Who knows what and how?

New Evidence about the Acquisition of Evidentials in Tibetan

Jay L Garfield

(Smith College, Harvard Divinity School, University of Melbourne,

Central Institute of Higher Tibetan Studies)

Jill de Villiers

(Smith College)

Contact J Garfield jgarfield@smith.edu, Dept of Philosophy, Smith College,

Northampton, MA 01063

Who knows what and how?

New Evidence about the Acquisition of Evidentials in Tibetan

Abstract

There is much interest in the relationship between the acquisition of evidentials and Theory of Mind. Tibetan is unusual in that it contains two distinct indirect evidentials, one encoding knowledge gained by *direct perception* of *indirect* evidence for the truth of the claim; the other encoding knowledge by more general inference. Attention to how and when children learn this distinction sheds light on the cognitive processes that drive the acquisition of evidentials. We tested 32 Tibetan children age 4-11 years in Tibetan exile communities in India and found that comprehension of this distinction is acquired at around age 9, and correlates strongly with linguistic sophistication and with productive use of both direct and indirect evidentials, suggesting that acquisition of the evidential system is a late stage of linguistic development.

1.Introduction

1.1 **General background**

Evidential morphemes are grammaticalized forms for expressing the nature of a speaker's evidence for a statement. As many as one-quarter of the world's languages have a specialized grammatical system for encoding what type of evidence the speaker has for a declarative statement (Aikhenvald 2004). In English, these meanings tend to be carried by propositional attitude verbs, such as "I guess that," "he knows that," "she saw that" by adverbs, such as "probably," "apparently," or even "evidently" or m=by epistemic modals, such as "must" or "could." "But evidentials differ from attitude predicates in three ways. First, an evidential always expresses the evidence that the *speaker* has for the sentence, never that of a second or third person.[1] Second, evidentials convey *how* the speaker's epistemic state was achieved rather than the *nature* of the epistemic state itself (such as the degree of certainty). Third, unlike attitude predicates, evidentials generally occur only in main clauses and do not introduce a syntactically subordinate clause. Fourth, compared to attitude verbs and adverbs, evidentials encode a restricted range of meanings across languages (Kalsang et al. 2013). Epistemic modals on the other hand are not necessarily egophoric, do not create embedded contexts and do not weaken assertoric force. Moreover while they may indicate how an epistemic state was achieved, they do not necessarily do so.

Evidentials generally encode a distinction between *direct* evidence such as that gained from witnessing or participating in the event, *indirect* evidence that serves as the

[1] Or in the case of a question, in some languages (including Tibetan, as we shall see below), the point of view of the addressee.

basis for an inference and/or ego evidence, the evidence someone has simply in virtue of embodying the state. For instance in Turkish, a single indirect evidential can encode a variety of different kinds of indirect sources of evidence. In Turkish the morpheme *-mIş* can reflect either hearsay or other forms of indirect evidence. Tibetan, however, is unusual in that it distinguishes sharply between two different types of indirect evidence, marking each with a distinct morpheme. This allows us to investigate in more detail children's understanding of the inferential bases of others' assertions.

There are at least two important theoretical reasons to study the acquisition of evidentiality beyond its intrinsic interest as a psycholinguistic phenomenon. First, evidentiality is only present in about one quarter of the world's languages, and apparently only in those related in some way to central Asian language groups. Secondly the acquisition of evidentials is anomalous in certain respects: receptive competence appears to follow rather than to precede expressive competence, and final mastery appears only near the onset of adolescence. Therefore, it may well be that the acquisition of evidentiality involves more interaction between general cognitive mechanisms and the language faculty than does the acquisition of other grammatical features.

Thirdly, the mastery of an evidential system—particularly one as rich as that represented in Tibetan—requires an understanding of the structure of inference and source of evidence, and the ability to attribute or to withhold attribution of inferences of various types to others and to oneself. Studying the acquisition of evidentials hence may give us a window into the development of children's understanding of these epistemic concepts. Thus, an exploration of the developmental track of evidentiality may shed light on aspects of the development of theory of mind in its broadest sense that are not

illuminated by the more usual study of the mastery of the attribution of propositional attitude verbs.

1.2 Two indirect evidentials

Tibetan is a Tibeto-Burman language spoken throughout the Tibetan plateau and also in exile Tibetan communities, including the large Tibetan exile community in India and Nepal. It is a head-final language, in which adjectives follow nouns, there are postpositions instead of prepositions, and the verb is at the end of the sentence. Almost every sentence has a copula or verb of existence as well as the lexical form encoding the particular verb meaning, and the copula is the form that encodes the marking of evidentiality.

The Tibetan evidential system comprises three categories: direct evidence, indirect evidence and ego evidence (immediate reflexive knowledge). Within each of these categories several different morphemes encode further subtle distinctions.

Evidentiality is marked in Tibetan by a post-verbal morpheme. In cases where the predicate has no other verb, this morpheme functions as the copula as well as the evidential marker. Most contemporary scholars agree that are three categories of evidence that may be marked, 'ego', 'direct' and 'indirect.' The morphemes marking these categories are shown in (1).

(1) ego: yin, yod
direct: 'dug, song, shag
indirect: red, yod gi red, yod sa red, yin sa red

neutral (or general knowledge): red, yod red[2]

Here we will describe in general terms the nature of ego, direct and indirect evidence.

Ego evidentials are used to report what Garrett (2001) calls "immediate knowledge," that is, knowledge that the speaker simply has because it pertains directly to the speaker, as a property or possession of the speaker. Ego evidentials are used to self-attribute properties or possession. Assertions with ego evidentials are restricted to those with first person subjects, since one cannot have immediate knowledge of someone else's personal experience.[3]

Direct evidentials are used to report eventualities that the speaker knows from having witnessed the eventuality with one of his senses. DeHaan (1999) characterizes direct evidence in general as involving events that are "in the same deictic sphere" as the speaker.

Indirect evidentials are used to report eventualities that the speaker knows from having made an inference from another related state of affairs. (For a formal semantic account of evidentiality see Kalsang et al (2013).)

