The following article, in the form of the author’s accepted manuscript with page numbers inserted in double brackets, has been published as


The article is under the copyright of John Benjamins Publishing Company, so for any permission to re-use or reprint the material in any form the publisher should be contacted.

[[45]]
Factual vs. evidential? - The past tense forms of spoken Khalkha Mongolian

Benjamin Brosig
Academia Sinica

Past tense forms of spoken Khalkha Mongolian distinguish between established (-sang) and non-established knowledge, which is then either based on direct (-lää) or indirect (-jee) evidence. Time of acquisition thus determines whether information source is marked, though vivid recollection (-lää) and deferred realization (-jee) overrule it. Conversely, attempted recollection in questions (-l=uu) doesn’t presuppose sensory perception. A fourth suffix (-v) is used if well-established events still surprise the speaker. These suffixes may also be used in a discontinuous fashion to refer to the future and then modally qualify predictions as inevitable (-sang), apprehended [but preventable] (-v), based on sensory evidence (-lää) or inferred (-jee). The distinction between unsourced -sang and sourced -lää/-jee is thus not about factual stance, but codes the extent to which information is consolidated in memory.

Keywords: direct evidence, discontinuous tense, indirect evidence, mirative, preventive, time of acquisition

* I’m grateful to Bettina Zeisler, Dolgor Guntsetseg, Östen Dahl, Henrik Bergqvist, Lamont Antieau and Maria Koptjevskaja-Tamm for comments on the draft of this chapter, to all my informants for their enormous contribution, and to all who supported me during this stay and later on. My research was financed by Stockholm University as my employer and the Sven och Dagmar Saléns stiftelse who funded my 7-month stay as a visiting researcher at the National University of Mongolia in Ulaanbaatar with 53270 SEK.
1. Introduction

Khalkha Mongolian has a complex evidentiality system. At the level of present tense, there is an opposition between two suffixes. Of these, -n indicates immediate access to an event either via direct firsthand perception or, mediated by perfect forms or the like, via direct access to the resultant state and indirect knowledge of the event itself. The contrasting -aa is used for events that have been established previously, be it through indirect evidence or previous personal perception (which can be made explicit through the use of additional sentence-final particles). This opposition is obligatory for progressives, and arguably so for resultatives and perfects – thus for all ongoing present situations. Habitual/generic propositions in -dag, mostly taken to be self-evident, are usually not marked for their degree of establishedness, and potential -n is always unmarked (see Brosig 2015a for details). [46] At the level of past tense, the seemingly neutral suffix -sang (which I will argue is used for established events) is used alongside the Direct Past -laa for immediate direct (sensory, participatory, speaker-internal) and the Indirect Past -jee for immediate indirect (inferential or secondhand) access to the event. In addition, the peripheral past tense suffix -v (for past and future) and the suffix -dag in a non-habitual use (for present and past) signal surprising events, though their place in the overall system cannot yet be specified. A sketch of the evidentiality system proper is given in Table 1:

<table>
<thead>
<tr>
<th></th>
<th>established</th>
<th>non-established</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>indirect</td>
<td>direct</td>
</tr>
<tr>
<td>past</td>
<td>-sang</td>
<td>-jee</td>
</tr>
<tr>
<td>ongoing present</td>
<td>-aa</td>
<td>-n</td>
</tr>
</tbody>
</table>

The system contains a number of yet unexplored tensions such as the parallel existence of the Direct Perfect -sang bai-n (with personally perceived result) and the simple past form -jee that signals indirect access, both of which can be used for inference from visual evidence. The basic system interacts with a number of external elements such as the speech complementizer verb ge- ‘say’ (Song 2002; Hashimoto 2004; Yáo 2007) and particles such as bilee for recollection (Brosig 2012), written aj for hearsay/inference, and the official suragtai (< source-COM) for hearsay. Finally, the participles -sang (Established Past), -dag (Habitual), -h (Future) and -aa (non-personal access) can all combine with the clitic =iin (< yum

1 Except for -jee, all suffixes are subject to vowel harmony. For instance, the realizations of what I write as “-sang” as a label are the forms -sang, -seng, -song and -sông depending on the immediately preceding vowel in the stem.
‘thing’), which has an unclear function, while the combined form =iin bai-n possibly indicates general reasoning.\(^2\)

In this chapter, I will restrict myself to describing the past forms of spoken Khalkha: the Established Past -sang, the Direct Past -laa, the Indirect Past -jee and, in addition, the peripheral Modal Past -v. While their meaning has been investigated quite often (e.g. Byambasan, Önöribayan, Pürev-Ochir, Sanjaa, and Janchivdorj 1987; Svantesson 1991; Hashimoto 1993a, 1993b; Kim 1995; Song 1997; Nelson, Davidson, Narantuya, and Nordby 1998; Kalchofner 1999, 2003; Jingan 2010; [[47]] Binnick 1979, 1990, 2012), the respective authors generally didn’t make proper use of available previous research or of spoken corpora. The set of uses they treat is thus generally too narrow. It is also too broad in the sense that evidence for very specific uses restricted to specific genres prevents the discovery of restrictions that obtain in other registers. However, the use of the past forms differs substantially between different genres. -v is widespread in formal written genres and functions as a general, evidentially neutral past. In spoken language, however, it is restricted to a few very specific uses, though this is ignored in part of the literature (e.g. Hashimoto 1993a; Song 1997). -jee is used in the news to relate secondhand information irrespective of whether the journalist just acquired it, and -laa can sometimes be extended to reliable secondhand knowledge (for a similar extension of -laa in a complex construction in Kalmyk, see Skribnik and Seezing 2014: 162).

While journalistic (as well as scientific) uses are fairly transparent to native speakers, they are often not able to explain non-evidential uses of -laa and -jee in narrative prose. -v has ceded most of its spoken uses to -sang (cf. e.g. Binnick 2012: 74), but it is less clear how exactly this suffix is used in different written genres. Due to a strong orientation towards morphological form classes, -sang as a participle is often not treated as part of the past tense system, but rather, e.g., as a perfect (Byambasan et al. 1987; Hashimoto 1993a, 1993b; Kim 1995; Song 1997; Jingan 2010). All past suffixes have also developed future uses that are more common in the spoken language. In legalese Khalkha, still other uses exist. It is not clear to me whether the uses of the past tense suffixes in literary prose follow one congruent system or whether they are the result of imperfect learning by new generations of writers and of dialect mixing. At any rate, it is a widespread delusion to assume that conclusions drawn from literary prose are valid for the oral language.

Given the state of research, the question of whether evidentiality is an obligatory category in Mongolian has rarely been addressed. Binnick (2012: 76–79, 110) seems to favor analyzing -sang as indicative of direct

---

\(^2\) As many particles combine with -aa, but rarely or not at all with -n, it is assumed that they help to subdivide the vast semantic area covered by -aa. It is not currently known to which extent such disambiguation might actually be necessary.
personal evidence, as does Xuèyàn (2013: 181) by applying the term “direct” to both -sang and -laa. Song (1997: 204) and Street assume Khalkha -v respective Middle Mongol -bA to “imply nothing about the speaker’s certainty or source of information” (Street 2009: 132). Kalchofner takes a middle road by assuming that -sang/-v represent an objectiv[izing], “factual” perspective, even though he seems to relegate it by and large to pragmatics. Similar tripartite systems with one arguably neutral marker exist in the Central Asian Turkic languages Kazakh, Kyrgyz, Uyghur and Uzbek, though these are not analyzed along purely evidential lines (Johanson 2003; Straughn 2011: 164–169). Since a strict evidential equation of -sang and -laa can flatly be rejected, serious hypotheses are (i) that -sang (i-v) is “evidentially neutral” and, thus, “non-evidential,” and (ii) that -sang explicitly presents a proposition as “factual” in the sense of valid and unrelativized (as Mushin (2001: 74) can be interpreted). In [[48]] this chapter, I will propose a third explanation, according to which the basic distinction is drawn between established knowledge (-sang), which may implicate a claim to validity, vs. non-established (mostly new) direct (-laa) or indirect (-jee) knowledge, which may relegate the evaluation to the addressee.

This chapter is structured as follows: In Section 2, I give a short overview of the linguistic evidence used in this chapter. Section 3 presents some usage statistics, and the sections 4–7 are devoted to the four past suffixes in past declaratives. Section 8 and 9 briefly address their uses in questions and future reference, respectively. Section 10 concludes the chapter.

