Milligan, K., Astington, J. W. & Dack, L. A. 2007. Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child development* 78(2): 622–646. doi: 10.1111/j.1467-8624.2007.01018.x

Öztürk, Ö. & Papafragou, A. 2008. Acquisition of evidentiality and source monitoring. In *Proceedings of the 32nd Annual Boston University Conference on Language Development*, H. Chan, H. Jacob & E. Kapia (eds), 368–377. Somerville MA: Cascadilla Press.

Papafragou, A., Li, P., Choi, Y., & Han, C. H. 2007. Evidentiality in language and cognition. *Cognition* 103(2): 253–299. doi: 10.1016/j.cognition.2006.04.001

Perner, J. & Ruffman, T. 2005. Infants' insight into the mind: How deep. *Science* 308(5719): 214–216. doi: 10.1126/science.1111656

Pillow, B. H. 1989. Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology* 47(1): 116–129. doi:10.1016/0022-0965(89)90066-0

Pratt, C. & Bryant, P. 1990. Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development* 61(4): 973–982. doi:10.2307/1130869

Schenner, M. 2010. Evidentials in complex sentences: foundational issues and data from Turkish and German. In *Evidence from Evidentials*, T. Peterson & U. Sauerland (eds), 28: 183–220. University of British Columbia Working Papers in Linguistics.

Shatz, M., Wellman, H. M. & Silber, S. 1983. The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition* 14(3): 301–321. doi:10.1016/0010-0277(83)90008-2

Southgate, V., Senju, A. & Csibra, G. 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science* 18(7): 587–592. doi:10.1111/j.1467-9280.2007.01944.x

Speas, P. 2010. Evidentials as generalized functional heads. *Linguistik aktuell: Amsterdamer Arbeiten zur theoretischen & angewandten Linguistik* 156: 127–150. doi:10.1075/la.156.10spe

Ünal, E. & Papafragou, A. 2016. Production-comprehension asymmetries and the acquisition of evidential morphology. *Journal of Memory and Language* 89: 179–199. doi:10.1016/j.jml.2015.12.001

Uzundağ, B., Taşçı, S., Küntay, A. & Aksu-Koç, A. 2015. Functions of evidentials in Turkish child and child-directed speech in early child-caregiver interactions. Paper presented at the *40th Annual Boston University Conference on Language Development*.

Wimmer, H. & Perner, J. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13(1): 103–128. doi:10.1016/0010-0277(83)90004-5