Adult speakers of Tibetan reliably distinguish two distinct indirect evidentials, *yod kyi red* and *yod sa red.* Each is used for reporting an assertion about a situation that the speaker did not directly witness. The forms represent a distinction between two different kinds of inference. On the one hand, one may infer that a state of affairs obtains

---

[2] Although these obviously have morphemes in common, there does not at this time exist a compositional analysis of the different meanings, so we will adopt the traditional approach and consider these to be distinct lexical items.

[3] Questions anticipating an ego-evidential in the answer also use ego-evidentials, in virtue of the reflection principle for questions in Tibetan. See de Villiers & Garfield, 2017).

based on general information, but not on the basis of any specific perceptible evidence directly related to that situation. Call this *general inference.*

On the other hand, one may infer that a state of affairs obtains on the basis of perceiving some specific evidence that is a reliable indicator of that situation. Call this *specific inference.* Suppose, for instance, I report that my friend Sonam is in his office. I might assert this on the grounds that his office hours are at this time, and that he is punctual about keeping them. This is an instance of *general inference.* On the other hand, I might *see* the light under his door, and his umbrella leaning against the wall in the corridor. Note that in this case I do not *see* Sonam in the office, and so I only *infer* that he is there. But in this case, my inference is based on specific perceptual evidence relevant to the situation. My assertion is hence in this case grounded in *specific inference.*

In Tibetan, *yod sa red* is the indirect evidential that marks *specific inference. Yod kyi red* marks general inference. One way of getting at the difference in their felicity conditions is by reference to the direct perceptual evidential *'dug.* If I report that Sonam is in his office using *yod sa red,* and you ask me why I say that, I can respond by saying something like, "his umbrella is in the hall (*'dug)".* That is, I can refer to a related state of affairs on the basis of which I am warranted in asserting using a direct evidential. If I use *yod kyi red* on the other hand, I implicate that no such state of affairs is available.

Suppose, for instance, that I say:

Tsi tsi pha gir yod kyi red

Mouse over there is [general inference]

(There is a mouse over there.)

I am asserting that there is a mouse over there, and letting my listener know that my grounds are general. It may be that there is a mouse who visits that spot at this time every day, or my daughter might have told me that she saw one there. If you ask me about my grounds, neither of these situations can be reported using *'dug.*

On the other hand, suppose that I say:

Tsi tsi pha gir yod sa red

Mouse over there is [specific inference]

(There is a mouse over there.)

The kind of situation that would warrant this would be my seeing the cat preparing to pounce, or fresh mouse footprints leading to the spot in question. I would report these by saying:

Zhi mi 'dis ri dwags tangs kyi 'dug.

(Cat that [instrumental] hunting is [direct-evidential]

(That cat is hunting.)

Or

Tsi tsi'i rjes pha gir 'dug

(Mouse's tracks over there are[direct-evidential]

(There are mouse tracks over there.)

Adult Tibetans regularly mark this distinction. But given its abstract character, the fineness of the distinction, and the unavailability in the conversational situation of the epistemic situations that warrant the choice of these evidentials, it is surprising that children *ever* learn to use them. Kalsang et al (2013) demonstrate that the semantics of the evidential system in Tibetan can only be represented by a rather complex situation

semantics in which both relations of accessibility and inclusion must be computed. It is not surprising that mastery of a system encoding such complex relations takes time. It is to the acquisition problem that we now turn.

1.3 The learning problem and its theoretical importance

The acquisition of propositional attitude verbs reveals a great deal about linguistic and cognitive development because the verbs are abstract, refer to invisible acts or mental states that must be inferred from behavior, and introduce complex grammatical constructions. For this reason, the acquisition of mental state verbs has been tightly linked to the cognitive developments that support the acquisition of theory of mind in the preschool years (Shatz, Wellman & Silber, 1983; Bartsch & Wellman, 1997; Astington & Baird, 2005; de Villiers & de Villiers, 2000; Milligan, Astington and Dack, 2007). The acquisition of evidentials promises to be revealing for the same reasons: these morphemes encode information about abstract epistemic relations between situations of utterance, evidence situations and reference situations. They thereby convey information about the epistemic states of utterers, states even more difficult to infer, since the facts that determine the felicity conditions of evidentials are absent from the conversational context. This suggests that children must distinguish between a range of highly abstract relations in order to understand and to use these morphemes correctly, and that they must encode this information grammatically if they are to use evidentials as information regarding the epistemic states of others or to use them felicitously to represent their own epistemic states.

But there are important disanalogies between propositional attitude verbs and evidentials as well. As we noted above, evidentials are egophoric: a child's use of an evidential encodes the evidence the *child* has for her statement, not information about other's knowledge, while propositional attitude attributions often convey information about other's mental states. The felicity conditions of evidentials, as Kalsang et al show, reflect the relations between discourse referents and evidence situation, and not the truth value of the sentence in which they occur or the speaker's degree of assent thereto. As a consequence, learning to understand and to produce evidentials appropriately may require not only Theory of Mind, but also sophisticated meta-cognitive reasoning. A child must not only know how to make inferences, but must also be able to recognize and to categorize the inferences that others have made and reported using evidentials. The process by which a child narrows down the possible meanings of even simple direct evidentials is hence likely to be quite complex (for discussion, see de Villiers, Garfield, Gernet-Girard, Roeper & Speas, 2009).

Previous studies of the acquisition of evidentials reveal three notable phenomena. First, some evidentials appear felicitously in spontaneous speech well before children demonstrate reliable *comprehension* of the meanings they convey. Second, when production is elicited in controlled conditions, children show earlier control of direct evidentials than of indirect evidentials. Third, 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. How can we explain the anomaly that productive competence with evidentials precedes receptive competence with evidentials, and even precedes the understanding of the information these evidentials apparently encode?

One possibility is that early production does not accurately reflect children's knowledge of the conditions of use. Evidential morphemes appear early in the speech of children. The morphemes carrying direct and indirect evidential meaning occur in 2 year olds speaking Turkish (Aksu, 1978), Korean (Choi, 1995) and Tibetan (de Villiers et al, 2009?). However, being obligatory in adult speech, these morphemes are extremely common in the input, and the child who uses them may not necessarily map them onto the right adult meanings.