2. Linguistic evidence

As the basis of my study, some kind of linguistic evidence had to be selected. Small corpora (as used, e.g., by Street 1963) usually don’t contain enough data to choose one of different competing explanations with sufficient certainty, which leads to underspecification. Examples created by linguists on the basis of structures in other languages (e.g. Shibagaki 2013, reviewed in Brosig 2013a; but even Brosig 2009) or different registers of the same language (e.g. Song 1997) tend to contain artifacts and should thus be avoided. Working with a very small number of informants (e.g. Matsuoka 2008; Binnick 2012, reviewed in Brosig 2013b) puts the linguist at the mercy of their preferences, biases, systematic oversights and uncontrolled sociolinguistic peculiarities. Eliciting a very large number of informant judgments via a questionnaire (e.g. von Heusinger, Klein, and Guntsetseg 2011), on the other hand, is not feasible for hypothesis-building because a set of pre-defined answers or even rating scales would have to be
used. In order to avoid atypical uses from written registers, a new and – due to practical constraints – necessarily small corpus consisting primarily of spontaneous spoken language usage had to be produced. I therefore tried to work with a small, but not extremely tiny corpus, and to partly compensate for the overall small data set through qualitative, open-ended elicitation from a modest number of informants.

The corpus I used spans over 589 minutes and mainly consists of relatively free conversation data from unscripted television programs (see Brosig 2015a for details), with speakers mostly between 35 and 55 years of age. It was transcribed by B. Zoljargal. I then semi-automatically annotated 4244 finite aspectual forms on all tense levels and selected a relevant subset (i.e. most instances of infrequent forms and usually no less than a third of examples for all but the most frequent forms) for elicitation. Internet materials and overheard utterances were occasionally used as well.

Elicitation was mostly done in Ulaanbaatar between April and November 2013 with 42 informants (34 of them female, 27 of them students, average age 26) for a total of 285 hours. Elicitation time was almost evenly divided between three informant groups: 31 people with 1–6 hours each, 9 people with 7–18 hours each, and my two main informants, G. Mönhsaran and D. Solongo, who together did 88 hours. The most active informants shifted the demographic biases of the study somewhat, so that 89% (rather than 81%) of elicitation hours were done with females, 80% (rather than 65%) with students, and the average informant age for each hour of elicitation was 23. During elicitation, informants were presented with contextualized original sentences in transcription and asked about semantic differences between them and similar tense-aspect-evidentiality forms that I considered worth comparing in the given context. Sentences were discussed with four informants on average, but the recurrence of morphemes in different combinations ensured that I could rely on data from more than ten informant judgments for most morphemes. The analysis was done by generalizing from my notes on their answers across both examples and markers. I focused on recurring patterns, ignoring isolated idiosyncrasies in

---

3 Questionnaires that require informants to formulate answers themselves are hard to use with new or modestly trained informants. Provided with instructions on what information to provide and exemplary answers produced by experienced informants, they still couldn’t produce useful data. In the very last phase of elicitation, experienced informants with 10 or more hours of interview-based elicitation were able to give useful answers to questions on grammatical semantics on their own, but this still required follow-up interviews of about one hour each (given a questionnaire that should have taken 3–4 hours to fill in).

4 While this approach yields decent information about meaning and connotation, it doesn’t give any information on how frequently different interpretations emerge. Thus, I have no quantitative data to check my intuition that uses like (1), (11) and (16) below are very frequent, while those in (7) and (12) are rare.
order to discover common principles from which most of their connotation- and context-based interpretations could be explained. While informants naturally differed a great deal in their focus on various details of different conversational situations and their interpretations thereof, no consistent differences suggesting distinct grammars for individual informants (or speakers in the corpus) were discovered with respect to past markers. [[50]]

3. Frequencies

The 10 hours of recordings contain approximately 60,000 words and 4244 annotated finite predicates. Of these, 1248 are past forms: the Established Past -sang (n=904), the Direct Past -laa (n=225), the Indirect Past -jee (n=50) and the Modal Past -v (n=69). Yet, this has to be put into context: as Table 2 indicates, -v is almost irrelevant in declaratives (n=1002), while -laa and -jee play a very minor role in interrogatives (n=215):

<table>
<thead>
<tr>
<th></th>
<th>declarative</th>
<th>interrogative</th>
</tr>
</thead>
<tbody>
<tr>
<td>-sang</td>
<td>75.5%</td>
<td>68.4%</td>
</tr>
<tr>
<td>n=757</td>
<td>n=147</td>
<td></td>
</tr>
<tr>
<td>-laa</td>
<td>19.0%</td>
<td>4.2%</td>
</tr>
<tr>
<td>n=190</td>
<td>n=9</td>
<td></td>
</tr>
<tr>
<td>-jee</td>
<td>4.8%</td>
<td>0.9%</td>
</tr>
<tr>
<td>n=48</td>
<td>n=2</td>
<td></td>
</tr>
<tr>
<td>-v</td>
<td>0.7%</td>
<td>26.5%</td>
</tr>
<tr>
<td>n=7</td>
<td>n=57</td>
<td></td>
</tr>
</tbody>
</table>

This table speaks against excluding -sang from the proper past tense system, while it indicates that the role of -v is marginal at best. But even if we follow the solid idea that -sang is evidentially neutral, while -laa and -jee are not (e.g. Svantesson 1991), it is not clear what motivates speakers to use the latter so sparingly. The situation in questions is different again. And while Table 2 only lists past uses, all past suffixes also have future uses of which 5 instances of -v and up to 26 instances of -laa were attested in the corpus. In the following, I will try to clarify most of these issues.

4. Established past -sang

The Established Past -sang is used to refer to past events that have been saliently established and reified within the memory of the speaker. They might or might not have been learned via the senses, but the way they have originally been established in the memory of the speaker has become irrelevant to the assertion of their truth. Presumably for such reasons,
Kalchofner (2003) attributed a “factual epistemological stance” in the sense of Mushin (2001: 74–77) to -sang, which disassociates the speaker and her source of information from the statement that thus becomes [[51]] perfectly objective and unchallengeable. Examples (1), (2) and (3) illustrate the epistemic peculiarities of -sang:

(1) en zeerd üree öngör-söng jil
D.PROX chestnut 3–5-year-old.horse pass-PRF.P year
byanhonGr-iin naadam-d mön airagd-sang
PLACE-GEN feast-DAT same finish.1st.to.5th-EST.PST

‘[The chestnut horse of the herder Budgaliin Buyannemekh from Möst community in Khovd county has number 14.] This chestnut also took a prize last year at the Naadam of Bayankhongor.’

(2) hoyr=n ger bül bol-song
two=3.POSS yurt group become-EST.PST

‘[I’m living a happy life. (...) I have four daughters.] Two of them have formed families. ’

(3) A: bii ug-aas=a a hot-iin hūn B: hot-iin?
1SG root-ABL=RP town-GEN person town-GEN
A: tiin hoto-d töösöng
so town-DAT be.born-EST.PST

A: ‘I am a city dweller.’ B: ‘City dweller?’ A: ‘Yes, I was born in the city [i.e. Ulaanbaatar]. ’

In (1), informants hold that the sports commentator claims to have reliable knowledge. As a reporter, he might possibly use -jee to convey hearsay information, for what it’s worth. Some think that he claims to have learned about the event recently, which is unlikely here. In (2), the speaker couldn’t use (-ch)-laa or -jee because he didn’t just find out about it. In (3), too, informants reject alternative forms: -laa would indicate that the speaker has recallable, recent direct knowledge about her own birth; -jee could be used by an orphan that just found out about the circumstances of her birth. However, she would use -sang if this process of inferring had already been completed satisfactorily. -sang is thus obligatory when the speaker talks about her established knowledge.

Examples (4) to (7) are meant to show that -sang can refer to both distant and recent events, that it is independent of the speaker’s access to

---

6 At the same time, Kalchofner linked -jee to the imaginative (relating something as made up, probably due to the use of -jee in fairytales and some other narratives), reportive and inferential and -laa to the perceptual experience and private experience stances of Mushin’s heuristics.

7 For the transcription from Cyrillic and of spoken examples, see Brosig 2015 fn 6 and 9.
the event, and that [[52]] both perfective and perfect contexts are possible (while imperfective uses require additional marking that will be illustrated later).