In Tibetan, they are copulas, hence pervasive in the language. In Turkish, they are also past tense morphemes, and children may only gradually connect them to evidentiality (Aksu-Koc, 1988). Without careful control of events and knowledge access, it is very hard to detect errors in spontaneous speech in everyday conversation with small children. The problem is that with two forms, one may appear to be mastered if it is used correctly whenever it is called for (e.g. *-DI*). But it is possible that these apparently correct uses simply reflect a bias to use that form. This will be apparent if this same form is also (-mIş) used when the other form is more appropriate. For this reason, experiments have recently revived the use of d` (d-prime) from signal detection theory (Green and Swets' (1966), to consider not only successes ("hits"), but also failures ("false alarms") (Evans, Viele, Kass, & Tang, 2002). Only by meeting a criterion of sensitivity is a form considered mastered. Hence it is possible that the age of mastery of the direct/indirect

distinction can be underestimated by attending to only the correct use of the direct morpheme, ignoring its incorrect uses.

Elicited production is a more reliable indicator of competence than measures of the emergence of the forms in spontaneous speech. Clever protocols have been devised to elicit evidentials in more controlled situations where their adequacy can be judged. Aksu-Koç (1988) showed Turkish children pictured stories in which a target event (e.g. the popping of a balloon) was either explicitly shown or had to be inferred from the perceived outcome (e.g. the popped balloon). Asked to relate the story, only after the age of four did her subjects prefer the Turkish direct evidential *-DI* for directly perceived events and the Turkish indirect evidential *-mIş* for inferred events. Ozturk & Papafragou (2007) used elicited production methods with Turkish speaking children to study the difference between the evidential markings of *see* and *hear about* versus *see* and *infer*. In production, children clearly controlled the form *-DI* (direct evidential) before they controlled the indirect (*-mIş*), the use of which was only appropriately mastered around age 7.

The data from comprehension studies amplify the concerns about the meaning attached to the morphemes by young children. Despite apparent control of the morphemes in speech, comprehension of others' use of the evidential morphemes has been found to be surprisingly difficult for children speaking these languages. In comprehension, Aksu-Koç (1988) asked her child participants to decide whether a doll who reported an event using *-mIş*/*-DI* had seen the event or had been told about it. Even six year olds had trouble using the evidential morphology of another speaker to make consistently correct inferences about source knowledge. In later work, Aksu-Koc¸ and Alici (2000) confirmed that six year olds could not differentiate the reliability

communicated by Turkish evidentials. For Ozturk and Papafragou's (2007) participants, comprehension was also later than production and showed the same difference across types of evidential, with children being better able to understand who saw something (-*DI*) than who either heard about or inferred something (-*mIş*).

Choi (1995) reported early appearance of the Korean evidential morphemes *–e* and *–ta*, signifying direct evidence, and -*tay* marking hearsay, just as Aksu-Koç had reported in the case of Turkish children. Papafragou, Li, Choi and Han, (2007) studied evidential use by three and four year old Korean speakers, with results similar to those reported for Turkish children. Papafragou et al set out to discover whether the meanings children assign to evidentials were the same as adult meanings, in experiments on elicited production and comprehension. For example, they asked children to report what puppets had said to them (using *–tay*) and to correct lies told by puppets (changing *–tay* to *–e*). In comprehension, they had to decide which of two puppets had uttered a particular sentence containing an evidential by remembering which puppet had directly witnessed or just heard about the contents of a box.

In this task, the children watched two dolls interact with a box in different ways. One looked in the box and the other was told about it in a whisper by another character. The dolls were then hidden by a curtain, and a disembodied voice emerged saying a sentence containing either the direct or the indirect evidential. The child's task was to assign the utterance to one of the two dolls. Papafragou et al. also used a truth-value-judgment task to evaluate a puppet who sometimes mis-used evidentials by mismatching them to his source of knowledge. Papafragou et al. find the emergence of a contrast between *–e* and *–tay* in production at age 3-4 years, but no evidence of comprehension in

either task used. They argue that since production requires only the child's perspective, it is less taxing than explicit judgment about someone else's perspective.

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. The child's "comprehension" of the evidential, they argue, is clearly being judged by an unusually high standard. For these reasons, researchers have argued that the children really know the meanings of the evidentials at an earlier age, but that their competence is being masked by the sophisticated task demands.

But it is also possible that the production data are consistent with the comprehension data. It may be that children, despite the use of these morphemes in their speech, are *not* fully in control of the meanings that evidentials convey until quite late in development, regardless of modality of testing. 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. Therefore, productive competence in evidentials may not be as clear as their early appearance suggests. Ünal and Papafragou (2016) argue against a methodological explanation for the asymmetry between production and comprehension, and instead emphasize that productive competence does not entail receptive competence.

Studying children acquiring Tibetan, de Villiers et al. also found that the direct evidential *'dug* emerges early in children's speech, at around age three years. The indirect forms are rarely used until age five years or so in spontaneous speech, or even in contrived circumstances designed to control the child's access to visual information. Comprehension tasks revealed the late development of a systematic contrast between direct and indirect evidentials, emerging only after age eight years or so. The authors proposed the following stages in the acquisition of Tibetan evidentials:

i) The child takes *'dug* to mark "here-and-now visible to both speaker and listener."

ii) The child takes *'dug* to mean speaker egophoric certainty.

iii) The child takes *yod kyi red* and *yod sa red* to mean "less certain."

iv) The child takes all three evidentials to mark speaker source of evidence.

The explorations of Tibetan in that study also shed light on the third claim in the existing literature on evidential acquisition to which we refer above. Recall that researchers studying Turkish and Korean argue that delays in comprehension or productive use of evidential contrasts cannot be due to conceptual delays, on the grounds that children show evidence of understanding source knowledge around age four years. But the only evidence for *this* claim is that children can answer appropriately when asked

questions such as "Which doll saw inside the box? Which doll knows what's inside the box?" (by having seen), etc.