(4) Yevrop-iin baildan_dagualagch Konstantinopl 1453 on-d Europe GEN conqueror Constantinople 1453 year-DAT Turk-iig ezel-sen-er dundad zuun tögs-sön Turks ACC occupy-PRF.P-INS middle century end-EST.PST ‘When the conqueror of Europe, Constantinople, subjugated the Turks in 1453, the Middle Ages ended.’

(5) teg-ee=l ter shaar ing-eed dee-sh=ee yav-aad do.so-C D.DIST balloon do.like.this-C up-ALL=RP go-C ög-söng. teg-eed joohom bai-jai-snaa buc-aad give-EST.PST do.so-C a.bit be-PROG-C just.after return-C minii ömön hür-ee d ir-seng teg-ee=l 1SG GEN before reach-C come-EST.PST do.so-C=FOC ing-ee=l minii ömön togt-jai-snaa do.like.this-C=FOC 1SG GEN before keep.still-PROG-C just.after bii haring ter-ii į av-h=uu, bai-h=uu ge-j 1SG but D.DIST-ACC take-FUT.P=Q let.be-FUT.P=Q COMP-C bod-jii-song. teg-eed ter dee-sh=ee yav-sang think-PROG-EST.PST do.so-C D.DIST up-ALL=RP go-EST.PST ‘Then, that balloon went up. Then, after a short while, it came back and came before me. Then, while [the balloon] was standing still before me, I wondered whether I should take it or leave it. Then it went up.’ [[53]]

(6) teg-ee=l tus-tus-iin-h=aa ajil~ajil-d do.so-C=FOC particular~DISTR-GEN-NOM=RP work-DISTR-DAT yav-cgaa-sang

\[\text{This is compatible with Binnick’s (2012) claim of there being a deictic difference between }\text{-sang/}-v\text{ and }\text{-laa/-jee: the latter are “deictic tenses” that “directly relate the time of the “eventuality (...) to the time of the speech act,” while the former are “anaphoric tenses” that have a deictic center “mediated by a reference time” (p. 102). He links this to the difference between characterizing a situation vs. an occurrence (p. 144) so that deictic tenses must be used at the beginning of narrative texts unless the sentence contains a very concrete time adverbial (p. 162–164), while anaphoric tenses require the existence of an established point in time to which they can relate their event as posterior (p. 104–107, 209). On the basis of my corpus, the very last claim can neither be confirmed nor rejected. It is quite likely, though, that these aspectual traits which are mentioned at several places of this paper exert an influence on the precise dividing lines between the usage of different evidential markers.}\]

\[\text{This is compatible with Binnick’s (2012) claim of there being a deictic difference between }\text{-sang/}-v\text{ and }\text{-laa/-jee: the latter are “deictic tenses” that “directly relate the time of the “eventuality (...) to the time of the speech act,” while the former are “anaphoric tenses” that have a deictic center “mediated by a reference time” (p. 102). He links this to the difference between characterizing a situation vs. an occurrence (p. 144) so that deictic tenses must be used at the beginning of narrative texts unless the sentence contains a very concrete time adverbial (p. 162–164), while anaphoric tenses require the existence of an established point in time to which they can relate their event as posterior (p. 104–107, 209). On the basis of my corpus, the very last claim can neither be confirmed nor rejected. It is quite likely, though, that these aspectual traits which are mentioned at several places of this paper exert an influence on the precise dividing lines between the usage of different evidential markers.} \]

\[https://mn.wikipedia.org/wiki/Дундад_зуун, retrieved 2014-03-24. The translation is what native speakers are likely to understand. The author might have intended Turk as an accusative subject (which would render it historically accurate), but this is hard to arrive at without an accusative on Konstantinopl.\]
[After getting up in the morning, the guest asked why there were so few people around. He received the response that the others usually leave around 6 AM, and he concluded:] ‘So they have all gone to their respective work.’

(7)\textit{zaa yav-sang} well go/leave-EST.PST

‘[Well, let’s proceed to the first horses. Horses of 3–5 years are very fast, probably we won’t catch up with them.] Well, we are off.’

(4) refers to an event hundreds of years ago; (5) happened in the speaker’s youth; (6) happened a few hours before and (7) just moments ago.\textsuperscript{10} The speaker witnessed (5) and (7), but not (4) and (6). (5) is narrative, while (6) is compatible with a perfect interpretation. In the absence of specifications for conditioned exceptions, these examples refute Binnick’s (2012: 110) claim that -sang is a remote and possibly firsthand past.

There is one notable idiosyncratic use of Past Progressive forms. Prototypically, they are used to express simultaneity or to signal an incomplete process such as \textit{bod-jii-song} ‘[I] was thinking’ vs. \textit{bod-song} ‘[I made up my mind and] thought.’ However, they can also be used for events that are claimed to be completed and without explicitly relating the internal structure of the event to other events. Rather, reference to the internal event structure just helps to emphasize that the speaker underwent such an experience, irrespective of whether it specifically characterizes her or not. Stylistically, informants describe experiential uses of the Past Progressive as lively, colloquial and informal, to the extent that school teachers stigmatize against them. In (8) and (9), even though the structure of the actual properties of the sentences is close to punctual, no iterativity is implied. This explanation also accounts for the possible interpretation of (10) as complete, a fact that Hashimoto (1995) overinterpreted as evidence for a resultative interpretation of the Khalkha Progressive. Historically, this use might be quite ancient and go back to an actual resultative construction, namely to Middle Mongol \textit{-ju a-} (Brosig 2014a), with \textit{bai-} replacing its predecessor \textit{a-} in the course of historical development in all its uses. \textsuperscript{[54]}

\begin{verbatim}
(8)bi anh höödöö alt-(an)d yav-aad teg-eed B-d
1SG first.time countryside gold-DAT go-C do.so=C NAME-DAT

eemeg bögü hoyor belegl-eed avchar-ch ög-chii-s=iin
earring ring two present-C bring-C give-PROG-EST.PST=MC
\end{verbatim}

\textsuperscript{10} (7) is problematic with regard to establishedness. The speaker might have opted for it to avoid \textit{-laa} which in one-word-sentences would yield future (see discussion in Section 5) or \textit{-ch-laa} which would be past, but might possibly relativize control.
‘I for the first time went to the countryside [searching] for gold, then I brought with me an earring and a ring and presented them to B. [We still have them, she just doesn’t put them on.]’

(9) myanga-n cagaan hon ’oo malal-dag malchin cagaa
    thousand-AT white sheep-RP herd-HAB.P herder white
    hüühen ge-j namaig zaluu bai-had
    girl COMP-C 1SG.ACC young COP-C.when
    hel-jii-s=iim=aa
    say-PROG-EST.PST=MC=EMPH

‘When I was young, they [once/repeatedly] called me the white [light-skinned/pure] herder girl who herds a thousand white sheep.’

(10) bi tan-d neg zahia bichi-j bai-san
    1SG 2SG.HON-DAT one letter write-C COP-EST.PST

‘I have written a letter to you [once].’ (Hashimoto 1995, his translation)

5. Direct past -laa

The suffix -laa marks past events directly accessed by the speaker. This includes the sensory perception of external events, introspection into internal states, or events executed by the speaker herself. -laa is most often identified as proximal (Song 1997: 208; Nelson et al. 1998: 121–126; Binnick 2012: 108–111). However, it was shown that precise cut-off points don’t exist (Song 1997: 169–183) and that it can also refer to distant events that the speaker remembers clearly (Svantesson 1991: 193).11 As a solution, I suggest that the issue is not temporal proximity but immediate cognitive access to an episodic representation of the event either due to the recency of its perception or to the speaker’s ability to remember it in detail even after a long time has passed. This contrasts with merely remembering an abstracted, stripped, reified representation of the event in terms of some of its most salient features. Episodic long-term memory, in turn, can provide the deictic anchoring for events in time mentioned by Binnick (see Footnote 6 above).

In the corpus, recent and distal examples are both attested, but the latter are not numerous: [[55]]

(11) önöödör bi tanai-d hono-h ge-j
    today 1SG 2PL.GEN-DAT stay.overnight-FUT.P COMP-C
    zor’-j ir-lee
    aim-C come-DIR.PST

11 Other approaches can safely be disregarded. Explaining -laa as a perfect (Kalchofner 2003) disregards its narrative function, defining it as a marker of epistemic certainty, but -jee as a resultative (Jingan 2010: 105–108) disregards their aspecto-temporally similar distribution.
[A knocks at the door, B opens. The following conversation starts: A: How do you do? B: How do you do? A: Excuse me, I want to ask you one thing. B: Ok?] A: ‘Today I have come in order to stay the night at your place [, Mr. Ganbaatar]!’