We argue that the use of such a task to measure the relevant cognitive skills severely underestimates the conceptual underpinnings of evidential understanding. Such skills may only suffice to underpin the correct and fortuitous use of the direct perception evidential, without even representing the fact that it is an evidential That is, taking it merely to mean "I saw that," is not actually representing an evidential system at all, which requires representing the different possible evidentials as a range of alternatives, and understanding the set of meanings they encode. However even *that* conclusion is not obvious given the confound of certainty and knowledge source (see discussion in de Villiers *et al*). For a true understanding of inferential (indirect) evidentials, a child must be able to understand the inferences other people are capable of drawing from circumstances and the *difference* between perception and inference. There is surprisingly little literature on when or how children make inferences from signs that something has happened, for example drawing the inference that someone has eaten breakfast from having seen eggshells in the sink, or that a big brother is home because his bike is lying on the path. Yet this is precisely the kind of inference required to use the evidential *yod sa red* correctly in Tibetan.

In the de Villiers *et al* study, a conceptual task used a children's book, *Anno's Hat Tricks* (Nozaki & Anno, 1985) . The participants are introduced to the characters of a hatter, Tom, Hannah[4], and the shadow of a child, which they are instructed to take to be their own shadow. Children were shown pages of the *Shadowchild* book which posed

[4] The characters were renamed with Tibetan names for the testing of Tibetan children.

increeasingly difficult inferential challenges. The children were asked to solve these puzzles and this provided evidence both of children's inferential sophistication, and in their explanations, of their metacognitive understanding of their own inferential ability. On each page, the hatter shows a different combination of red and white hats, and by looking at that combination and at Tom's and Hannah's hats, the participant is asked to deduce the color of her own hat, shown on the page as a shadow.

For example, the hatter says that he is taking two red hats and one white hat and placing one on Tom and one on Shadowchild (the shadow). The participants then learn from Tom that he knows the color of his own hat to be red. (see Appendix A)

Sample Question: What color is your hat, Shadowchild?

The participant must infer that Tom saw that the participant's hat was a color that allowed him to ascertain that his own hat was red. Had Tom seen a red hat, then another red hat and a white hat would have remained, and in that case he would not be able to determine the color of his own hat. Since Tom reported that he *did* know the color, he must have been able to see that Shadowchild's hat is white.

The book presents progressively more difficult scenarios, and children were tested on all but the most challenging. This task did not require children to know evidential meanings. However, notice that in order to solve these problems one must attend not only to one's own evidence, but also to what other people say in response to *their* own evidence. This is the ability that is required in order to use inferential evidentials appropriately because Tibetan distinguishes between specific and general inference, and so requires children to represent the precise premises they use in coming to the judgment reported in the sentence involving the evidential.

The evidence from that study revealed that the ability to reason from other's inferences was strongly correlated with the children's likelihood of correctly using an indirect evidential - *yod sa red,* an evidential the mastery of which is rather late in the developmental sequence- in a completely different task. It appears that inferential and metacognitive capacities and the mastery of the evidential system go hand in hand, providing some evidence for the view that acquiring mastery of the evidential system involves an interaction between grammatical and more general cognitive capacities.

In this paper we discuss a further challenge: at what age can a child differentiate not just direct from indirect sources of evidence, but two different kinds of indirect evidence, based on different inferential premises? Just as in the case of simple direct evidentials, the learning path is complex: in order to learn the meaning of the evidential used around her, a child must not only be able to draw such an inference but recognize when someone else is implicating that the reference situation[5] is accessible from, but not contained in, the information situation that justifies their statement. This entails that that *other party has drawn an inference* on the basis of perceptual signs, not on the basis of more general information. Furthermore, the two kinds of inference must be differentiated, with the more abstract form, *yod kyi red*, being used when there is no immediate perceptual sign, but some other source of evidence such as habit, hearsay, or general agreement.

[5] i.e. the situation that the sentence asserts to obtain.

In our experimental work on the acquisition of evidentials in Tibetan, we ask two questions, in order to understand the cognitive development that underlies children's understanding of the distinction between specific and general inference:

a) When can Tibetan children differentiate between the two forms of indirect evidential and show that they understand the different bases of inference?

b) Do they use one form before the other?

2. Methods

2.1 Stimuli

The task comprised two parts: The first part uses nine stimuli, each of which is an illustrated short story about two characters, Dolma and Tashi. In each story, an event occurs, or a situation obtains, that neither the participant nor either character witnesses directly (e.g., in one case, we are told that mom is holding a surprise present behind her back; in another we are told that a loud noise was heard coming from the other side of the fence). In each case, one character reports what occurred using *yod sa red,* and the other reports what occurred using *yod kyi red.* The participant is asked, "Who saw something?" The correct conclusion is that the character who used *yod sa red* based the inference on a perceptual sign. This forced choice task is augmented by asking the participant to explain his/her answer. Any fully competent adult speaker of Tibetan would recognize that the use of the specific versus the general inference evidential signals having seen a particular piece of evidence. Children's explanations of their responses give evidence of their metacognitive understanding of the conditions that led to their choice.

The second part used six stimuli, each of which is an illustrated short story in which a single character draws an inference about a situation not currently perceptible. In half of the cases the character reports the conclusion of the inference using *yod sa red,* and in half the conclusion is reported using *yod kyi red.* In each case, the child is asked "why did he say that?" This task allowed us to examine children's sensitivity to the felicity conditions of the two evidential forms.

In addition, in order to enable us to measure linguistic sophistication we presented participants with three wordless picture narratives, one drawn from the DELV (Seymour, Roeper & Villiers, 2003) and two similar stories. The participant was simply asked to tell us the story represented by the sequence. All three stories involved complex social situations designed to elicit the use of mental state verbs, specific referential devices and temporal language. See appendix D for sample stimuli. All stimuli were presented by native Tibetan speakers using a laptop computer.

2.2 Participants

Participants were 32 Tibetan children age 4-11 years (mean age 7 yrs, 9 mo), 17 boys, 15 girls, each of whom was a native Tibetan speaker. Children were drawn from two primary schools serving the Tibetan exile community in Dharamsala, India. None had any evident disability, and all attended regular classes.

2.3 Procedures

Each child was tested individually in one or two short (approximately 15 minute) sessions by two native Tibetan speakers, one male, one female, in a relatively

quiet area of their school. Stimuli were presented on a laptop computer. All sessions were videotaped and later transcribed by a native Tibetan speaker and entered into a Filemaker Pro database for scoring. In each case, children were presented with the stimuli, and were asked the relevant questions by the experimenter. Of the 32 children, only 25 contributed narratives due to time constraints.