(12) teg-eed ih_surguul-’aa tögs-löö shinjleh_uhaan-ii
do.so-c university-RP finish-DIR.PST science-GEN
akdem-d erdem_shinjilgeen-ii ajiltn-aar och-loo
academy-DAT research-GEN worker-INS go.to-DIR.PST
mikroskop shagaig-aad doloon jil suu-sööng
microscop peer-C seven year sit-EST.PST.EMPH
‘Then I graduated from university. I entered the Academy of Sciences as a scientist. Staring through microscopes, I sat there for seven years.’

(13) dûrûûn ohin gar-laa.
four girl emerge-DIR.PST
‘[In principle, I thought that I want to father a boy. But it didn’t happen.] Four girls were born.’

(11) was uttered within minutes of the speaker’s arrival at that place. An alternative -sang would be interpreted as completed, which leads some informants to reject it, while others take it as a fully established fact and thus more insistant on not being send away. The events referred to in (12) lie at least seven years in the past. Additionally, -laa here emphasizes the sequence of events, while -sang might just enumerate unrelated facts. From (13), informants don’t understand that the birth was recent, but instead perceive more emotional affectedness on the part of the speaker than would be expressed by the more aloof -sang.

While informants consistently describe -laa as being less remote than -sang, aspectually complex forms such as Progressives combined with -sang and especially -laa are usually perceived as being more remote than either on its own. The Direct Past Progressive is most saliently used for recollecting the details and inner structure of directly witnessed past events; in (14), the emphasis is on how he brought the bag from Oceania rather than on the origin of the bag. Temporal distance is only implicated, though, as recent past events that require reference to their temporal structure can be marked as Direct Past Progressive: in (15), both the loss of confidence as non-completed and as non-established subjective perception are crucial for reasons of politeness. [[56]]

(14) kualûa lumpa avstral-aas avchir-jii-laa.
PLACE PLACE-ABL bring-PROG-DIR.PST
‘[Well, this there is the rucksack of my smallest grandchild.] I was bringing it here from Kuala Lumpur, from Australia.’

(15) zaâ bayrallaa ügûû ge-e=l hel-ch-h=üü ge-ed
well thanks EX.NEG COMP-C=FOC say-BOU-FUT.P=Q COMP-C bas doto-r ev=giü=l bol-jii-loo shüü
also inside-LOC concord=EX.NEG=FOC become-PROG-DIR.PST DP
‘Well, thank you. I was starting to feel uncomfortable that you might say “No”.’

Alternatively to proposing a gesamtbedeutung for past tense uses of -laa and -jee as suggested here, it would also be possible to propose that these suffixes assume a different meaning within complex forms. However, -laa occurs only 22 times and -jee only 9 times in four different aspectual constructions. On the basis of so few tokens, the advantages and disadvantages of this alternative approach cannot be assessed.

-laa and -jee can rarely be exchanged, as this invariably changes the source of information. The relationship between -sang and -laa is more complicated. Reasons for which informants refuse to exchange them are related both to temporal distance and evidentiality:

(16)badraag=aa ald-laa shüü. _ badraa gar-aad
NAME=RP lose-DIR.PST DP NAME exit-C
yav-cn=uu
go-BOU.EST.PST=Q
‘We have lost our Badraa! _ Did Badraa go out?’

(17)nad deer=chen üs urg-aad ehel-ch-lee
1SG.DAT on=STC hair grow-C begin-BOU-DIR.PST
[‘Now what is happening with me?] On me, hair has begun to grow. [Under my armpits, hair has grown (uraG-ch-laa). On my balls here, hair has begun to grow (urg-aad ehel-lee).” About such things, children worry a lot.”]

(18)hümüüx yaa-j övd-d-iig chii odoo med-lee
person do.what-C hurt-HAB,P-ACC 2SG now recognize+know-DIR.PST
[After hitting X once, Y asked: “Does it hurt?” X said: “A bit.” Y hit X again and asked again: “Does it hurt?” Upon X answering “It hurts”, Y said:] “Now you have understood how people feel pain ...”

(19)manai aav say-han (...) _ öngör-söng.
1PL.GEN father recently-DIM pass-EST.PST
‘My father recently passed away.’

In (16) and (17), the assertors have just noticed a new state of affairs that they don’t properly understand, thus -sang cannot be used. In (18), -sang is possible but would express Y’s conviction, while -laa rather refers to Y’s impression from X’s reaction. This indicates that -laa with the second person can even be used for emphasizing [[5]] the speaker’s sensory experience when the addressee is aware of this experience herself (contra
Binnick 2012: 71–73). Soldiers and other people who receive orders likewise have to confirm them with forms such as med-lee ‘understood!’ or sons-loo ‘I heard it.’ Thus, -lāa cannot be replaced by -sang when it explicitly indicates the realization of new information by the speaker. On the other hand, even an event explicitly marked as (relatively) recent such as (19) cannot easily receive -lāa (as informants state) if the speaker is already familiar with it.

A special function of -lāa is to establish fictive situations. This might be interpreted as a form of direct access to a mental representation of an event just created in the speaker’s mind and thus not yet firmly entrenched into it. In (20), the speaker doesn’t refer to a concrete event, but just invents it. Similarly, -lāa is used in counterfactual contexts, such as (21), for situations that might have occurred in the past under certain conditions but didn’t:

(20) jisheeq=n hoyor baigulG centavar-t or-loo shdee
example=3.POSS two organization NAME-DAT enter-DIR.PST DP
zaa (...) huv’-iin kampan or-hod tolboh herege-tei
well private-GEN company enter-C. when fee necessity-COM
shd
DP
’[I earn money through advertisement.] For example, two organizations are about to enter [my game show] Centaur. Well, if (...) a private company enters [the program], it must pay a fee.’

(21) chii bai-h=giig bai-sang bol yag un-laa (overheard)
2SG be.at-FUT.P=EX.NEG COP-PRF.P if exactly fall-DIR.PST
’If you hadn’t been there, I would just have fallen.’

-lāa is commonly used to refer to an imminent future for which visual or speaker-internal evidence exists (see Section 9). Nelson et al. (1998: 121–126, see also Chuluu 1995: 97–100, 1998: 155–156, 169) tried to predict this variation in time reference from aktionsart: -lāa can have near future interpretations with agentive accomplishments (id- ‘eat’ plus specific object); agentive nonconclusive predicates (togl- ‘play’); stative/stance (am’dar- ‘live’); punctuals (naitaa- ‘sneeze’); movement and weather predications. The copular auxiliary bai-, change-of-state verbs (urga- ‘grow’), and verbs with partially affected entities only yield recent past. Some punctual predications can only mean past, as their imminence cannot be indicated by firsthand evidence in the present. The copular auxiliary, weather predications and mental state verbs (ergelz- ‘doubt,’ tan’- ‘recognize[+know]’) may refer to past events that hold up to the present time, e.g. Neg cag boroo or-loo one hour rain come-DIR.PST ‘It was/had been raining [for] an hour.’ Statives such as am’dar- are unable to highlight the beginning of an event and thus exclude the present.
This argument is mostly sound but requires some modifications: stative verbs don’t easily allow for future reference: Bold Ulaanbaatar-t (arvan jil) amˈdar-laa [[58]] ‘Bold lived in Ulaanbaatar (for ten years)’ cannot have the future interpretation ‘Bold is about to start living in Ulaanbaatar’ claimed by Nelson et al. (1998: 122) even if the temporal adverbial is absent. The same holds for -laa in combination with analytic aspectual forms. Informants don’t agree on the interaction between -laa and directed activities (cf. Croft 2012), which lexicalize directed change without a concrete end point. While Nogoo ötgön urga-laa ‘The grass grew thick’ (Nelson et al. 1998: 123) is not accepted as future, some informants do accept nogoo=ch udalgii ötgön urga-laa green=FOC soon thick grow-DIR.PST ‘even the grass will grow thick soon’ or margaash dulaar-laa ‘it will get warm tomorrow.’ Finally, it is not the partially affected object in Bat jaahan mah id-lee NAME a.little meat eat-DIR.PST ‘Bat ate some meat’ (Nelson et al. 1998: 125) that yields past, cf. the future interpretation of Bi odoo jaahan mah id-lee ‘I’ll eat some meat now’ with the adverbial odoo ‘now.’

For non-stative predications, there seem to be more crucial issues than the aktionsart. While only 10% of 120 uses of -laa in my corpus refer to the future, informants who were asked to replace other past markers with -laa frequently reported a change to future meaning. Past reference is usually clear from past tense adverbials, dependent clauses (past, habitual, goal), past sentence chains, the presence of the suffix -ch-, and overall context. Future uses are mostly found in relatively short sentences such as (40) in Section 9 and with future adverbials. In some cases it is irrelevant whether or not an accomplishment has already come about. If a child ‘has / is about to burst into tears’ as indicated by the verb form uil-laa, adult intervention is advisable in any case, so the speaker doesn’t need to disambiguate.

6. Indirect past -jee

The suffix -jee marks past indirect access to information via inference or hearsay (Svantesson 1991: 193; Song 1997; Binnick 2012). However, it is surprisingly infrequent (n=46–50), which needs to be explained. I propose that its hearsay use is highly restricted and that -jee is only usable when talking about information inferred from sensory evidence for the event (or possibly hearsay) that doesn’t yet belong to the established knowledge of the speaker.

First, all uses in the corpus seem to relate to inference. Even a small set of retellings I investigated is rather dominated by -sang, sometimes followed by the quotative verb ge-, and by the Present Perfect -sang bai-n. The Indirect Past -jee is widely used in journalistic broadcasts and texts, but this usage might be register-specific: [[59]]
(22) M. suman-d öchgdör oroi Shüüh-in shiidver place arrow-DAT yesterday evening court-GEN decision güıcetgel-in alban-ii ajiltan-g buuda-j, implementation-GEN service-GEN worker-ACC shoot-C am’ nas-iig=n’ horoo-son hereg gar-chee12 life year-ACC=3.POSS reduce-PRF.P case come.up-INDIR.PST

‘In M. municipality, the crime happened that an employee of the Service for implementing court decisions was shot to death yesterday evening.’

(22) is taken from a short report in an online newspaper, and the reporter who wrote it is not likely to have witnessed it or done any more research than to call the responsible police office in the distant rural community where it happened.

More commonly, -jee is used when drawing inferences. These can be trivial or complicated, but the reasoning is always recent:

(23) chii coor-oo=güü bai-n ge-seng
2SG be.pierced-INDIR.P=EX.NEG COP-DIR.PRS say-EST.PST
coor-cong bai-jii-jee, asGar-cha-j
be.pierced-BOU.PR.F COP-PROG-INDIR.PST spill-BOU-INDIR.PST
‘You said, “It isn’t punctured.” It was punctured alright, it has spilled out.’

(24) en ajil bol ündseng-d=ee shiide-gd-jee ge-j
D.PROX work TOP basis-DAT=RP decide-PASS-INDIR.PST COMP-C
siül-in Gurav dörvö-n jil-iin naadm-aas hoi-sh
tail-GEN three four-AT year-GEN celebration-ABL after-PROS
hara-gda-j bai-n=aa see-PASS-C COP-DIR.PRS=EMPH
‘After the Naadam celebrations of the last three or four years, it is apparent that this work has basically been decided upon.’

(25) zaa ing-eed ulaambaatar hoto-d üül ajillagaa yav-uu-j
well do.so-C PLACE town-DAT action activity go-CAUS-C
baig-aa barag alban bus yanhni gazar
COP-INDIR.P almost official ID.NEG brothel
shin-eer neme-gd-lee ge-j ard tümen-d=ee hel-hed odoo
new-INS add-PASS-DIR.PST COMP-C people-DAT=RP say-C.when now
emzegl-meer bai-n=aa. neg-düger horooll-in
be.sensitive-WISH.P COP-DIR.PRS=EMPH one-ORD district-GEN
ar-iin zam deer bol nege-n tom yanhni gazar
north-GEN way on TOP one-AT large brothel [[60]]
baiguula-gd-jee ge-j

found-PASS-INDIR.PST  COMP-C
‘Well, now, it’s a sensitive thing for me to tell the Mongolian people that new, certainly unofficial brothels operating in Ulaanbaatar were added. That on the Northern Road in the First District, one large brothel has been established.’

In (23), a past resultant state (having been pierced) and a punctual past event (the spilling) are inferred from the trivial result of the punctual event (spilled matter). Clearly, the “lack of commitment” that Song (1997: 183–193) associated with -jee is not present here (though using -jee in (1) and (2) would lead to such an implicature). Rather, it is understood that this state of affairs was discovered just now. In (24), the speaker takes recourse to past evidence for a conclusion he draws at the moment of speech. Using -sang instead would have indicated an established, entirely unproblematic fact, and -l aa would have claimed an interaction with the decision-makers. In (25), the reporter recently discovered with his own eyes that new brothels exist there, but must (trivially) infer that they were established.

The co-occurrence of -jee with the first person is limited. It is possible if an event at a different place affected the speaker as in (26), if the realization of some parts of the situation is deferred (cf. Jingan 2010: 87) as in (27) where only the adverbial constitutes new information, or with events that the speaker had forgotten such as (28). It is not uncontrollable verbs as such that matter (Binnick 2012: 66–70).

(26) beltegl=ee hii-j bai-had sonin deer bii
   training=RP do-C COP-when.C newspaper on 1SG
gar-ch-jee
   go.out-BOU-INDIR.PST
   ‘While doing my training [there], I appeared in the newspaper.’

(27) bi tedn-ii=i nege-n adil bodo-j yav-jee.
   1SG 3PL-GEN=FOC one-AT like think-C go-INDIR.PST
   ‘I thought as if I was one of them.’ (Binnick 2012: 69, internet)

(28) bi baga-d=aa ene nom-iig unsh-jee
   1SG small-DAT=RP D.PROX book-ACC read-INDIR.PST
   ‘I must have read this book when I was small.’ (Binnick 2012: 68, constructed)

-jee is perfectly compatible with recent events, as shown by (25) and (29), which refutes accounts that focus on -jee as a marker of remoteness such as Hashimoto (1993a) and Binnick (2012).

(29) A: oo manai ohin unt-chi-j
    INTERJ 1PL.EXCL.GEN girl sleep-BOU-INDIR.PST [[61]]
B: unt-chi-j=uu? B: th
It also becomes clear from examples like (22), (26), (27) and (28) that -jee doesn’t have any resultative meaning as claimed by Kalchofner (1999, 2003). Jingan’s (2010: 86–89, 107) attempt to ascribe a resultative basic meaning to notions such as resultant states, explanations, surprise, conjecture and reinterpretation is not convincing either. While at least trivial inferences can be ascribed to all of his examples, explanatory force and surprise are indicated by other contextual devices, and ascribing a resultative interpretation (instead of connotation) to höld-jee ‘froze’ (which is actually already coded by forms like the Direct Perfect höld-söng bai-n and the Direct Resultative-Continuative höld-ööd bai-n (cf. Brosig 2015a)) seems inappropriate. There might have been such a resultative in Proto-Mongolic (Brosig 2014a), but as it had already turned into a perfective past for indirectly accessed events by the 13th century (Street 2009), an analysis as a contemporary resultative would be diachronically improbable (cf. Bybee, Perkins and Pagliuca 1994: 105; Tatevosov 2001).

One issue that cannot be fully addressed on the basis of my data is the role of non-visual sensory evidence. It seems that prototypical, easily identified sounds as in nohoi huc-laa ‘the dog barked’ would most commonly be referred to with the Direct Past even if the agent is not in sight. In cases where an additional element of inference is involved, speakers have to choose the marker that matches the evidence they perceive as primary. Informants agree that a person sitting in her room, upon hearing a noise from the street, could say mashin-ii osol bol-chiho-j car-gen accident become-BOU-INDIR.PST ‘an accident occurred.’ Apparently, the connection between the noise of an accident and the accident itself cannot be so straightforward as to give primacy to sensory access. In an arguably slightly more straightforward case, upon hearing the door bang, a speaker uttered to me gar-ch-laa ‘she left.’ Neither of us could actually see the person leaving, but as there were only the three of us in the flat, the sound could only be interpreted in one way.