3.Results

3.1 Scoring

Part I was scored as follows: Each participant received 1 point for answering each initial question correctly (i.e. choosing the character who used *yod sa red* as the one who saw something). Participants were also scored for appropriate use of evidentials in their explanations. As a consequence, each child could receive a maximum of nine points for correct evidential comprehension, and a maximum of nine points for correct evidential production.

Part II was scored as follows: Each participant received one point for each cogent explanation of why the character in the scenario used the evidential s/he did. For example, the following is a cogent explanation for the use of *yod sa red:* The child is asked why the character in the story said, using *yod sa red,* that there is a cat*:*

ga re byas nas lab na...zhi mi gis skad nas meao...lab red pa, 'o 'di byas nas.

Because it said… cat [instrumental]said meow… it said. That's why.

"Because the cat said 'meow,' that's why."

The child here cites perceptible evidence – the noise – to explain the inference to an unseen cat, and hence explains the use by the character in the story of the *yod sa red* evidential.

On the other hand, the following is *not* a cogent explanation of the character having used *yod sa red*:

de a' sem chen zshan pa red bsam nas ani 'di yod kyi red lab red

there animal different [past] thought [ablative] and is [general inferential] say [past].

"Having thought it was a different animal, he said it is."

Leaving aside the general incoherence of the response, there is no mention of evidence at all.

In addition, each participant received one point for appropriate use of evidentials in his/her explanation. As a consequence, each child could receive a maximum of six points for correct evidential comprehension, and a maximum of six points for correct evidential production. The scoring yielded a measure of success for the judgment task for each evidential: 0-3 on *yod kyi red*, and 0-3 on *yod sa red*.

In order to derive a rough index of linguistic sophistication, a score of mean verbs-per -sentence was derived from the three transcribed narratives. Mean length of utterance is inappropriate for a language like Tibetan which permits a great deal of elision and deletion, and in any case loses its utility for children above age four (Brown, 1973). Mean verbs-per-sentence has promise as a more linguistically neutral index of linguistic sophistication (Deen, 2005).

3.2 Data

We first report the data from the first task on receptive understanding of the distinction between the two indirect evidentials. Participants were divided into three age groups (Group 1:4-7 years, mean =5.38, n=13; Group 2: 8-9 years, mean = 8.50, n=10; Group 3: 10-11 years, mean =10.33, n=9). An ANOVA taking percentage correct as the dependent variable, with age group as the independent variable revealed a significant age effect ($F[2,29]=10.66$, $p<.001$, partial $eta^2=.424$). See Figure 1.

Figure 1 here

A single sample t test comparing the means of each group with chance performance revealed that Group 1 was significantly worse than chance (primarily due to the number of non-responses), mean=.15, sd=.28, $t(12)=-4.39$, $p<.001$; Group 2 was not significantly different from chance, mean=.50, sd=.34, $t(9)=.009$ (n.s.). Only Group 3 exceeds chance significantly, mean=.74, sd=.28, $t(8)= 2.62$, $p<.04$.

We next report data from the second task examining the reasons children give for the use of each of these evidentials. Using the same age groups, a 2X3 repeated measures ANOVA taking evidential type as the repeated measure and age group as the independent variable revealed a significant age effect ($F[2,29]=5.2$, $p<.012$, partial $eta^2=.264$) as well as a significant type effect, with percentage correct on *yod kyi red* significantly better than percentage correct on *yod sa red* ($F[1,29]=5.37$, $p<.03$, partial $eta^2=.156$). See figure 2.

Figure 2 here

We now report correlations between performances on the task in Part 1 and the task in Part 2 . There is significant partial correlation, controlling for age, between correct answers on these two tasks ($r(29)=.480$, $p<.007$). There is a significant partial correlation as well, also controlling for age, between participants' correctness on the first task, asking "who saw something" and participants' use of appropriate evidentials in their explanations of their answers ($r(29)=.536$, $p<.003$), indicating that receptive and expressive use of these evidentials are strongly related. There is a weaker, but still significant, correlation between correct use of evidentials in explanations of answers and performance on the second task, in which participants are asked why an evidential was used ($r(29)=.363$, $p<.05$).

There are also correlations, again controlling for age, between correct use of evidentials in explanations and participants' linguistic sophistication as measured by mean verbs per sentence ($r(18)=.436$, $p<.055$ in the first study and $r(18)=.453$, $p<.05$ in the second). No other measures correlate significantly.

4. Discussion

The first experiment - examining the development of the ability to differentiate between the meanings of the two indirect evidentials - showed that children's understanding of this distinction increases significantly as they get older. In fact, children do not perform better than chance on demonstrating an understanding of the distinction between *yod kyi red* and *yod sa red,* until age 10 or 11 years, long after Theory of Mind and basic inference abilities are in place.

The second experiment confirms that sensitivity to this distinction develops late, but reveals more detail in the acquisition track. Tibetan children master the use of *yod kyi red* before they do *yod sa red*. This result, in tandem with those of de Villiers et al 2009, suggests that Tibetan children initially develop a bipartite understanding of the Tibetan evidential system as marking only the distinction between *'dug* (direct) and *yod kyi red* (indirect). At this stage, they use the most general indirect evidential, *yod kyi red*, in all situations which call for an indirect evidential. It is only later, at about age 10, that they recognize the subtle distinction between *yod kyi red* and *yod sa red*. The distinction between direct and indirect evidentials only requires an understanding of whether another person was witness to an event or not. The *yod kyi red/yod sa red* distinction, however, requires the listener to understand the *nature* of the inference that the speaker draws for his/her claim and the relation between the evidence and the claim for which it is evidence.