7. -v

The literature primarily describes -v as a neutral past suffix as used in the storyline of written texts. No such usage is attested in my spoken corpus, so no review is necessary. The use in performatives mentioned by Svantesson (1991: 192) [[62]] mainly occurs in written contexts like the section hüchin=güi-d too-co-ν validity=EX.NEG-DAT count-MIR.PST ‘declared void’ in classified ad newspapers whereby one can declare lost legal documents to
be invalid. In spoken narratives, -v seems to be used to express that a non-recent event surprised the narrator when it happened and presumably still contravenes her conception of the regular course of events. The narrative preceding (30) contains a number of devices commonly used to catch the attention of the listener. The last syllable of caca-v receives a markedly rising intonation. After the utterance, the addressee made an exclamation of surprise, to which the speaker answered bii=ch gaih-aad ‘It surprised/surprises me, too.’ (31) shows the optional variant -v=aa.

(30) teg-eed saihan suu-jii-san=chin genthen
   do.so-c nice sit.down+sit-PROG-PRF.P=STC suddenly
   minii urt tal-aas neg nōhōr _ nūür-līuū us
   1SG.GEN front side-ABL one comrade face-ALL water
caca-v
   sprinkle-MIR.PST
   ‘[I want to relate an interesting story. In April, you see, I was travelling in Laos and Thailand, just at their New Year’s time. Laos, Cambodia, Thailand and Vietnam have one single moon calendar. That was quite strange, as we celebrate our New Year already in February. Once, we were sitting at a banquet at New Year’s Eve - what kind of expensive food do they have - they had put roasted chicken, egg and various kinds of food on the table.] And then, while we were sitting there pleasantly, suddenly one guy in front of me sprinkled water towards my face!’

(31) negen saihan nar-tai öldör bi naiz-uud-tai(g)-aa
    one nice sun-POSS day 1SG friend-PL-COM-RP
cug bai-j bai-laa, genet minii shūd
    together COP-C COP-PST.DIR suddenly 1SG.GEN tooth
una-v=aa, teg-snee genet ahiad neg
    fall-MIR.PST.EMPH do.so-c.just.after suddenly again one
una-v=aa (internet)
    fall-MIR.PST.EMPH
    ‘At a nice, sunny day I was being together with my friends. Suddenly, one of my teeth fell out! And immediately afterwards, another one suddenly fell out!’

8. Questions

Due to limitations of space, I will be extremely brief here. -sn=uu (i.e. -sang=uu) is used for neutral polar questions. The form -s=iim=uu (-sang yum=uu in formal language) presumes a (usually positive) answer. Interrogative -v has been analyzed as referring to recent events (Binnick 2012: 95–102), but this isn’t supported by my corpus. Chuluu (1998: 136–137) proposes that it presupposes a certain state of affairs and is more polite.
My sparse data seems to support this, but then it isn’t clear how [63] it differs from -s=-iim=uu. Both -sang and -s=-iin are used on their own in content questions, but informants who reflect about it generally dislike -sang, possibly because the written language requires the presence of the question particle be. This being said, I will refrain from a proper analysis of these forms and restrict myself to the evidential qualities of -l=uu and -jee.

The shortened interrogative form -l=uu is used if the speaker cannot remember an event that she witnessed (Kullmann and Tserenpil 1996: 187) in self-directed questions, e.g. (32). In content questions, the equivalent form is -laa:

(32)bii soiz=oo shiree-n deer tav'-l=uu?
1SG scissor=RP table-AT on put-DIR.PST=Q
‘Did I put the scissors on the table?’ (elicited)

(33)hamg-iin siül-d hüüg=ee hezee zagan-laa? odoo
all-GEN end-DAT son-RP when scold-DIR.PST now
mart-ch-jee
forget-BOU-INDIR.PST
‘When did I scold my son the last time? I’ve forgotten by now.’

The speakers don’t need to have actual firsthand knowledge; it is sufficient if they knew a fact once and cannot remember it. Speaker B in (34) never witnessed the event but used to know about it. While (35) was clearly witnessed (as the speaker tries to recall details of the process), informants indicated that this wouldn’t be clear without the Progressive (when the speaker would only try to recall the event as such).

(34)A: bii neg sar-ii=n sunga-j ög-sii=shd.
1SG one month-ACC=RP prolong-ACC give-EST.PST=DP
B: dahí-j sunG-l=uu?
repeat-C prolong-DIR.PST=Q
A: dahí-j sunga-j ög-sii=sh=d tiin
repeat-C prolong-C give-EST.PST=MC=DP so

(35)taa neg songuul’-d örsöld-jii-l=üü,
2SG.HON one election-DAT compete-PROG-DIR.PST=Q
yaa-laa
do what-DIR.PST
‘You were taking part in one election, how was it/what did you do?’

Kullmann and Tserenpil (1996: 187–188) claim that the long form -l=uu(y)=uu is used for questions about an imminent future. But this, as such, is definitely not its main use. The form is not attested in my corpus at all,
but most examples on the internet indicate that it expresses a (usually strongly) incredulous, ironical attitude towards [[64]] a claimed event. This is often, but by no means always combined with an (often somewhat restrictedly) contemptuous stance towards its assertor.

(36) tom bank-iig bathüü huurmach bichig barimt-aar
   big bank-ACC NAME false paper evidence-INS
   huur-laa=yyu
   deceive-DIR.PST=Q
   ‘Can it really have been Batkhüü who deceived the big banks with forged documents? [If they can grab Sugar and Ganzorig and interrogate them, they will no doubt reveal the big guys.]’

Most such uses are found in comments on articles, not within the articles themselves, and all informants consider this usage rather informal, roughly comparable to more formal forms such as –sang ge-j=üü (see below). Originally, the idea behind this use might have been to rhetorically claim that the addressee can improbably have any direct knowledge of such an unlikely event.

There is weak evidence for a non-ironic use, possibly restricted to some verbs, intonation patterns and an older speaker generation. For (37), the only such example I could find online, three informants (average age 35) suggested that Non-Established Present Prospective forms such as -h ge-j bai-n=uu are semantically similar, with differences pertaining either to aspect (actual future vs. prospective) or politeness (informal/intimate vs. polite). For three other informants (average age 22), however, the writer still seems to suggest that the addressee cannot be serious about going to China without learning the language first and should thus learn from the writer first.

(37) hyatat(a)-d sur-haar yav-laa=yyu?13
   China-DAT learn-CVB.CNS go-DIR.PST=Q
   ‘Are you about to go to China to study? [Doctoral student with degree from China trains children who want to study there.]’

Questions based on -j=üü ask for inferential evidence. Judging from the few corpus examples such as (38), this evidence lies with the addressee. This example also indicates that a speaker can mark a statement as based on indirect evidence at least in cases when the available direct evidence has been challenged.

(38) Observing an unexpected event, X states: uu-ch-la, ‘[BB] drank [the vodka]!’ providing firsthand evidence. Y is skeptical and asks: uu-ch-

j=uu? ‘Did he drink it?’ asking whether there is any circumstantial evidence. X checks that the glass is indeed empty and then draws the inference: uu-chih-j. ‘He drank it.’ [[65]]

A different type of question is based on the form in -j=üü of the verb ge- ‘say’ and used to verify statements or intentions (Binnick 2012: 93; cf. Jingan 2010: 103–104), as illustrated in (39):

(39)A: za en gadaad pasport=ch nootoi shüü well D.PROX foreign passport=FOC fishy like
B: hulih=ch bai-j magadgii ge-j=üü fake=FOC be-CVB maybe say-INDIR.PST=Q
A: ‘Well, this foreign passport looks suspect.’ B: ‘Do you mean that it might even be fake?’

Diachronically, the person who infers here is the speaker, and the original meaning might be paraphrased in a literal way as ‘I infer that you are saying PROPOSITION, are you?’ Questions based on inferentials in Middle Mongol were speaker-centered (Street 2009), so this use might be fossilized if it is old enough.

9. Futures

In the preceding sections, past tense uses of -laa, -v, -jee and -sang were discussed. However, all of these suffixes also have a future use that by and large reflects their evidential properties. -laa covers intentional and predicted imminent future events that can be foreseen from speaker-internal or visual evidence. -v is used for future events that might come as an unpleasant surprise for the addressee. These two suffixes are relatively frequent in future contexts and might be argued to exhibit fuzzy temporal semantics. -jee and -sang are very rare in future contexts and then refer to scheduled or predicted events that can be inferred or taken for granted, respectively. In all non-prospective future contexts, the past suffixes compete with the Potential -n, and in prospective contexts they compete with a specialized Prospective construction (see Brosig 2015a).

The only past suffix very frequently found in future contexts is -laa, and this use is always noted in the literature. Song (1997: 192–193) argues that -laa is better suited for expressing an inevitable imminent future event due to the commitment entailed by firsthand evidentiality. However, the existence of circumstantial visual (sensory?) or speaker-internal evidence for the imminence of the event could explain this use, too. With first-person subjects, what is expressed are usually actual decisions of the speaker. The Direct Prospective form -h gej bain would not express intentionality, while
Potential -n could not convey that the speaker intends to act immediately.[[66]]

(40) zaa tan-iig en gal_toogoon-d=chon üld-ee-lee
   well 2SG-ACC D.PROX kitchen-DAT=STC leave-CAUS-DIR.PST
   bii ger_bül-eer hes-lee
   1SG family-INS wander.around-DIR.PST
   ‘Well, I’ll leave you behind in the kitchen. I’ll walk around a bit in the house.’

(41) za za bii zövshöör-löö.
   well well 1SG accept-DIR.PST
   [Father: ‘I want to marry you to the girl of my choice.’ Son: ‘I want to marry the girl of my own choice.’ Father: ‘But that one is the daughter of Bill Gates.’ Son: ‘Oh, is that so?’] Well well, I accept.  

(42) zaa urag-sh=aa yav-laa
   well forward-ALL=RP go-DIR.PST
   ‘Well, [the race horse] is about to make/has just made progress.’

In (40), the speaker announces his decision to directly commit an action. This is somewhat similar to the performative use (noted by Kalchofner 2003) as in (41). Due to lack of data, little can be said about its particular properties vis-à-vis other performative devices such as the Direct Present Progressive and Adhortative. In (42), it doesn’t make an important difference whether the horse is only about to make headway or has already improved its position among the racing horses. Such temporal underdeterminedness is quite characteristic of Khalkha live reports.

The corpus contains one example where a future event in -laa depends on a condition, and it can even be used together with a concrete time measure adverbial:

(43) joohon tata-s-hiig-eed yava-h=güi bol odoo zam-d=aa
   a.little pull-?-do-C go-FUT.P=EX.NEG if now way-DAT=RP
   or-loo
   enter-DIR.PST
   ‘If [the race horse] doesn’t fall back a little, it’ll enter the way now.’

(44) odoo ugasaa yag guch döchö-n minut-iin daraa
   now in.any.event exactly 30 40-AT minute-GEN after
   och-loo=shdee
   go.to-DIR.PST=DP
   ‘Now I’ll arrive after exactly 30–40 minutes anyway.’

---

14 [http://ganaas.blogspot.se/2007_05_01_archive.html](http://ganaas.blogspot.se/2007_05_01_archive.html), retrieved 2014-06-01, first part of a joke
If -v refers to future events, it expresses a preventive meaning (cf. Song 1997: 162–163): the speaker issues a warning that a certain event might occur, arguably imagining the adverse event as if already completed in her mind (Jingan 2010: 106). The presence of a long vowel is common but (contra the contemplation of Chuluu [67] 1998: 138) not obligatory: its absence is perceived as insistent and commanding, while its presence is more polite (cf. Brosig 2015a on -n=aa) and, in this case, apparently the more neutral variant. Pure apprehensive meanings would be expressed through different constructions (see Brosig 2005b: 120-122).

(45) like dar-h=aa martv=aa15
    like press-FUT.P=RP forget-MIR.PST=EMPH
    ‘Don’t forget to press [the] “like” [button].’ (heading to an online video)

(46) neg joohong temcel-memcel bolhoor hööy chii
    one little contest–DOWNPLAYING because hey 2SG
    iih-chi-v=ee al-ul-ch-v=aa (...) ge-j
    die-BOU-MIR.PST=EMPH kill-CAUS-BOU-MIR.PST=EMPH COMP-C
    bod-dog boluu
    think-HAB,P MP
    ‘Just because there’s one small contest, might they possibly be thinking:
    “Watch out, you might die! You might get yourself killed!”

In (45), the speaker reminds the addressee to “like” a certain online post. One could have translated this as ‘You might forget to press “like”,’ but martvaa is fairly idiomatic. Overall, -v in its preventive use seems to combine with verb stems that already indicate an event that is likely to be evaluated negatively, as holds for iih- ‘die’ and aluul- ‘get killed’ in (46).

For -jee and -sang, future uses seem to be extremely rare, and neither is attested in the corpus. The former is mentioned by Byambasan et al. (1987: 167–168, source of (47)) and Jingan (2010), while the second is not attested in the literature at all.

(47) Ta=ch mongol uls-iin bayar naadam-d
    2SG.HON=FOC Mongolian state-GEN celebration feast-DAT
    orolc-jee take.part-INDIR.PST
    ‘You will certainly take part in the national celebration of Mongolia.’

(48) naada-h=chin evider-ch=dee! (overheard)
    D.MED-NOM=2.POSS get.damaged-INDIR.PST=DP

15http://www.cardflourishes.co.uk/video/_ecg7_thqA8&feature=youtube_gdata_player, retrieved 2014-03-27
In (47) and (48), the perfective evidential meaning of -jee is apparently dissociated from the past meaning and turned into a certain inference about the future given a certain course of events. Jingan (2010: 107) describes this use as a highly likely future result given a certain present state as a premise. In (49), the only unambiguous future use of -sang that I observed in interaction, informants tend to assume a certain development beyond the speaker’s influence and, probably, liking. In the actual conversational context, though, the writer rather seemed relieved that she wouldn’t have to wait anymore. This leaves us with what is perhaps best described as an authoritative prediction on the part of the writer.

Finally, there is the question of how to explain discontinuative tense uses of past suffixes in future contexts. Given the data at hand, the question can only be put synchronically. In the case of -v, there is a component of surprise common to both future (for the addressee) and past uses (for the speaker). The surprise is dissociated from the speaker in some way, in the future use by presupposing it in the addressee and in the past use by ascribing it to the speaker especially at an earlier point in time. Given that the neutral past use still found in writing is absent from spoken language, this dissociated surprise might be taken as the core meaning of -v. One would yet have to assess how its underdescribed, but not infrequent interrogative uses fit into the picture.

In the case of all four suffixes, one might wonder whether the notion of pastness has been replaced by the notion of perfectivity. Perfectivity tends to be incompatible with the present, to which these forms indeed cannot refer. The use with complex aspectual forms and stative verbs could still be understood as delimiting these at some point in the past. For -laa, temporal semantics might actually have been weakened by the extension of its future uses: it is understood as future with non-stative aktionsart if none of the many possible contextual clues to the contrary is present. On the other hand, -jee and -sang presumably have to be coerced into future uses by a small set of unknown contextual factors and, due to their low frequency, might be semantically more specific than suggested here. Diachronically, it wouldn’t be unexpected if particularly -sang, which possibly was a perfect in Middle Mongol (Brosig 2014a) and might still be somewhat less suitable for situating historical events at a concrete point in time, is hard to convert into a tense-neutral perfective. The difference between pastness and tense-neutral perfectivity is thus gradual, shifting context by context and
construction by construction. Decontextualized one-word sentences containing -jee and -sang are firmly on the past side of this divide, as are the great majority of constructions containing them as well as the majority of constructions containing -v and -laa. Thus, reanalyzing these four suffixes as temporally neutral or, worse, reanalyzing Khalkha as devoid of actual tense would throw out the baby with the bath water. Only for -v, which has already lost most of the semantic space it once occupied and is on the verge of dropping out [69] of the past/perfective paradigm entirely, dispensing with a tense-based analysis might almost be possible. A similar case can be found in Khorchin (Brosig 2014b) where the cognate of -laa seems to have lost most of its disposition to form part of complex aspectual forms and, with the loss of evidentiality, might indeed have become an underdetermined marker of temporal proximity in the past or future.

10. Conclusions

The existence of two evidentially marked and one unclear form among the past tense suffixes of Mongolian has been a well-established fact, at least for specialists. Yet, existing descriptions failed to capture the conditions under which a certain form can or must be chosen. Relying on existing research, a non-native would overgeneralize the use of any of these past forms. In this chapter, I proposed an analysis that allows restricting the use of -sang to established knowledge and -laa and -jee to (by and large) direct respective indirect non-established knowledge. To the extent they are known, I discussed particularities such as an imaginative use of -laa that constitute related constructions with their own conditions of usage. Overall, my analysis provides more restrictive definitions of the Khalkha past tense forms than previous work. While the corpus is too small to contain all particular uses common to the community of speakers and the precise conditions of use still require refinement, I hope that the categories used for analysis will turn out to be descriptively more adequate than those suggested in previous analyses.