Taking these tasks together, and examining the correlations between performance on them suggests further conclusions. We see that metalinguistic understanding of the distinction between the two indirect evidentials is strongly correlated with appropriate expressive use, and that expressive correctness is strongly correlated with general linguistic sophistication, both controlling for age. While this may appear to be at odds with earlier findings that expressive use precedes receptive use, it may well not be, and this for two reasons. First, it is highly likely that much early felicitous use of evidentials does not reflect a full grasp of their meaning, but might simply reflect a confusion of evidentials with epistemic modals, and indeed there is a sufficient overlap between the felicity conditions of statements with evidentials and the

truth conditions of statements using the epistemic modals with which they might reasonably be confused, that this could account for apparent early expressive competence. Moreover, we are examining the mastery of a subtle distinction between inferential evidentials, and there is no prior evidence that the felicitous use of these evidentials precedes their comprehension.

This also suggests a refinement of the developmental sequence sketched in de Villiers *et al*., (2009). We propose that the final stage, in which the child takes all three evidentials to mark speaker source, in fact resolves into two distinct stages. It is only after the child represents the fact that *yod kyi red/yod sa red* are indirect evidentials that the child comes to represent the *distinction* between the kinds of indirect evidentially they encode. So the developmental sequence we propose runs as follows:

i) The child takes *'dug* to mark "here-and-now visible to both speaker and listener."

ii) The child takes *'dug* to mean speaker egophoric certainty.

iii) The child takes *yod kyi red* and *yod sa red* to mean "less certain."

iv) The child realizes that *'dug* and *yod kyi red/yod sa red* are evidentials, representing the evidential system as a whole.

v) The child distinguishes among types of indirect evidence.

It thus appears that the understanding of the complete evidential system in Tibetan is a late stage in general linguistic development, requiring metalinguistic capacity, metacognitive reasoning, and a mature understanding of the nature of inference. This suggests that the acquisition of evidentiality draws on a much broader suite of general cognitive capacities than does most other grammatical development. This fact is

probably explained by the complex semantics of the evidential system explored in Kalsang et al (2013).

References

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

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, & Alici, D. M. (2000). Understanding sources of beliefs and marking of uncertainty: Child's theory of evidentiality. In E. V. Clark (Ed.), *Proceedings of the 30th Annual Child Language Conference* (pp. 123-130). Stanford: Stanford University.

Aksu-Koç, A. (1988). *The Acquisition of Aspect and Modality: The Case of Past Reference in Turkish*: Cambridge University Press.

Astington, J. W., & Baird, J. A. (2005). Representational development and false-belief understanding. In J. W. A. J. A. Baird (Ed.), *Why language matters for theory of mind* (pp. 163-185): Oxford University Press.

Bartsch, K., & Wellman, H. (1997). *Children Talk About the Mind*: Oxford University Press.

Brown, R. W. (1973). *A First Language: The Early Stages*. Cambridge: Harvard Univ. Press.

Choi, S. (1995). The development of epistemic sentence-ending modal forms and functions in Korean children. In J. Bybee & S. Fleischman (Eds.), *Modality in Grammar and Discourse* (pp. 165-204). Amsterdam: John Benjamins.

Davis, C., Potts, C., & Speas, M. (2007). The Pragmatic Values of Evidential Sentences. In M. Gibson & T. Friedman (Eds.), *Proceedings of Semantics and Linguistic Theory 17*. Ithaca: CLC Publications.

Deen, K. U. (2005). *The Acquisition of Swahili*. Amsterdam/Philadelphia: John Benjamins.

DeLancey, S. (1986). Evidentiality and Volitionality in Tibetan. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The Linguistic Coding of Epistemology* (Vol. 20, pp. 203-213). Norwood, New Jersey: Ablex Publishing Corporation

de Villiers, J. & Garfield, J. (2017) Evidentiality, Questions and the Reflection Principle in Tibetan: What do Children Learn when they Learn About Evidentiality? In F. Nihan Ketrez, Aylin C. Küntay, Şeyda Özçalışkan, Aslı Özyürek *(eds) Social cognition and environment in Language development: Studies in honor of Ayhan Aksu-Koç.* John Benjamins series Trends in Language Acquisition

de Villiers, J., Garfield, J., Gernet-Girard, H., Roeper, T., & Speas, M. (2009). Evidentials in Tibetan: Acquisition, Semantics and Cognitive Development (pp. 29): Smith College.

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

Evans, J. L., Viele, K., Kass, R., & Tang, F. (2002). Grammatical morphology and perception of synthetic and natural speech in children with specific language impairments. *Journal of Speech, Language, and Hearing Research, 45*(2), 494-504.

Garrett, E. (2001). Evidentiality and Assertion in Tibetan. Los Angeles: UCLA.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Kalsang, Speas, M., Garfield, J & de Villiers, J.G. (2013) Direct Evidentials, Case, Tense and Aspect in Tibetan. *Natural Language and Linguistic Theory,* Volume 31, Issue 2, Page 517-561

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.

Nozaki, A., & Anno, M. (1985). *Anno's Hat Tricks*: Philomel Books.

Ozturk, O., & Papafragou, A. (2007). The Acquisition of Evidentiality in Turkish, *31st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Papafragou, A., Li, P., Choi, Y., & Han, C. (2007). Evidentiality in language and cognition. *Cognition, 103*(253-299), 253-295.

Papafragou, A., & Ozturk, O. (2007). Children's acquisition of evidentiality, *31st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Seymour, H. N., Roeper, T. & de Villiers, J. G. (2003). DELV-ST (Diagnostic Evaluation of Language Variation) Screening Test . San Antonio TX: The Psychological Corporation.

Shatz, M., Wellman, H., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition, 14*, 301-321.

Speas, M. (2004). Evidentiality, Logophoricity and the Syntactic Representation of Pragmatic Features. *Lingua, 114*(3), 255-276.

Ünal, E., & Papafragou, A. (2016). Production-comprehension asymmetries and the acquisition of evidential morphology. *Journal of Memory and Language* 89: 179-199.

Wellman, H., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child Development, 75*, 523-541.