The issue of whether or not -sang can actually be understood as an evidentially “neutral” form remains a difficult question. Informants have suggested a certain level of speaker confidence when using -sang, which is, at any rate, not given with -jee. I would suggest that there is no confidence implied by -laa either, as the speaker is just referring to her own perception as the source of the information. Thus, for both -laa and -jee the factuality or reality of a proposition is relativized in terms of the reliability that the addressee might attach to firsthand and inferred/secondhand evidence (cf. Izvorski 1997). When -sang is permissible in the same contexts as -laa and -jee, it would by paradigmatic contrast implicate an unhedged claim to the validity of the proposition. On the other hand, the form seems too frequent
to express any “epistemic support” (Boye 2012) for a proposition on its own. But the speaker must (for some reason) have accepted a proposition in -sang, perhaps often after consideration, which might be taken as a clue by the addressee that she, too, might have some reason to trust this information. “Factuality” might thus indeed be linked to an epistemic value, but more in the eye of the beholder than for the speaker. [[70]]

Areally, the evidentiality system of Khalkha Mongolian reflects Tibetan systems such as Ladakhi where assimilated, fully asserted authoritative knowledge expressed by the auxiliary yod contrasts with non-assimilated, seemingly non-authoritative knowledge within which immediate perception is expressed by hdug and inference is expressed by special inferential markers, while hearsay is mostly expressed by verba dicendi (Zeisler 2017) - like Khalkha ge-. The major difference is that authority and participation are much less significant in Khalkha, so that -laa can easily be used for events that the assertor undertook just recently and which are not yet well established mentally, and for decisions that the assertor is just taking, while hdug is not easily used in contexts that are controlled by the assertor.

Like Tibetan, Khalkha cannot be exhaustively described by the notions suggested by Aikhenvald (2004), who assumes visual, other sensory, inferential, assumptive, hearsay and quotative information sources, or by the improved system of Plungian (2011: 33–37), which adds participatory and “endophoric” (i.e. speaker-internal) forms of direct personal access to information. A model that can capture the Khalkha system is suggested by Tournadre (2008: 297–298), though, who presents four different parameters that interact in Tibetan evidentiality systems: source of information, access to information, field of knowledge, and time of acquisition. He thus understands “source” in a more restrictive sense so that it can either be the speaker herself or somebody else, while access to information is possible either via the senses [of the speaker] or via inference. Furthermore, he separates “endophatic” access to the speaker’s internal state as a form of sensory evidence from participatory (“personal”) evidence, which, if defined as the voluntary (Haller 2004) or active, conscious participation (Rule 1977: 71) of an assertor (i.e. the speaker in declaratives and the addressee in questions, cf. Creissels 2008: 12), indeed goes beyond the notion of access to information favored by Plungian. The difference between -laa vs. -jee (and present -n vs. -aa) can be captured by both models, though in a slightly different fashion: In Plungian’s model, -laa would be neatly

16 Using the notion **MODE OF ACCESS** instead of **SOURCE OF INFORMATION** was already argued for by Michael (2006: Section 4).
17 It is maybe important to point out that Tournadre doesn’t claim that all these parameters belong to a domain of evidentiality, an idea against which Zeisler (2017) explicitly cautions.
identified with direct personal and -jee with indirect evidence. In Tournadre’s model, one would have to define -laa as marking sensory evidence of the speaker, while -jee is used for evidence that is either non-sensory or, if a quotative use indeed survives in spoken Khalkha, the evidence of somebody else. The most crucial parameter for Khalkha, though, is the “time of acquisition”, which differentiates between newly acquired (-laa/-jee) and assimilated information (-sang), which is specific to Tournadre’s model. However, I interpret this parameter in terms of the consolidation of information in memory, [[71]] which Dahl (2013) links to evidentiality and remoteness distinctions, rather than in terms of the absolute remoteness of the time when the information was acquired. Common knowledge, which Rule (1977: 71) and Oswalt (1986: 36) call “factual,” then naturally groups with assimilated information. Spoken past declarative -v is only retained for evidentially neutral, non-new, but nevertheless not fully assimilated information, i.e. a small semantic area that none of the other three past suffixes can properly refer to.

There are several questions left for further research. First, there are a number of registers other than free conversation, both in oral and written language, and their conventions influence each other. Analyzing these registers would substantially increase our knowledge about different uses of past suffixes in Khalkha. Moreover, the corpus materials didn’t allow a proper exploration of how senses other than vision are treated within the evidential system. As devices other than the past suffixes were not dealt with here, the expression of inference (-jee vs. -sang bai-n) and hearsay or the way speakers choose their epistemological stance also require further investigation. Finally, the form -s=iin, which is related to assertivity, and the suffix -ch-, which combines with all past suffixes and is related to notions such as boundary-orientation, displeasedness, lack of control, and surprise, have largely been ignored in this chapter due to lack of space. While their meaning doesn’t seem to be evidential, these forms are central to the notions that are regularly and paradigmatically expressed in Khalkha conversation.

**Acknowledgement**

This research was conducted in the years 2013-2014 when I was employed as a doctoral student by Stockholm University, and its results were revised again in 2015. For seven months in 2013, the National University of Mongolia in Ulaanbaatar received me as a guest researcher, and my stay in Mongolia was funded by Sven och Dagmar Saléns stiftelsen with 53270 SEK. I am grateful to Bettina Zeisler, Dolgor Gunsetseg, Östen Dahl, Henrik Bergqvist, Lamont Antieau and Maria Koptjevskaia-Tamm for comments on the draft of this chapter, to all my Mongolian informants for
their enormous contribution, and to all who supported me during my stay in Mongolia and later on. [[72]]

**Abbreviations**

In the following, only glosses that differ from the Leipzig glossing rules are provided.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>attributive case (-n)</td>
</tr>
<tr>
<td>BOU</td>
<td>boundary-emphasizing aktionsart modifier</td>
</tr>
<tr>
<td>C</td>
<td>convert</td>
</tr>
<tr>
<td>CNS</td>
<td>consequence (for converbal suffix -haar)</td>
</tr>
<tr>
<td>D</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DIM</td>
<td>diminutive</td>
</tr>
<tr>
<td>DIR</td>
<td>immediate direct evidence (-n, -laa)</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>EST</td>
<td>established (-sang)</td>
</tr>
<tr>
<td>EX.NEG</td>
<td>existential negation</td>
</tr>
<tr>
<td>FOC</td>
<td>focus clitic</td>
</tr>
<tr>
<td>ID.NEG</td>
<td>identity negation</td>
</tr>
<tr>
<td>INDIR</td>
<td>immediate indirect evidence (-jee)</td>
</tr>
<tr>
<td>INTERJ</td>
<td>interjection</td>
</tr>
<tr>
<td>DP</td>
<td>discourse particle or clitic</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual (-dag)</td>
</tr>
<tr>
<td>MED</td>
<td>medial (close to addressee)</td>
</tr>
<tr>
<td>MC</td>
<td>modal clitic (=iin ~ yum)</td>
</tr>
<tr>
<td>MIR</td>
<td>mirative (-v)</td>
</tr>
<tr>
<td>MP</td>
<td>modal particle</td>
</tr>
<tr>
<td>ORD</td>
<td>ordinal number</td>
</tr>
<tr>
<td>P</td>
<td>participle</td>
</tr>
<tr>
<td>PROS</td>
<td>prosecutive case (-sh)</td>
</tr>
<tr>
<td>RP</td>
<td>reflexive-possessive clitic</td>
</tr>
<tr>
<td>STC</td>
<td>stance-marking clitic (cognate with =2POSS)</td>
</tr>
<tr>
<td>TOP</td>
<td>topic particle (bol)</td>
</tr>
<tr>
<td>X_Y</td>
<td>X and Y only receive one gloss, or two English words gloss one word in the source language.</td>
</tr>
<tr>
<td>X_ Y</td>
<td>After uttering X, the speaker paused.</td>
</tr>
</tbody>
</table>

**References**


Plungian, V. 2011. Types of verbal evidentiality marking. In G. Diewald, & E. Smirnova (Eds.), *Linguistic realization of evidentiality in European languages* (pp. 15–58). Berlin: De Gruyter Mouton.