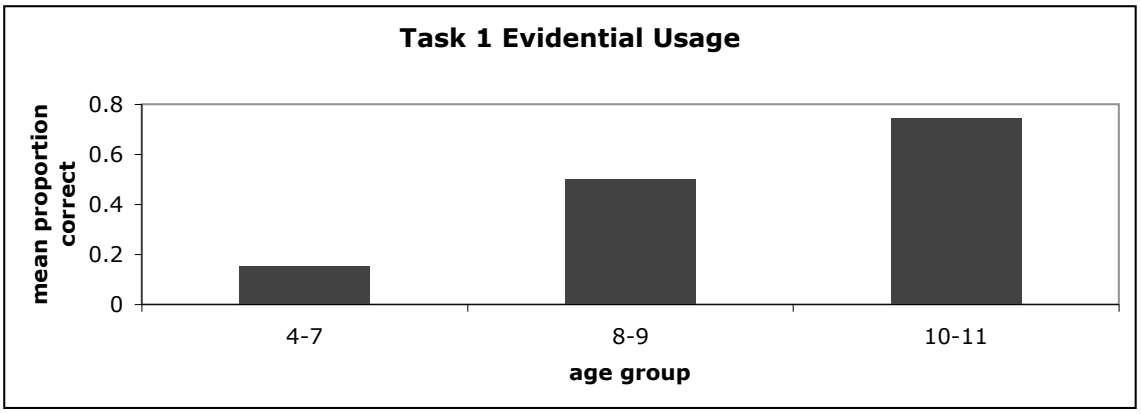


FIGURE 1 Mean proportions of correct evidential usage across groups 1-3.

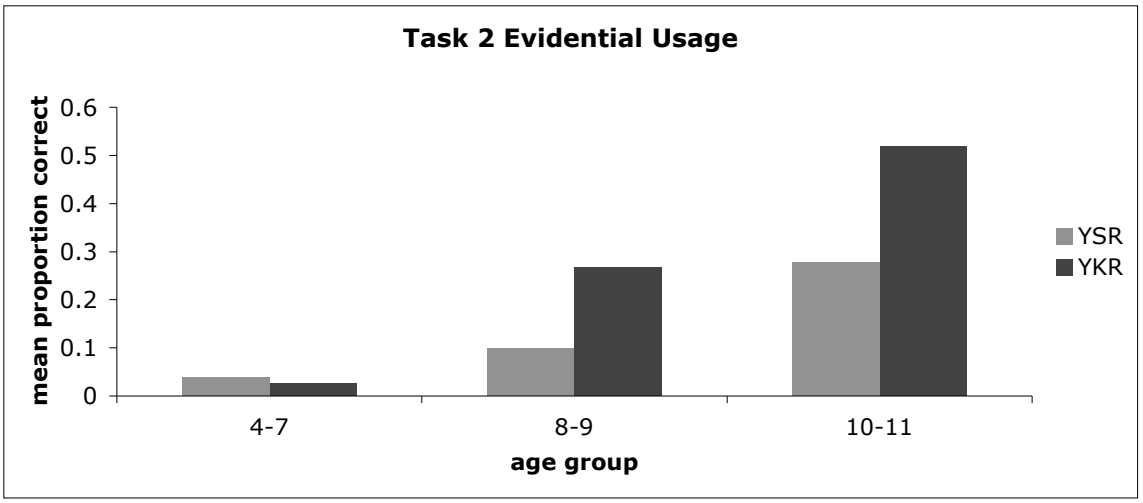
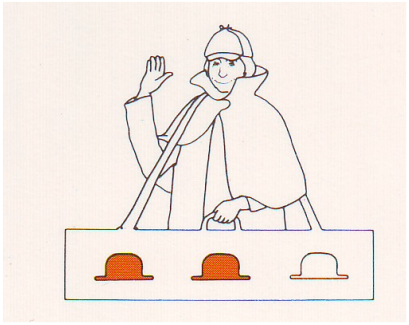


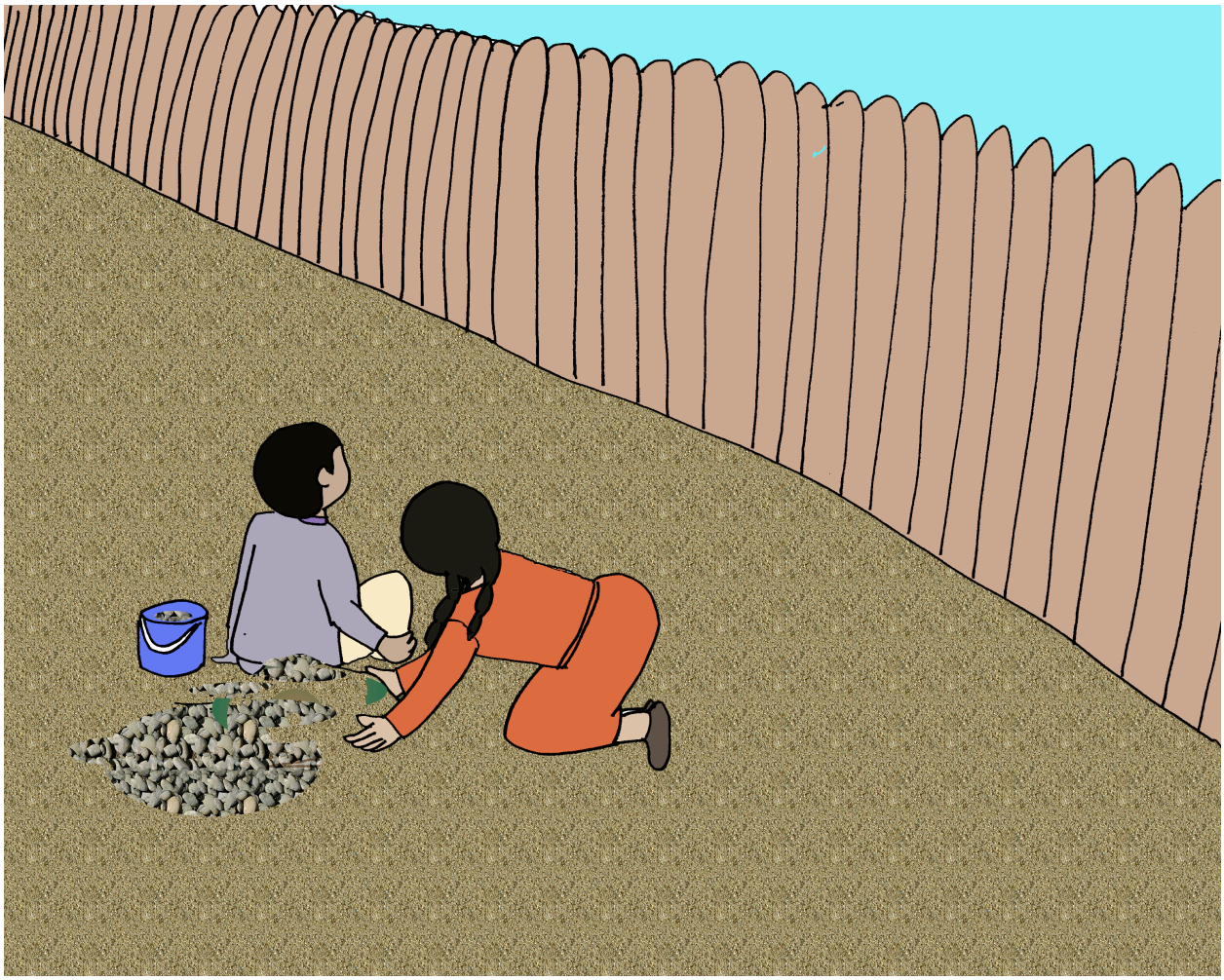
FIGURE 2 Mean proportions of correct usage of *yod sa red* and *yod kyi red* across groups 1-3.

Appendices

Appendix A



Appendix B

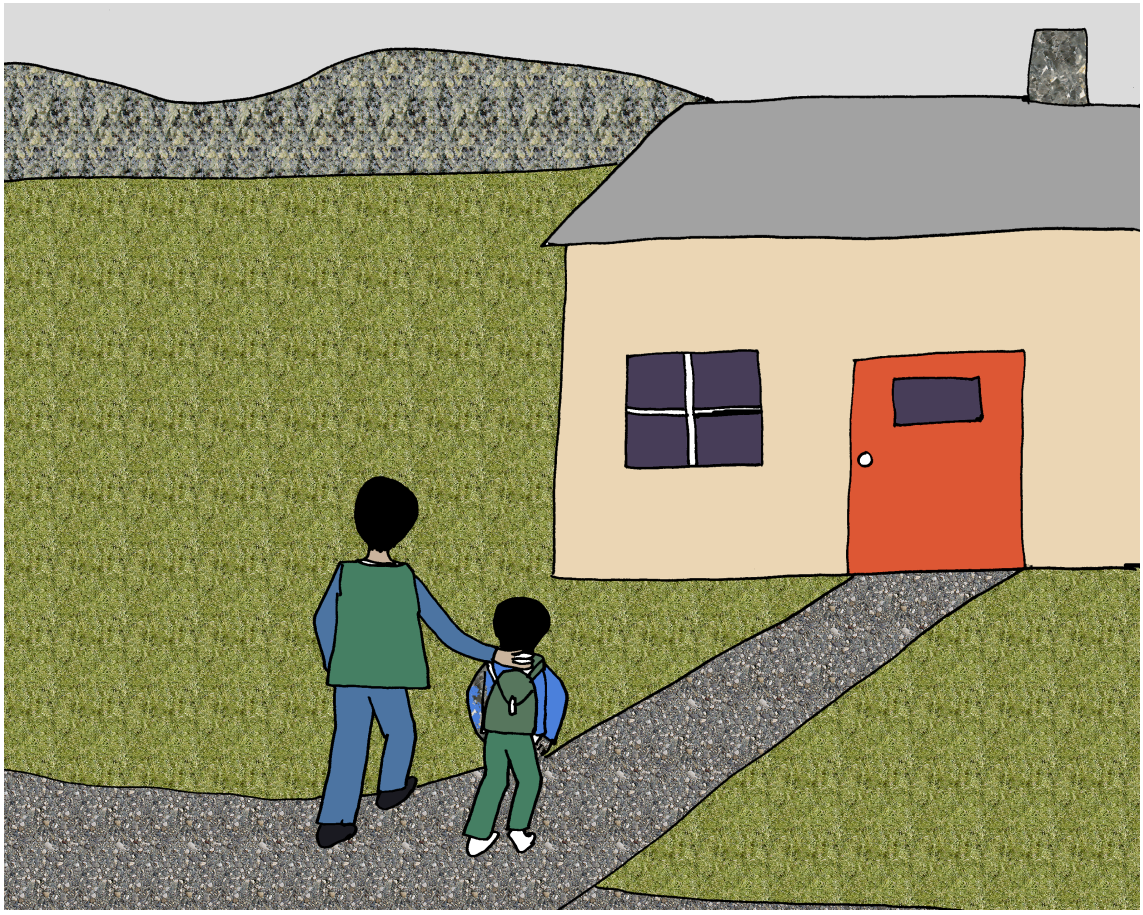


Sam and jenny are in the backyard playing. They hear a loud noise on the other side of the fence.
Sam says, “It looks like it’s a motorcycle making that noise.”
Jenny says, “It must be a tractor making that noise.”

Who saw something?

bsam sgrub dang ‘jigs med gnyis rang khyim gyi rgyabs ngos kyi ldum rar tsed mo tse bzhin gnas la kong gnyis kyi glo bur du ra skor gzhan pyogs nas skad sgra chen po zhig thos skabs
bsam sgrub kyis, “de sbag sbag gis skad rgyabs pa ‘dra po ‘dug”
‘jigs med kyis, “de tan tan rmod ‘khor gyi skad rgyabs pa yin sa red”
su yi gang mthong pa red dam

Appendix C



Andrew's dad drives him home from school every day at 3:00. Andrew and Dad drive up to the house and park the car.
Andrew says, "It looks like Mom is home."

Why do you think he said that?

ag lod dus rtags par chu tshud gsum pa zin pa dang pa' lags kyi slob gra nas nang bar du mo tras skyel gyi yod. ag lod dang kho rang gi pa' lags kyi snum 'khor nang bar du gtang nas sgo rar bkag nyar bus.
ag lod kyi, "a ma nang du tan tan yod sa red"

khyed rang gi bsam par kho'i de ci'i pyir lab pa red dam.

Appendix D